HPC activities in LCG-FR

Catherine Biscarat, LPSC/IN2P3/CNRS
Technical coordinator of LCG-France

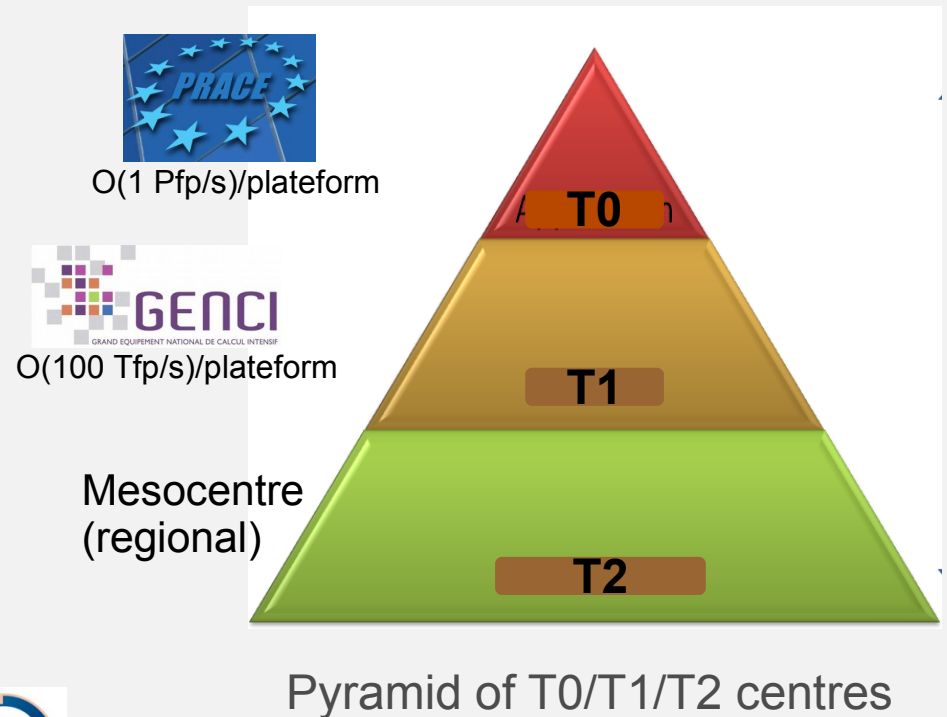"FR-cloud, Regional Centres Meeting",
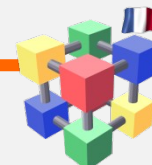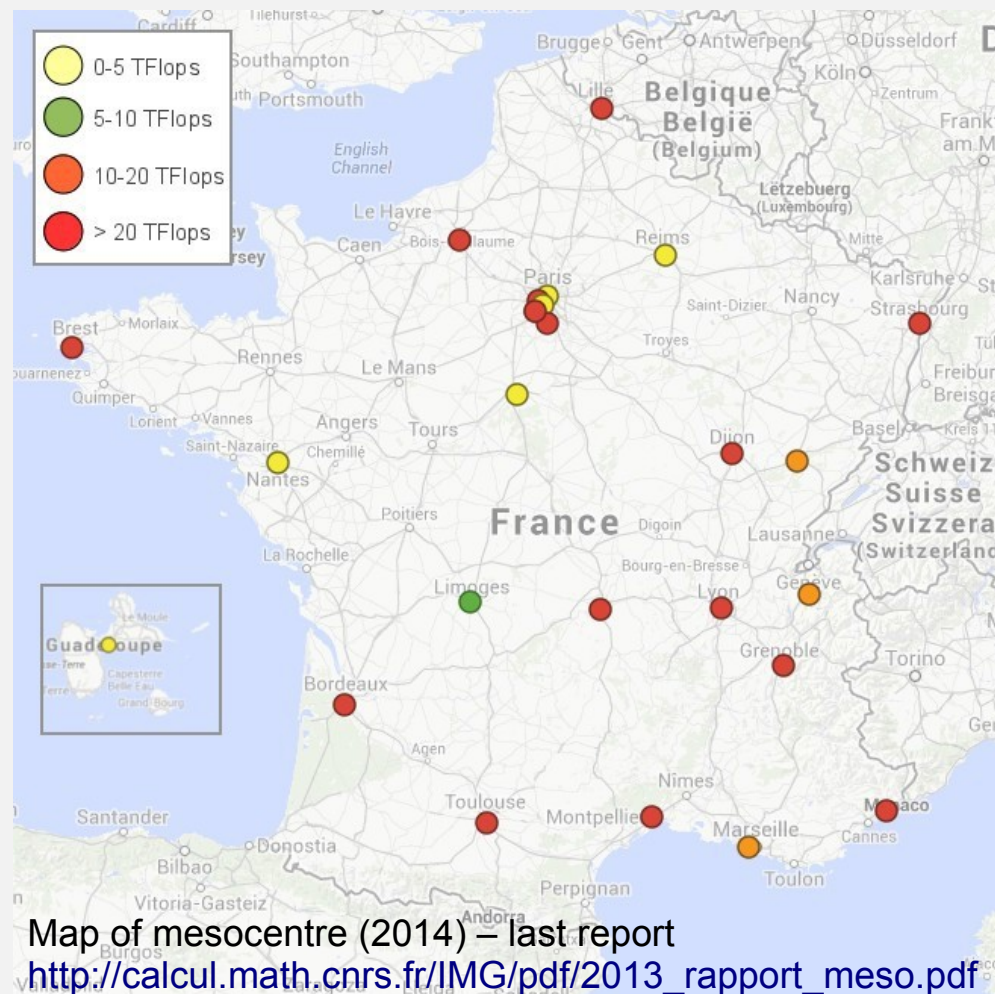11-12 April 2018, Hong Kong
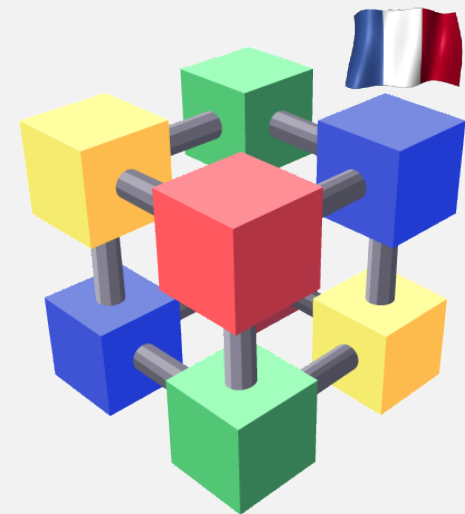
HPC in France

