



GPU and HPC usage

Frédéric Derue, LPNHE Paris

Calcul ATLAS France (CAF) meeting
CC-IN2P3 Lyon, 1st April 2019

ATLAS usage of GPU

- **Meeting to come :**
 - CC-IN2P3 GPU workshop, 3-5th April, Lyon [*indico*]
 - ⇒ see my slides on the today's agenda

HPC in lab and universities

- **Maison de la Simulation (Orsay/Saclay, IRFU, LAL) [[link](#)]**

Machine Poincaré [[link](#)] part of [Equip@Meso](#)

IBM Computer, mostly made of iDataPlex dx360 M4 servers :

- 92 nœuds de calculs “poincare[001-092]” (clallmds dans LoadLeveler) :
 - 2 processeurs Sandy Bridge E5-2670 (2.60GHz, 8 cœurs par processeur, soit 16 cœurs par nœud)
 - 32 Go de mémoire par nœud
- 4 nœuds GPU “poincaregpu[01-04]” (clgpu dans LoadLeveler) équipés de :
 - 2 processeurs Sandy Bridge E5-2670
 - 64 Go de mémoire par nœud
 - 2 GPU Tesla K20 (Cuda Capability 3.5, 4.8 Go de mémoire par GPU)
- 4 frontales interactives “poincareint[01-04]” équipées de :
 - 2 processeurs Sandy Bridge E5-2670
 - 32 Go de mémoire par nœud

HPC in lab and universities

- **Virtual Data (IRFU, LAL, LLR) [[link](#)]**
Ressources matérielles
 - 2 serveurs, chacun doté de 2 cartes NVidia K20 et d'Infiniband :
 - serveur Dell R720,
 - carte mère sandy-bridge 2xE5-2650, 64 Go RAM, 1 To disque,
 - 2 cartes Intel K20M, 2 liens Infiniband QDR.
 - 2 serveurs, chacun doté de 2 cartes Intel Xeon Phi et d'Infiniband :
 - serveur Dell R720,
 - carte mère sandy-bridge 2xE5-2650, 64 Go RAM, 1 To disque,
 - 2 cartes Intel 5110P, 2 liens Infiniband QDR.
 - 1 serveur doté de 6 cartes NVidia Titan et d'Ethernet 10G :
 - serveur Transtec CUDA 8230,
 - carte mère ivy-bridge 2xE5-2650 v2, 128 Go RAM,
 - 6 cartes NVidia Titan, 2 liens Ethernet 10G.

HPC in lab and universities

- CIMENT (Grenoble / LPSC) [[link](#)]

Name	Vendor	Number of cpu cores	Total memory	Max memory/node	Total storage (net)	Computing network	Total Gflop/s	Accel Gflop/s *	Purchase date
Dahu	Dell	1280	7680 GB	192 GB	189.6 TB	Omnipath 100Gb/s	43000	0	2018-01-10
Bettik	Dell		448 GB	64 GB	560 TB	Omnipath 100Gb/s	0	0	2017-01-10
Luke	Dell	1148	6992 GB	512 GB	1062.52 TB	Infiniband QDR	32313	0	2014-05-01
Froggy	Bull	3236	15456 GB	512 GB	118.55 TB	Infiniband FDR	88250	24380	2013-05-31
Ceciccluster	Dell	216	432 GB	24 GB	12.5 TB	Infiniband QDR	1961	0	2011-12-01
Gofree	Dell	336	2016 GB	72 GB	30 TB	Infiniband QDR	3178	0	2011-01-01
Global_storage	Dell		456 GB	32 GB	920 TB	10Gb/s ethernet	0	0	2010-09-01

-*: 'Accel' means accelerators, such as GP-GPUS

- Total number of cores: 6592
- Total op/s: 168.7026 TFlop/s
- Accelerators op/s *: 24.38 Tflop/s

Froggy (vieux cluster):

GPU nodes: 9 Bull R421-E3 computing nodes:

cpu: Intel E5-2670 + GPU K20 (Sandybridge + Kepler) 2.6 Ghz

cores: 2x8 ; memory: 32 GB ; computing network: Infiniband FDR

Dahu (nouveau cluster)

