

Ressources de Calcul au LLR

A. Sartirana

- Historiquement nous avons commencé à **gérer des moyens de calcul** (notamment **grille**) pour permettre au **groupe CMS** d'être **présent et actif dans WLCG**
 - ❑ en **2005 LLR rejoint** le site T2 WLCG **GRIF**
 - ❖ au fil du temps le LLR est devenu le **«pole-CMS» de GRIF**;
 - ❑ le site grille local a été un **ressource précieuse et très utilisée**;
- ... au fil des années, **le calcul a évolué** et nous essayons d'**adapter nos offres** aux nouveaux **besoins** at aux nouvelles **technologies**
 - ❑ **autres groups** ont exprimé des besoins (pas forcément "grid-compliant");
 - ❑ **HPC, manycores, Cloud, ...**

- L'Ecole polytechnique met à disposition une **salle machine complètement équipée**
 - ❑ **600kW** puissance électrique/clim dont **400kW ondulée**.
 - ❖ fluides/maintenance/monitoring... fournis par l'X;
 - ❑ **~200m²** partagés avec autres (~12) labos
 - ❖ LLR a actuellement 12 racks (sur ~64. max possible 70);
 - ❑ **renouvelé Q3/2014**
 - ❖ contrib. P2IO;
- réseau:
 - ❑ **Renater, Saphir, UPSay**



Plus de détails dans la présentation de **P.Hennion**

- Site **Tier-2 WLCG distribué**, constitué de la fédération de 6 labos HEP en IdF
 - ❑ aujourd'hui ressources grille dans 5 labos;
- **~18k slots** de calcul et **~8PB de stockage**;
- **~30%** de l'activité grille en **FR**
 - ❑ 4 VO's **LHC**; (~20 VO's totale);
 - ❑ pas seulement grille: collaborations aussi sur des **projets cloud**.



➤ **Gestion collaborative** des moyens

- ❑ **comité technique** (~10 personnes = ~5.5 FTE);
- ❑ réunions mensuelles, visio hebdomadaires, mails,...
- ❑ **partage** de **connaissance** et **information**;

➤ **choix techniques** homogènes
et **services** partagés

- ❑ Quattor;
- ❑ monitoring.



➤ LLR est un des 6 **membres du Tier-2 GRIF**

- ❑ **cluster** HTCondor (+ CREAM) unifié avec le LAL (~5.5k cœurs dont ~2.5k au LLR)

 - ❖ DELL R4X0 et C6X00, 3GB/slot, 10gbps network;

- ❑ **1.5PB** de stockage (DPM)

 - ❖ serveurs de fichiers SuperMicro, (2x)10gbps nw;

- ❑ réseau LAN 2x10gbps en cours d'upgrade a 100gbps;

- ❑ le "pole" **CMS** de GRIF

 - ❖ contacts avec les ops central de CMS;

 - ❖ services pour CMS: Xrootd EU redir, CMS Middleware readiness testbed;

- ❑ autres VO's "locals" actives: ILC, T2K, Hess, Harpo, (LHCb).



- Une partie des moyens de calcul grille est dédiée à un **ressource locale de type Tier-3**
 - ❑ ~500 cœurs (**variable selon les besoins**);
 - ❑ partage le stockage T2 + NFS (~50TB);
 - ❑ **soumission** batch **directe** avec HTCondor et aussi accès **interactive**;
 - ❑ **ressource important pour le groupe CMS**;
 - ❑ nous avons de projet pour **rendre plus dynamique** le passage des ressources entre T2 et T3
 - ❖ basées sur **cloud et/ou dockerization** des WN's;



- Depuis 2012 on gère aussi des moyens de calcul **HPC**
 - **poussé par le projet Galop** qui nécessitait d'une ressource local pour développer/tester SMILEI, son code PIC
 - ❖ reproduire "en petit" les conditions sur PRACE et GENCI;
 - ❖ CMS aussi: analyse Higgs dans les canaux VBF et ttH;
 - les clusters sont **mutualisés avec le labo LSI**
 - ❖ partage de l'utilisation;
 - ❖ achats communs;
 - ❖ collaboration dans l'administration.