- GENCI (*Grand Equipement National de Calcul Intensif*)

- **T0 and T1**: call for projects twice a year
  - No gigantic machines (top500: >40)
  - IDRIS (HPC centre for the French National Centre for Scientific Research CNRS).
  - CINES (the computing centre for the French Ministry of Higher Education and Research)
  - TGCC (Very Large Computing Centre) of the French Alternative Energies and Atomic Energy Commission (CEA)
- **T2**: access to the resources is more flexible
  - Integrated in regional politics
  - Some of them integrated in the national project Equipex@meso (GENCI) to promote efficient jump to T1

O(1 Pfp/s)/plateform

O(100 Tfp/s)/plateform

Mesocentre (regional)

T0

T1

T2

Pyramid of T0/T1/T2 centres

HPC in France

Map of mesocentre (2014) – last report
http://calcul.math.cnrs.fr/IMG/pdf/2013_rapport_meso.pdf

Directions

- HPC applications have more and more data to take care of
  - New challenges: big data services, grid and cloud, network

- **European landscape:**
  - EOSC (European Cloud Infrastructure): interconnect of all computing resources
  - EDI (European Data Infrastructure) : interconnect of largest HPC centres, high bandwidth (data)
    → EuroHPC : investment in ExaScale supercomputers (FR has signed up early 2017)
    - Co-design applications/processors

- **France landscape:**
  - GENCI takes care of the renewal of the three HPC national centres:
    - Computing power = 6,8 Pflop/s 2016 (x400 in 10 years)
    - 2017-Occigen-CINES (3,5 PF); 2019 -Irene-TGCC (9 puis > 20PF); 2019-IDRIS
  - Growing interest in coupling grid/cloud/HPC (data management)
  - Some national project are under discussion (HC, cloud, grid: "IR-T2").