3 nodes with 4GPU Tesla V100 with nvlink

HPC in lab and universities

- **MUST (Annecy / LAPP)** [[link](#)]
 - WLCG site
 - 16 servers connected with Infiniband QDR
 - 3 computing nodes equipped with GPU type NVIDIA K80
 - 1 computing node equipped with GPU type NVIDIA V100

HPC in lab and universities

- **MesoCentre AMU (Aix-Marseille / CPPM) [link]**

memory: 31 TB, storage : 1.3 PB, computing : 338 Tflops/s from 5236 cores
(nœuds fins, nœuds à forte mémoire (SMP),
nœuds GPU et nœud dédié au calcul sur carte Xeon PHI)

new cluster (2018)

Noeuds de calcul Sky Lake :

- ✓ 112 noeuds Dell PowerEdge C6420 ;
- ✓ 32 coeurs par nœud, processeur Intel® Xeon® Gold 6142 (Sky Lake) à 2.6 GHz ;
- ✓ 192 Go de RAM par nœud;
- ✓ 100 Go sur disque local SSD.



Noeuds de calcul GPU :

- ✓ 7 noeuds Dell PowerEdge C4130 (24 coeurs par noeuds) Intel(R) Xeon(R) CPU E5-2680 v3, 320 Go de RAM par noeud et 2 cartes bi-GPU NVIDIA Tesla K80 (4992 coeurs CUDA);
- ✓ 2 noeuds Dell PowerEdge C4130 (28 coeurs par noeuds) Intel(R) Xeon(R) CPU E5-2680 v4, 320 Go de RAM par noeud et 2 cartes NVIDIA Tesla P100 (3584 coeurs CUDA).



- **Computing Department of University (LIS) – through PhD co-tutelle**

The LIS cluster gathers 12 hosts with 24 CPUs (188 cores) and 23 GPUs.

Tesla K80: 10

Tesla K20m : 1

Tesla K40c: 2

Tesla K40m : 2

GeForce GTX 1080 Ti: 8

HPC in lab and universities

- **Institut des Sciences du Calcul et des Données**

Sorbonne Université (Paris / LPNHE) [[link](#)],

- [mesu-alpha](#) : 19 Tflop/s. SGI UV2000 machine part of Equip@Meso
64 blades, each of which made of 2 motherboards (or nodes), equipped with one Intel Xeon (Sandy Bridge) of 8 cores and with a local memory of 128Gb
- [mesu-beta](#) : 59 Tflop/s. SGI ICE-XA server made of 3012 high performance computing cores, distributed in 144 nodes of 24 core and with a local memory of 128Gb
- but old OS SuSE Linux Enterprise Server (SLES 11 SP2),
ROOT does not compile, not allowed to install Singularity

But also at LPNHE

- two processeurs Intel® Xeon Phi™ Coprocessor 5110P (8GB, 1.053 GHz, 60 cores)
- two cards GPU K2200, each with 5 multiprocessors of 128 cœurs,
4GB GDDR5, thus 1280 cores available

HPC at CC-IN2P3

Presentation of P-E Macchi

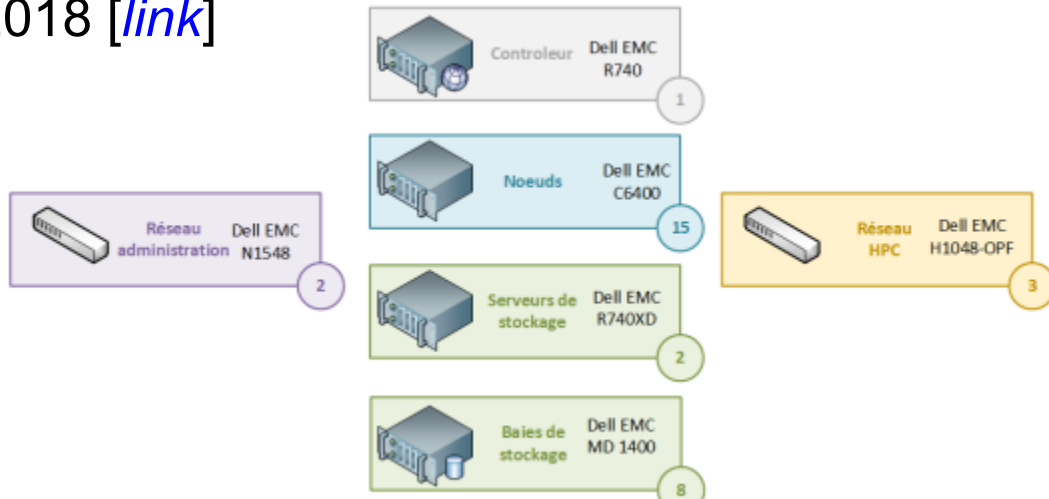
Journées Informatique IN2P3/IRFU, 3 octobre 2018 [[link](#)]

