

Kick-off ANR Compass

Contribution MDLS

Pierre Kestener

CEA-Saclay, DSM, France
Maison de la Simulation

Meudon, 14 mars 2013





MAISON DE LA SIMULATION



- **Laboratoire de recherche pluridisciplinaire** autour de la simulation numérique
- **Unité de service** ouverte sur les communautés offrant notamment une expertise et une aide aux développements applicatifs HPC pour les machines GENCI¹ et PRACE²
- **Un pôle d'enseignement et d'animation scientifique en calcul intensif** (Master M2S, PRACE - PATC)

¹Grand Equipement National de Calcul Intensif

²Partnership for Advanced Computing in Europe



- **Activité de développement/portage GPU pour PRACE:** code DP (spectroscopie ...)
- **Formation continue - PATC**³
 - **Large couverture des thématiques HPC:** bases de programmation parallèle (MPI, OpenMP), les accélérateurs, outils d'analyse de performance, profiling/debug/tracing, environnements runtime, bibliothèques parallèles haut-niveau (algèbre linéaire: PETSC, MUMPs, Pastix, Scotch, ...)
 - Programmation GPU (décembre 2012)
 - OpenMP/OpenAcc (janvier 2013)
 - A venir: **workshop accélérateurs (GPU + Intel MIC)**, avec Rob Farber
- **Formation initiale**
 - ENSTA - 2ème année - IN203 - Parallélisme / GPU
 - ENSTA - 3ème année - B13-3 - Prog hybride et multi-cœurs

³PRACE Advanced Training Center

- Thèses en cours

- **Feng XING (2011):** Méthodes de décomposition de domaines sur GPU pour l'équation de Schrödinger
- **Langshi CHEN (2012):** Méthodes de Krylov parallèles hybrides asynchrones
- **Fan YE (2012):** Nouveaux algorithmes numériques pour l'utilisation efficace des architectures multi-coeurs et hétérogènes (résolution de grands systèmes linéaires)



- Implications

contributeur Task1 (1.4) - **Computing working group** : document de définition infrastructures hardware / software, méthodes pour le développement logiciel, ressources matérielles de calcul GPU

coordonnateur Task3 - **Infrastructures** : outils collaboratif de type Forge, site web COMPASS, maintenance, compatibilité MAOS

- Task 3.1: outils collaboratif: Redmine ou Renater (FusionForge libre pour les labos de recherche)

contributeur Task5.3 - **Tomography / Learn and Apply**: implantation GPU des algorithmes, tests, optimisation, benchmarks

contributeur Task5.5 - **Supervisor**: implantation GPU des algorithmes, tests, optimisation, benchmarks

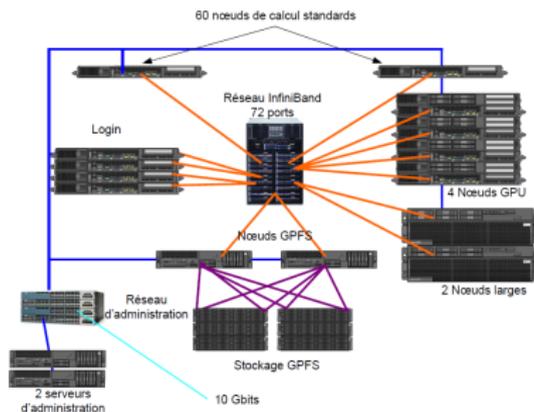


- **COMPASS: besoin en ressources de calcul GPU**

- **Proposition initiale:** chaque partenaire achète son équipement (GEPI, IPAG), éventuellement mutualisé dans GRID500 (via cluster ADONIS)
- **Usage:** simulations de différents instruments, AO PSF lib, XAO , benchmark de nouveaux algorithmes, tomographic reconstruction, pyramides, ...
- **Proposition retenue:** mutualisation des ressources pour les partenaires de COMPASS, **enveloppe budgétaire réduite (45k €)**

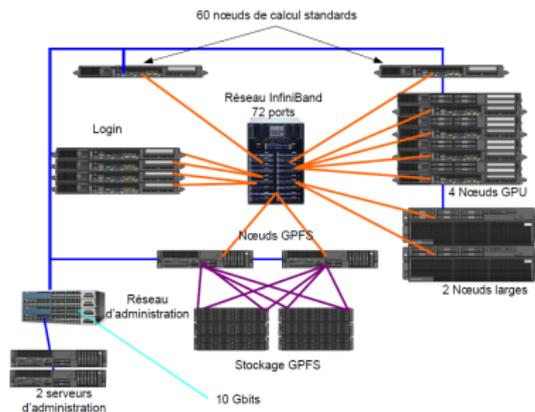


- Cluster Poincare de la MDLS - hébergé par l'IDRIS (Centre de calcul national du CNRS) ⁴
 - machine achetée par MDLS et CEA/IRFM (Cadarache) chez IBM
 - puissance de calcul crête CPU: ~ **30 TFlops**
 - mise en service en novembre 2012; 30 utilisateurs (50-100 nominal)
 - machine de **développement** mais **pas de production**;



⁴La machine principale est Turing (BlueGene/Q - 700 TFlops - 31ème au top500)

- Cluster Poincare de la MDLS - hébergé par l'IDRIS (Centre de calcul national du CNRS) ⁴
 - machine administrée par l'IDRIS (droits root non-délégués);
 - ouverture de compte comme sur leurs autres machines;
 - ports réseau fermé par défaut; ne peut pas servir à l'hébergement d'outils collaboratif



⁴La machine principale est Turing (BlueGene/Q - 700 TFlops - 31ème au top500)

- Cluster Poincare de la MDLS - hébergé par l'[IDRIS](#) (Centre de calcul national du CNRS) ⁴
 - **Description matérielle d'un noeud GPU**
 - 2 CPU (8 cœurs) Intel(R) Xeon(R) CPU E5-2670 @ 2.60GHz
 - 2 GPU NVIDIA K20 (CUDA hardware version 3.5) + 6GB RAM
 - mémoire RAM carte mère: 4GB/cœur = 48 GB par noeud
 - **Environnement logiciel**
 - OS: Redhat Server 6.3; installation *disk-less*
 - les 4 noeuds sont accessibles soit directement en interactif (via connection ssh) soit par le séquenceur de job (LoadLeveler).
 - **Projet COMPASS:** ajouter 4 noeuds supplémentaires dédiés (devis IBM: 11k €/ noeuds) **identiques** aux noeuds actuels; les noeuds de la config actuelle seront marginalement disponibles pour COMPASS.
 - Le budget actuel ne permet pas d'augmenter la capacité de stockage de Poincare; on peut envisager d'allouer un espace de l'ordre de 2 TB mais il faudra prévoir de transférer ce qui doit être conservé de manière pérenne

⁴La machine principale est Turing (BlueGene/Q - 700 TFlops - 31ème au top500)