FR Initiatives for LHC computing

- Several regional initiatives
  - LPSC (LCGFR T2)
  - LPNHE (LCGFR T2_GRIF)
  - Subatech (LCGFR T2)

  } Presented in Beijing in 2016
  LPSC & LPNHE use regional HPC (a few 1k cores)
  for local analysis - Interconnect to ATLAS central
  production on hold (RH)

- New: the CC-IN2P3 has build a demonstrator with a national facility

HPC – Subatech (1)

- A project by the Subatech LCG-T2 centre and a regional HPC centre (CCIPL)
  - In the global direction of mutualisation, grant (regional) obtained for 2016-2020
    - disk: localised at Subatech   http://www.ccipl.univ-nantes.fr/
    - CPU: localised in a new Datacenter ~10 km from Subatech (3440 CPU cores, LCG share=400 HT-core equivalent to Subatech T2)
  - Aim to declare resources in the site pledges

> Subatech ALICE
>
> J. Barbet
> L. Aphecetche

Risks

- Subatech is a "power user", two different units (directions and people) → need cooperation
- Security issues: had to establish minimum incident response procedure

Gains

- Infrastructure is taken care of by CCIPL (& cooling cost)
- Opportunistic resources (still under study)

- Configuration
  - Pilot jobs to be submitted via the batch system
  - An "ALICE-Box" inside the new DC
    - No additional services (CE)
    - Network gateway for exporting the data
    - A 10 Gbps link has been deployed
  - SLURM batch system, WNs are linux x86_64
  - CMVFS available on the WN (tarball deployed)
- In production since may 2017
  - New Monalisa queue Subatech_CCIPL
  - Only MC production jobs

Current setup
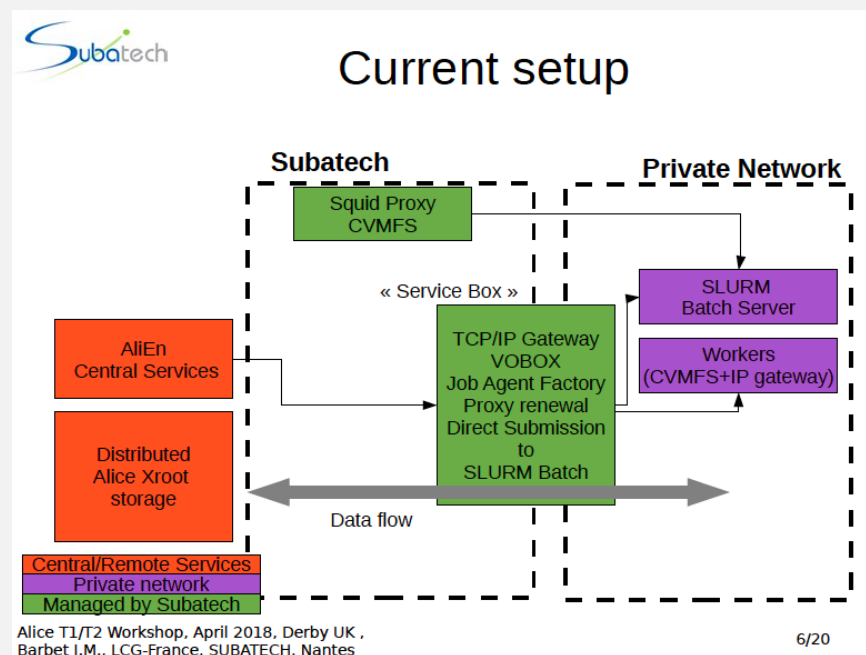
ALICE T1/T2 workshop, this week, UK