- HPC cluster : 16 C6320 Infiniband - 512 cores
Intel(R) Xeon(R) CPU E5-2698 v3 @ 2.30GHz, 128 GB memory
- GPGPU cluster : The GPU jobs are run on CentOS 7 and the current platform consists of 2 types of GPUs:
 - 10 Dell C4130 with 4 GPUs and 16 CPU cores per machine :
 - 2 Xeon E5-2640v3 (8c @2.6 Ghz)
 - 128 GB RAM
 - 2 Nvidia Tesla K80 → 4 GPU Nvidia GK210 with 12 GB DDR5
 - InfiniBand between nodes
 - 6 Dell C4140 with 4 GPUs and 20 CPU cores per machine :
 - 2 Xeon Silver 4114 (10c @2.2 GHz)
 - 4 NVidia Tesla V100 PCIe → 4 GPU Nvidia with 32 GB HBM2
 - No InfiniBand
- [Howto](#) for ATLAS users

HPC at CEA

• Intermediate size machine

Presentation of G. de Lahaut, Journées Informatique IN2P3/IRFU, 3 octobre 2018 [[link](#)]



364 HDD
8 SSD
2400 Coeurs
750000€
100 Go/s
11 To RAM
60 Nœuds

Outils	rôles
TrinityX/Ansible	- Contrôleur - Configuration des images
Luna	- <i>Provisioning</i>
OpenLDAP/SSSD	- Authentification
Zabbix	- Monitoring
Slurm	- <i>Scheduler</i> HPC

• TGCC [[link](#)]

Host the (old) Curie machine, and (new) Joliot-Curie machine [[link](#)], part of PRACE

Related projects at IN2P3 and IRFU

- **Activités et vision pour le domaine HTC / HPC**

[[link](#)] Feb 2017

- **11èmes Journées Informatique IN2P3/IRFU**

[[link](#)] Oct 2018, in particular [this presentation](#) by D. Chamont

- Programmation Parallèle et Performance du Calcul

(D. Chamont, G. Grasseau) [[link](#)]

→ Devenu la liste électronique et le groupe/wiki Codeurs Intensifs sur le [GitLab IN2P3](#)

→ A l'origine du montage de projets IN2P3

- [ComputeOps](#)

- [Reprises](#)

HPC at IDRIS

- **Previous tests**

- see [*this talk*] of A. Filipcic at Jamboree meeting on 5th March

FR

- IDRIS
 - successful tests in 2018
 - But now offline

- successfully tested by Manoulis in 2018

- see technical reports prepared by Manoulis on the agenda

- final report [[link](#)]

- running ATLAS applications in an HPC center: IDRIS of CNRS [[link](#)]

- Interface CCIDRIS pour les jobs Grille Cas d'utilisation d' Atlas [[link](#)]

HPC at IDRIS

- **New Jean Zay machine at IDRIS** [[link](#)]

Hardware configurations of the Jean Zay supercomputer

Jean Zay is an HPE SGI 8600 computer which is partitioned into two sections. It is composed of scalar nodes and converged nodes, or more precisely, "converged accelerated hybrid nodes". These hybrid nodes are equipped with both CPUs and GPUs, each with the latest available generation, which will permit the usages of both HPC and AI. The second partitioned section contains more than 1000 GPUs. All of the compute nodes will be interconnected by an Intel Omni-Path network and access a very high bandwidth shared file system.