- ❑ **Tornado:** 288 coeurs, IB EDR (100Gb/s)
 - ❖ 2 x E5-2697 v4 @ 2.30GHz, 1.8GB/core ram;
- ❑ **JollyJumper:** 672 cores, IB FDR (56 Gb/s) IB connection
 - ❖ 2 x E5-2697 v2 (12cores), 2.7GHz, 2GB/core ram;
- ❑ **Rantanplan:** 128 cores, QDR (40Gb/s) IB connection
 - ❖ 2 x E5620 (4cores), 2.4GHz, 6GB/core ram;
- ❑ **SPS NFS storage** (15TB) 10Gb/s + IB FDR network.

- ❑ Services (ldap, batch, gitlab, etc) sur dockers ou VM;
- ❑ a l'étude l'utilisation de containers (singularity) pour avoir des env personnalisées pour les utilisateurs
 - ❖ actuellement on utilise « module »;
 - ❖ liason avec ComputeOps.

➤ **GridCL (2012)**: R&D GPGPU/manycore avec **OpenCL** (P2IO)

❑ **moyens** manycore/GPU hétérogènes pour tester le code

❖ la **portabilité** étant un de points clé de OpenCL;

- ✓ 2 noeuds avec 2 Nvidia K20 GPU chaque;
- ✓ 2 noeuds avec 2 Intel Xeon Phi chaque;
- ✓ 1 noeuds avec 6 Nvidia Titan GPU chaque;
- ✓ 1 noeuds avec 2 AMD FirePro GPU chaque;



❑ activité de **développement** par les membres du proj.
Aussi externe (e.g. CERN);

❑ **training**: JDEV '13, IN2P3 School '16;

❑ "petites" productions;

❑ **outil** pour évaluer **nouvelles technologies**

❖ e.g. deep learning sur GPU, docker/singularity avec GPU.

- Nouveau projet **P2IO: ACP** (2017)
 - ❑ adaptabilité' du code a **différents plateformes** via générateurs de code/templates;
 - ❑ adaptabilité' des **environnement** via **containerisation**;
 - ❑ **deep learning**.



Objectifs en ligne
avec les projets
ComputeOps et
Métamorphe

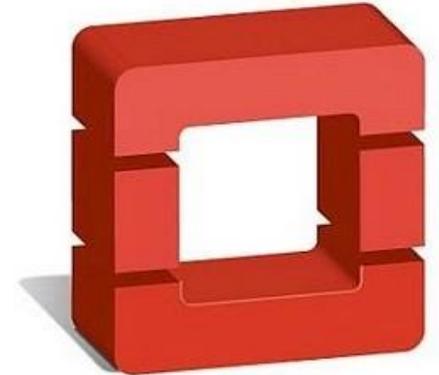
Nous allons acheter une **plateforme** de développement pour ce projet, avec des **architectures de nouvelle génération**: Nvidia Pascal et Intel KNL

➤ Projet de **cloud Openstack**

- ❑ premier **prototype** en **Q4/2017**;
- ❑ possible utilisations en prod:
 - ❖ "backend" pour le **SI**;
 - ❖ **partage dynamique des ressources** de calcul entre différentes plateformes;
 - ❖ cloud de calcul (fédération);

➤ projet P2IO: plateforme de stockage **CEPH distribuée**

- ❑ hébergeurs: **LLR, LAL/IPNO, IRFU**;
- ❑ **CSNSM, IAS, IMNC, LPT**.



- Le SI du LLR gère plusieurs ressources de calcul: Grille, HPC, many-core, (Cloud)
 - ❑ ressources de proximité: créées à la demande des utilisateurs et en interaction très proche avec eux;
 - ❑ ressources mutualisées (quand possible);
- Synergies avec APC et LPNHE
 - ❑ les activités grille sont déjà, de facto, fédérées dans le contexte de GRIF;
 - ❑ projet cloud très proches entre LLR/LPNHE;
 - ❑ proximité entre GridCL/ACP et le projet ComputeOps.