



Le calcul IN2P3 hors de ses sentiers battus

*Randonnée à travers les groupes de travail, projets,
technologies, infrastructures, la formation...*

Groupes de travail "offline" RI3

Nouveaux Langages (C.Meessen, S.Binet)

- D, **Go**, **Rust**, *ADA*, *C++11*.
- Mais ce que veulent les scientifiques : du **Python**, du Python, du Python, (*du Fortran*)...
- En embuscade : Julia, **Scala**...
- Le **fonctionnel** revient en force => Prochaine Ecole IN2P3

Programmation Parallèle et Performance du Calcul

(D. Chamont, G. Grasseau)

- Devenu la liste électronique et le groupe/wiki Codeurs Intensifs sur le GitLab IN2P3.
- A l'origine du montage de projets IN2P3 de R&D transverse : ComputeOps et Reprises.

Projets DCAP de l'IN2P3

Data and Computational Science, Algorithms and Programming

- BETA, Frédéric Suter (CC-IN2P3)
- DIRAC, Andrei Tsaregorodtsev (CPPM)
- **Geant4**, Marc Verderi (LLR)
- PRESOFT, Geneviève Romier (CC-IN2P3)
- **Machine Learning**, Balázs Kégl (LAL)
- **DecaLog**, Gilles Grasseau (LLR)
 - **Reprises**, David Chamont (LAL)
 - **ComputeOps**, Cécile Cavet (APC)
- Workflows4NP, Eric Legay (CSNSM)

Où est la "data science" ? ROOT ?

Des volontaires pour le **Quantum Computing** ?

Objectifs Reprises

1. Explorer les technologies prétendant concilier une écriture de haut-niveau **portable** et **pérenne**, avec une **génération de code** spécifique à chaque matériel (OpenCL, OpenACC/OpenMP, Pythran, Kokkos, ...).
2. Améliorer la **reproductibilité** des résultats en contexte parallèle, et réexaminer le crédo de la double précision, en s'appuyant sur les outils d'**arithmétique stochastique** (Verrou, Verificarlo, Cadna).

Expérimentations Reprises

- Technologies de conteneurs
 - **Doker**, Singularity..
- Vectorisation
 - **Génération de code.**
 - **UME, Vc, xtensor & xsimd**, TensorFlow.
 - **OpenMP.**
- Accélérateurs/coprocesseurs
 - **OpenCL** sur **GPU** et **FPGA.**
 - **OpenACC/OpenMP**
 - Kokkos, HPX.
 - TensorFlow.

Objectifs ComputeOps

1. Comparer les technologies de conteneurs : docker, rocket, lxd, udocker, singularity, shifter... Y'a-t-il un avantage décisif à dédaigner l'outil dominant (**Docker**) et à se tourner vers les alternatives "HPC" (**Singularity**, Shifter,...) ?
2. Etudier l'**interopérabilité** des technologies. Notamment les **images** et les **recettes**. Est-ce qu'un utilisateur pourrait développer sur son poste en Docker, puis déployer sur centre de calcul en Singularity ?
3. Vérifier les performances et valider la **compatibilité** des conteneurs avec la grille, la vectorisation, les **accélérateurs** de calcul, l'**infiniband**...

Progrès ComputeOps

- Sélection d'applications pilotes
- Comparatif des technologies
 - Mise en place d'un Singularity Hub.
- Compatibilité Singularity-HPC
 - GPU : OK.
 - Infiniband : optimistes.
 - OpenMPI : il faut exactement la même version sur l'hôte et dans le conteneur...

SuperComputing Reproducibility Initiative

Applications pilotes

- Reconstruction de traces, dans ACTS (et le Tracking Kaggle)
- Code multi-GPUs d'analyse MEM (Matrix Element Method)
- Algèbre linéaire sur FPGA pour ALICE-Upgrade
- Corsika : simulation de gerbes atmosphériques CTA
- Compression d'image pour CTA
- Calcul dynamique non linéaire
- Code de cosmo pour LSST
- Geant4, Gate... pour la dosimétrie
- Simulation neutronique
- SMILEI, code Particle-In-Cell pour la simulation des interactions laser-plasma.
- CMS MEM : analyse multi-noeuds multi-GPUs des productions de Higgs sur le canal ttH.
- HAhRD : machine-learning sur GPU, basé sur Tensorflow, pour reconstruire les images du détecteur HGCal de CMS.
- Calcul de taux de capture d'électron, pour l'astrophysique, en Fortran90 avec des directives OpenACC.
- LDC : système de simulation pour LISA.
- Simulation Geant4 pour les expériences de détection de la matière noire (Xenon & DarkSide)
- ...

Geant 4 & Geant-V

- Geant-V
 - A revisité la structure de données, et prévoit une beta fin 2018, visant une performance x2... (1/3 par vectorisation).
 - La vectorisation de la physique EM avance, et certains modeles hadroniques sont repensés.
- Geant 4
 - Révise son modèle de tâches, pour utiliser les threads C++11, se séparer de TBB ?, sous-traiter certains calculs aux GPUs...
 - Aimerait profiter des 2/3 du x2 Geant-V.
 - Il y a des idées pour utiliser du machine learning pour de la simulation rapide de gerbes.
 - Des acteurs de l'IN2P3 suivent et accompagnent ce mouvement.

ROOT

- Géométrie vectorisée
- Des classes apparaissent, qui ont une inspiration "Data Science" : TDataFrame, RDataFrame.
- Peu/pas de contribution IN2P3 ?

Infrastructures

- Plateforme R&D P2IO (hébergée au LLR)
 - 2 noeuds 2xK20, 2 noeuds 2xXeonPhi
 - 1 noeud 6xTitan, 1 noeud AMD
 - 1 Noeud 2xV100
- Au CC
 - HPC cluster : 512 coeurs, Infiniband.
 - GPGPU cluster : 20 noeuds C4130 2xK80, Infiniband.
- Mésocentres
- Au CERN : le techlab

Formation

Quelques Ressources

- [Natation Synchronisée](#) : tutoriels IN2P3.
- [Coding Pool](#) : catalogue de tutoriels en ligne.
- [FoCaL](#) : formations en calcul.
- [Les écoles listées par HSF](#) : CERN School of Computing, Bertinoro School of Computing, GridKa, Alghero School, Lund COMPUTE, CINECA, Machine learning in HEP...

Proposition de projet **e-coding**

- Recenser l'offre de tutoriels en ligne (Coding Pool).
- Collecter la demande, produire du contenu.
- Faciliter la mise en oeuvre (conteneurs, JupyterCloud).
- Valoriser auteurs, formateurs et apprenants (FP).