- 1528 XA730i compute nodes, with 2 Intel Cascade Lake 6248 processors (20 cores at 2.5 GHz), or 40 cores per node
- 261 XA780i converged nodes, with 2 Intel Cascade Lake 6248 processors and 4 Nvidia V100 SXM2 GPUs (32 GB each)
- Cumulated peak performance of 13.9 Pflops/s
- Each node will have 192 GB of memory
- Omni-Path interconnection network 100 GB/s (1 link per scalar node and 4 links per converged node)
- 5 frontal nodes (with 2 Intel Cascade Lake 6248 processors and 192 GB of memory per node)
- 5 visualisation nodes (with 2 Intel Cascade Lake 6248 processors per node, 1 GPU Nvidia P6000 and 192 GB of memory per node)
- 4 large memory nodes (with 4 Intel Skylake 6132 processors with 12 cores at 3.2 GHz, 1 GPU Nvidia V100 and 3 TB of memory per node)
- Spectrum Scale system of parallel files (ex-GPFS)
- Parallel storage device with SSD disks (GridScaler GS18K SSD) with a capacity of 1 PB

Provisional calendar

- Installation during the first six months of 2019
- First usage by selected projects (« Grands challenges ») during summer 2019
- Opening of usage to all users anticipated at end-summer 2019

- as ATLAS we can be interested to use it, for production, for analysis
- need to see how to access this machine

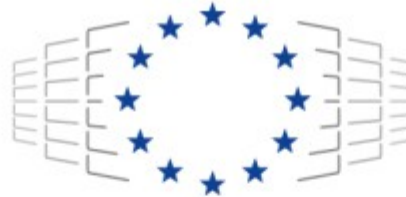
Prospects with existing HPC centers

see [*this talk*] of A. Filipcic at Jamboree meeting on 5th March

- Many were successfully tested in the past, but were not put into steady production
 - Some are small
 - Many cannot provide significant resources without approved allocation
- Allocations in PRACE are limited in time and scope - mostly tuned for single scientists or small groups
 - PRACE allocated 16B hours in total in 10 years, ATLAS uses 3B/year
 - We cannot expect large allocations in near future

EuroHPC (1/4)

See [[this talk](#)] of A. Filipcic at Jamboree meeting on 5th March and [[more detailed talk](#)] at WLCG 2019 meeting



EuroHPC
Joint Undertaking

<https://eurohpc-ju.europa.eu/>

A legal and funding agency

- 27 Participating States + EU
- Site: Luxembourg
- Budget: ~1 B€ (half from EU)
- Operational: 11/2018 to 2026

Mission: Establish an integrated world-class supercomputing & data infrastructure and support a highly competitive and innovative HPC and Big Data ecosystem

EuroHPC JU
EuroHPC JU Participating States

EuroHPC JU
Participating States

Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and Switzerland.



Sweden joined as 26th country a few weeks ago, Switzerland as 27th last Friday

EuroHPC (2/4)

- Total budget ~400M euro, 50% contribution from member states, some coming out Q1-Q3 2019
- Extreme Scale Technologies and Applications (90M)
 - Extreme scale computing technologies (hardware, software, methods and algorithms for key applications)
 - HPC applications to ensure European leadership
- Widening HPC skills and use (30M)
 - Increase the knowledge and human capital and upraise HPC capabilities, including through the creation of national HPC Competence Centres and their networking and coordination across the Union
 - Federating European supercomputing resources
 - Support to SMEs
- European Processor Initiative (40M)
- Competence centers

2019-2020

1. Acquire an integrated world-class supercomputing infrastructure

EU budget
≥ 270 M€

- Pre-exascale & petascale machines accessible via PRACE

2. R&I for a competitive HPC/BD ecosystem

EU budget
≥ 180 M€

- Calls for R&I: exascale technologies and systems (incl. low-power processor); applications and use; training and skills

Access to the JU machines:

- free of charge for publicly funded R&I
- allocation of access time based on calls and peer reviews
- up to 20% of access time for pay-per-use commercial services