Thoughts

- Integrating an HPC centre to behave like an EGI site (A/R, pledges, security) is difficult
- Accounting is recorded in APEL to match the pledges (since April 1st) - Hand made script
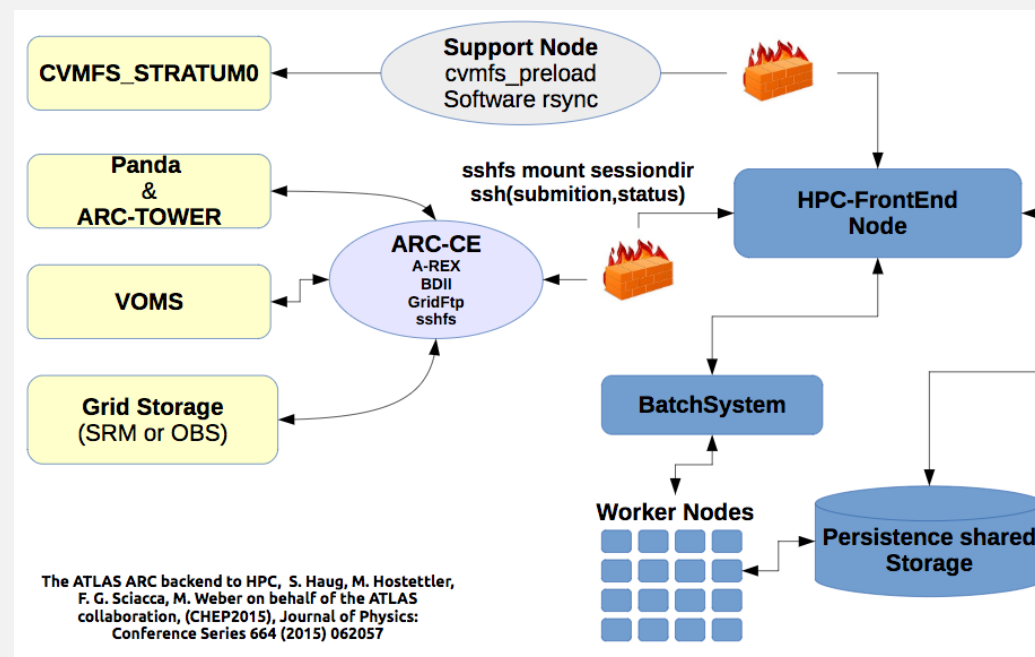
HPC – CC-IN2P3 (1)

The framework

- Collaboration between IDRIS and CC-IN2P3 (November 2016)
- IDRIS is the national CNRS centre for HPC
  - <u>TURING</u>: IBM Blue Gene/Q (~100k cores, Power PC)→ does not suits ATLAS/HEP applications
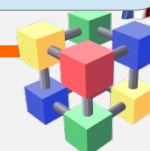 - <u>ADA</u>: IBM processor Intel E5-4650 (~10k cores, EL6) → suits ATLAS/HEP applications

The setup

- ARC-CE: ssh + SW delivery + interface to LRMS + stage in/out of the data
- SW: lack of local scratch disk to install cvmfs cache → Preload-cache + periodic rsync + pre-compiled parrot client on WN

The ATLAS ARC backend to HPC, S. Haug, M. Hostettler, F. G. Sciacca, M. Weber on behalf of the ATLAS collaboration, (CHEP2015), Journal of Physics: Conference Series 664 (2015) 062057

E. Vamvakopoulos
LCG-France workshop, spring 2017
https://indico.in2p3.fr/event/14042/
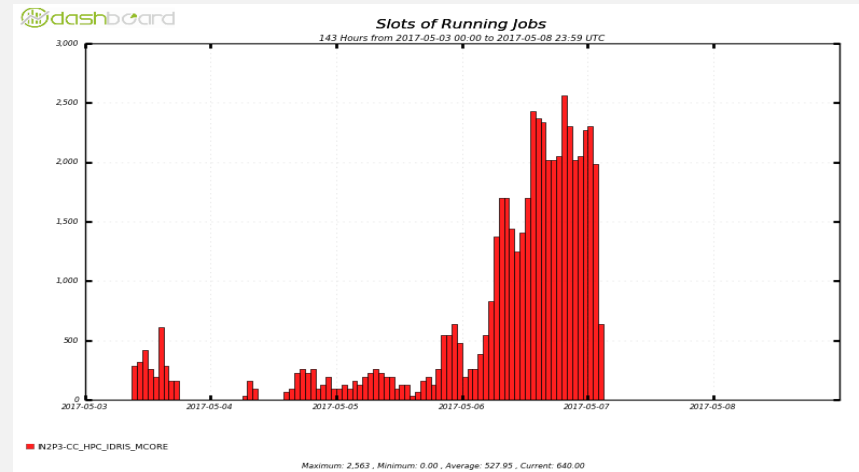
CC-IN2P3
ATLAS

