

MACHINES HPC, GPU, PHI: PROMOUVOIR DE NOUVELLES HABITUDES DE CALCUL AU LPNHE ?

Tristan Beau, Frédéric Derue



MATÉRIEL AU LABORATOIRE

Le matériel orienté HPC est a priori à destination de tests et développements en vue d'une production sur un autre site.

- Grille de calcul (non orientée HPC)
- 2 machines à 16 cœurs physiques, pour lesquelles OpenMP ou MPI peuvent être mis en œuvre
- 1 machine support de 2 cartes GPU K2200 (640 cœurs chacune, 4 GB)
- 1 machine support de 2 cartes Intel Phi (4x60 cœurs chacune)

PREMIERS TESTS

Pour commencer, traitement du cas classique d'intégration MC pour l'estimation de π .

Nombre d'itération identique pour tous les tests ou bien normalisé pour les résultats avec une *.

Architecture	Software / compilation	Temps (secondes)
PC mono cœur	Python (pas d'optimisation)	976
PC mono cœur	Python (avec numba)	35
PC mono cœur	GCC (pas d'optimisation)	38
PC mono cœur	ICC (optimisation au plus)	29
PC 4 cœurs OpenMP	ICC	9.3
PC 4 cœurs OpenMP	GCC	7.7
PC 4 cœurs MPI	GCC ou ICC	~10
1xPhi 240 cœurs, OpenMP	Intel	4.0
Mono GPU	Nvcc (implémentation naïve)	6.3
Mono GPU	Nvcc (meilleure implémentation)	0.23
Mono GPU	Nvcc + thrust	0.055*
Bi-GPU	Nvcc (optimisation non optimale encore !...)	0.04*

APRÈS LES PREMIERS TESTS, QUESTIONS QUI RESTENT OUVERTES

- Avec les GPU
 - Optimiser les appels aux ressources de plusieurs cartes GPU ?
 - Projection de temps d'exécution pour nos projets ?
 - Possibilité d'utiliser des parties de codes standard ? Quelle architecture de code ?
- Avec les Phi
 - Communication entre les Phi ?
 - Optimisation du temps de calcul entre machine hôte et Phi ?
 - Utiliser les codes déjà existant ou bien ré-écriture partielle requise ?
- Plus généralement :
 - Quel type d'architecture cible pour nos problèmes ?
 - Quel rapport gain / temps de développement ?

MACHINES PARALLÈLE DE L'ICS

- Matériel (cf http://asclepios.dsi.upmc.fr/wp/?page_id=214&lang=fr)
 - Machine MESU Alpha (1000 cœurs, OpenMP) et MESU Beta (1600 cœurs, MPI, et des GPU)
- Tests
 - Sur code séquentiel MemTk parallélisé avec MPI.
- Principal intérêt
 - disponibilité instantannée de la machine, nombre de cœurs important
- Limitations effectives ou potentielles
 - Structurelle (accès aux données)
 - Logicielle (e.g. absence de CVMFS fonctionnel)
 - Matérielle (I/O limités, GPU non accessibles pour l'instant)

CAS D'UN CODE MEM POUR LA MESURE DE LA MASSE DU QUARK TOP

... MEM ?

Matrix Element Method : description

- For each event a probability is defined as a function of the probabilities to be signal like (P_{sig}) or background like (P_{bkg}) :

$$P_{evt}(m_{top} | x) = f_{t\bar{t}} P_{sig}(m_{top} | x) + (1 - f_{t\bar{t}}) P_{bkg}(m_{top} | x)$$

$f_{t\bar{t}} \equiv \frac{N_{sig}}{N_{sig} + N_{bkg}}$

- The probability to be signal or background like is related to the differential cross section $\frac{d\sigma}{dx}(m_{top})$ through :

$$P_i(m_{top} | x) = \frac{1}{\sigma_{obs}^i(m_{top})} \frac{d\sigma^i}{dx}(m_{top} | x) \quad \text{with } i = \text{signal or background}$$

$\sigma_{obs}^i(m_{top}) = \epsilon^i(m_{top}) \sigma_{th}^i(m_{top})$

- Computing the differential cross section is the heart of the method :

k : observables of the final state at the reconstructed level

$|M_{t\bar{t}}(p_1, p_2, x, m_{top})|$:
Matrix element e.g.
 $p_1 p_2 \rightarrow t\bar{t} \rightarrow b\bar{b} e \nu_e \nu_\mu$

f_{pdf} : parton density function

$$\frac{d\sigma^i}{dx}(m_{top} | k) \propto \sum_i \int dp_1 dp_2 d\Phi |M_i(p_1, p_2, x, m_{top})|^2 f_{pdf}(p_1) f_{pdf}(p_2) TF(x | k)$$

x : observables of the final state at the parton level

p_1 and p_2 : incoming parton

$d\Phi$: phase space for process

$TF(x | k)$: Transfer Function from k to x

CAS D'UN CODE MEM

- Plusieurs code, plusieurs version, aucune orientée vers du calcul HPC
- Production pour les thèses avec MadWeight 2 :
 - Aurélien (2014), environ 3 Mh sur la grille
 - Sylvestre (2015), environ 10 Mg sur la grille

Dans la pratique, il faut a minima 4 mois pour refaire tourner l'ensemble du code, ce qui est trop pour des investigations fines
- Orientation vers de nouvelles versions (MadWeight 5), de nouveaux codes (MemTk de Berlin), ou une écriture complète ou partielle d'un code MEM (initiative au LLR, à laquelle le LPNHE est/était associée via une demande ANR qui a échoué)

MISE EN COMMUN DE L'EXPÉRIENCE HPC AU LPNHE

Objet : réunir un petit groupe de personnes pour l'émergence de l'utilisation des techniques HPC. Dans un premier temps, mise en commun de l'expérience dans un wiki : <https://gitlab.in2p3.fr/lpnhe/HPC/wikis/home>

À court terme, développement et/ou intégration d'outils pour les besoins de physique.

Premiers membres actifs du groupe au LPNHE :

- Tristan Beau – MCF - (ATLAS)
- Olivier Dadoun – IR - (Service informatique)
- Frédéric Derue – CR - (ATLAS)
- Jean-Philippe Lenain – CR - (HESS/CTA)
- Victor Mendoza – IR - (Service info)