Existing PRACE HPC infrastructure



Distributed Supercomputing Infrastructure

26 members, including

5 Hosting Members

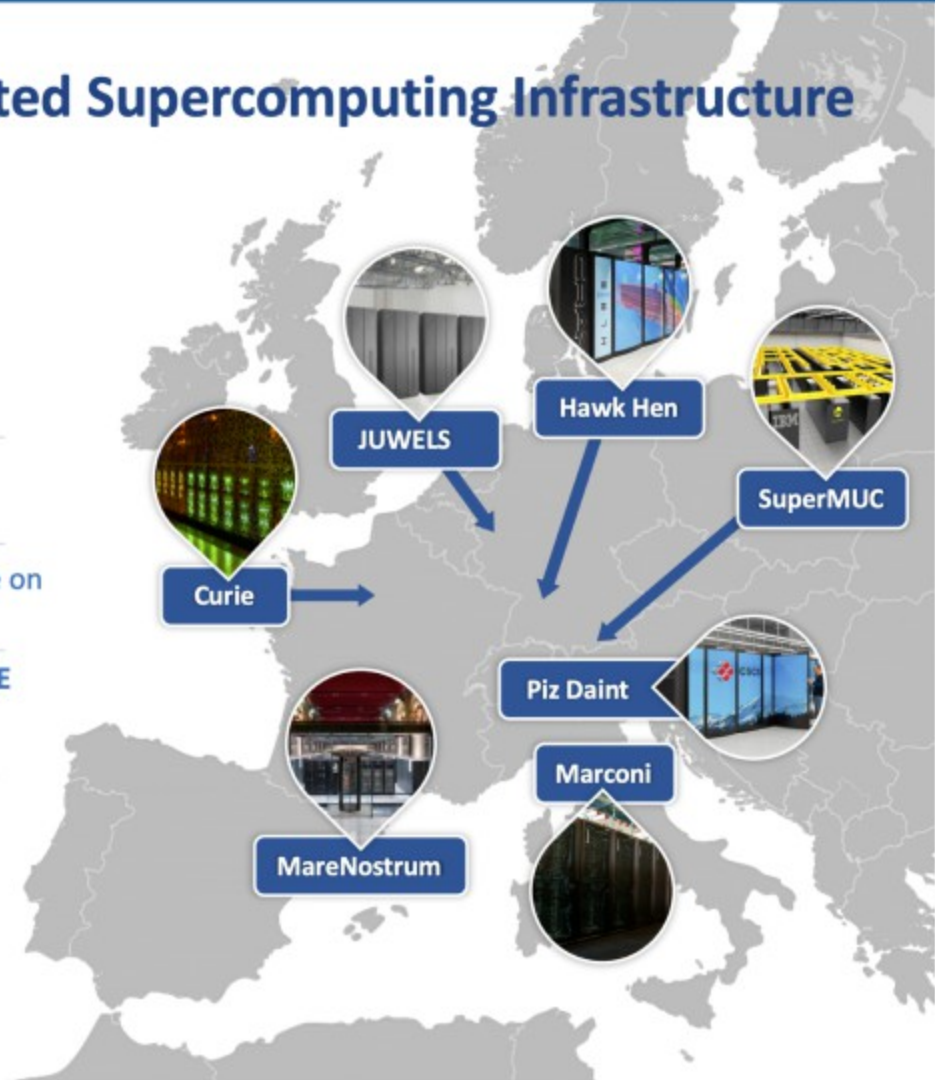
(Switzerland, France, Germany,
Italy and Spain)

652 scientific projects
enabled

110 PFlops/s of peak performance on
7 world-class systems

>12.000 people trained by 6 PRACE
Advanced Training Centers and
others events

Access prace-ri.eu/hpc-acces



EuroHPC (4/4)

- Significant funding from most members states
 - Some in procurement
 - Some in-kind, manpower
 - WLCG funding will be affected
- Not clear yet what architecture, possibly a mixture of
 - CPU partitions
 - GPU + FPGA partitions
 - To address intense CPU compute, AI and BigData processing
- The HPC machines need to be on top of TOP500 list
 - Uniform CPU+GPU machines are not at all excluded (Summit), although the CPUs will likely be x86_64

- HPCs will be connected with each other:
 - Big machines will be built by consortia of many countries
 - Opportunity for ATLAS distributed computing institutions to participate in R&I calls, especially data management
- 50% of resources will be managed by JU, 50% by participating countries

- EuroHPC pre-exascale and peta-scale computers will be available end of 2020
 - Some resources even in 2019

CERN/PRACE

- **Workshop CERN/PRACE** Oct. 2018 [[link](#)]
 - look also at conclusions [[link](#)]
 - conclusions at WLCG management board [[link](#)]