E. Vamvakopulos

HPC – CC-IN2P3 (2)

Resources available & used
- No limit on max running jobs
  - Pre-emptive queue (best effort)
- 32 multi-cores job
- This is a demonstrator at this step

Next steps
- Tune BDII publication
- Have started to include the Event-service (fault tolerance)
- The efficiency of the current backfill mode is under investigation

Future of this collaboration
- ADA machine is going for decommission at the end of 2018,
- New machine : question about its compatibility with atlas pre-compile production software
- More news this fall

(My) Final thoughts

HTC-HPC initiatives
- WLCG is questioning the effort to invest
  - Each HPC centre is different and needs quite some involvement

In France
- A very few regional initiatives (Tier-2/HPC-cluster in region)
- A collaboration with a national centre (CC-IN2P3/IDRIS)

Regional initiatives
- A way to sustain LCG-T2
  - HR at sites, though we have to maintain the grid at its very best at the same time
- Gives access to innovative technology (GPU)
- Aligned with the national politics (EOSC)

National initiatives
- This has to be discussed at high level (EuroHPC)

Additional material

IRENE = CURIE 2

The new high-end HPC platform successor of CURIE at TGCC: 9 Pflop/s for PRACE and for French scientific and industrial Open Science

- ❑ Atos/BULL SEQUANA system
- ❑ 9 Pflop/s peak for the first phase
- ❑ Mixed Intel processor technology :
  - ➢ KNL 7250
  - ➢ Skylake 8168 24-core 2.7 GHz
- ❑ Close to 400 TB of distributed memory
- ❑ Next generation interconnect
- ❑ Up to 500 Go/s multi-level Lustre
- ❑ Provisioned by GENCI & operated by CEA

❑ Availability H1 2018 - Grand Challenges

PRÉSENTATION GENCI 6-09-2017

http://calcul.math.cnrs.fr/spip.php?article285
Intervention GENCI

Equip@meso (2011 + 2012)
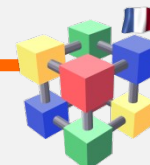
- CIMENT is the regional T2-HPC centre in Grenoble
  - created in 1998, under the umbrella of GRICAD
  - ~ 10 clusters, ~6000 cores, 130 Tflop/s (2/3 equip@meso)
  - ~250 users on average
- Collaboration between CIMENT and the LPSC started in 2010
  - Share of resources and expertise
- CIMENT is unique:
  - A home made computing grid on top of the cluster
  - Increase the clusters occupancy
  - A data grid also available: iRODS

- CIGRI
  - No central applications or data management
 - Opportunistic jobs (OAR job scheduler)

LPSC
ATLAS
ALICE

C. Biscarat
J. Stark
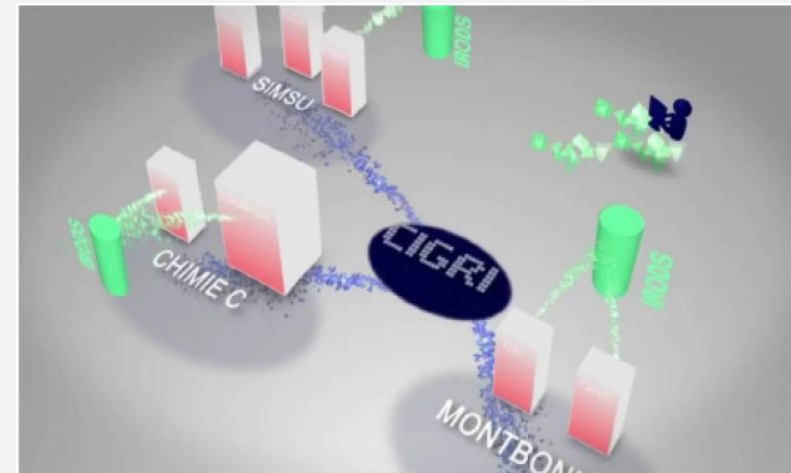
- LPSC usage of CIGRI: like a local T3 (ssh by a few users)
- First for Dzero: detector simulation (2011) – W mass measurement
- > 2012: ATLAS event generation (DIPHOX), upto 1k cores used at once (~ LPSC-T2 size)
- 2017: agreement with GRICAD to interconnect CIMENT with ATLAS
  - Technology: "a la ATLAS", ARC-CE, tar cvmfs
  - Gaps: OAR, manpower, recognition
  - Project on hold

Usage by HPC owners (blue) and opportunistic (grey)
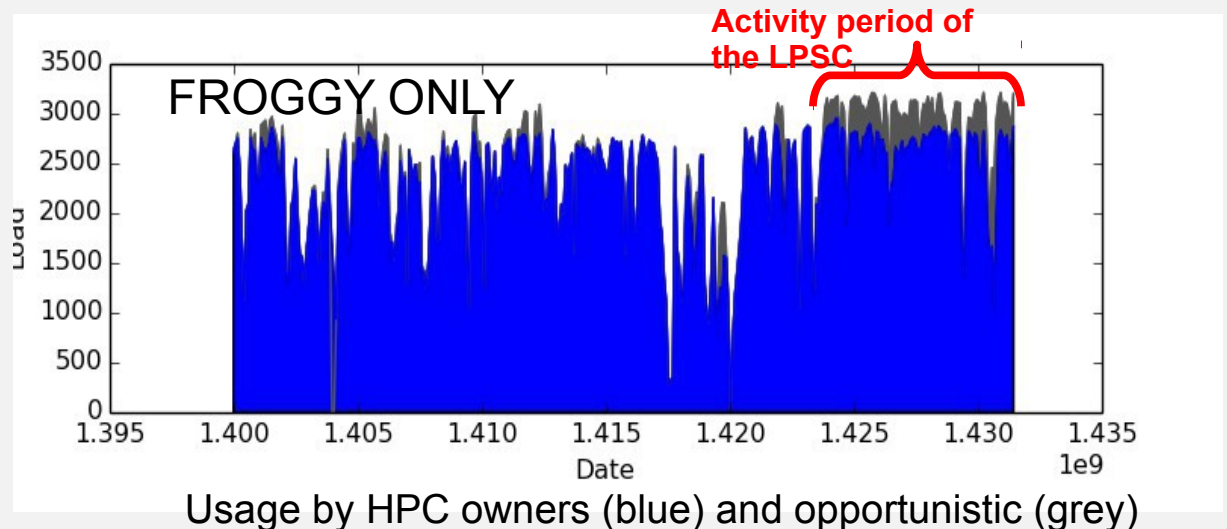
HPC – LPNHE (1)

- LPNHE is part of GRIF (lab in the heart of Paris)

Machines

- A local HPC cluster with GPU and Xeon Phi
- Mostly for training purpose
- A more powerful GPU card for LHCb online tests
- Additional cores at *Sorbonne Universités* (same campus)
- ~1000 cores, 20 Tflops, 16 MB shared memory (Equip@Meso)
- ~2000 cores, 60 Tflops

LPNHE
ATLAS

A. Bailly-Reyre
F. Derue
V. Mendoza

http://iscd.upmc.fr/expertise/mesu/)

Access to this resource

- ssh access for University members,
- limited local support (2 persons at university (one being at LPNHE), but mainly through company),
- Took a long time to install SW like CVMFS
- Not (yet) link to the local grid Tier2

HPC – LPNHE (2)

Use cases
- ATLAS : access to more resources for final analyses, speed Fits or ML
- LHCb : real-time analysis, ERC grant (V. Gligorov), in collaboration with LIP6 (lab of Informatics of University) ;
- Other users interested : CTA, cosmology ....

- ATLAS aims to include the clusters in the production system (middle term)

Current issues
- OS = Suse Linux Enterprise Server, quite old → ROOT does not compile – library pb
  - Need to deploy singularity/docker
- cvmfs installed (but took long): a service node with a shared area

Other possibility to better serve non LHC communities: Cloud
- prototype developed since arrival of a new engineer (2017), Openstack;
- build from "old" machines from grid Tier 3, few tens of VMs and TB;
- aim to be delivered in the coming weeks for some local users (first), then extended to more communities if successful