UL HPC Newsletter

Editors:	Dr. S. Varrette and Prof. P. Bouvry
Authors:	**The UL HPC Team** *aka:*
	↪ S. Varrette, P. Bouvry (Managers)
	↪ H. Cartiaux, V. Plugaru (FSTC)
	↪ S. Diehl (LCSB)
Date:	2016-02-02
Version:	v1.1
	http://twitter.com/ULHPC
	http://github.com/ULHPC
	hpc-sysadmins@uni.lu

In this issue…

Welcome to issue #2 of the UL HPC Newsletter.

📡 In the information and computational science age, most university research units and centers need to perform complex computations and process very large amounts of data. The UL HPC facility encompasses a High Performance Computing environment and high density data storage.

This second issue of the UL HPC newsletter is the occasion to review some statistics of the platform and its current capacities, while detailing the milestones completed in 2015, whether in terms of hardware or software developments. We will also come back to the two HPC Schools organized last year, and offer testimonials of achievements obtained on our infrastructure. Moreover, the planned developments for 2016 will be introduced. Many exciting challenges are waiting for us ahead, among which a new data center with cutting-edge cooling technologies, new official missions and mandates from the rectorate and the business opportunities around the HPC services we deliver since 2007.
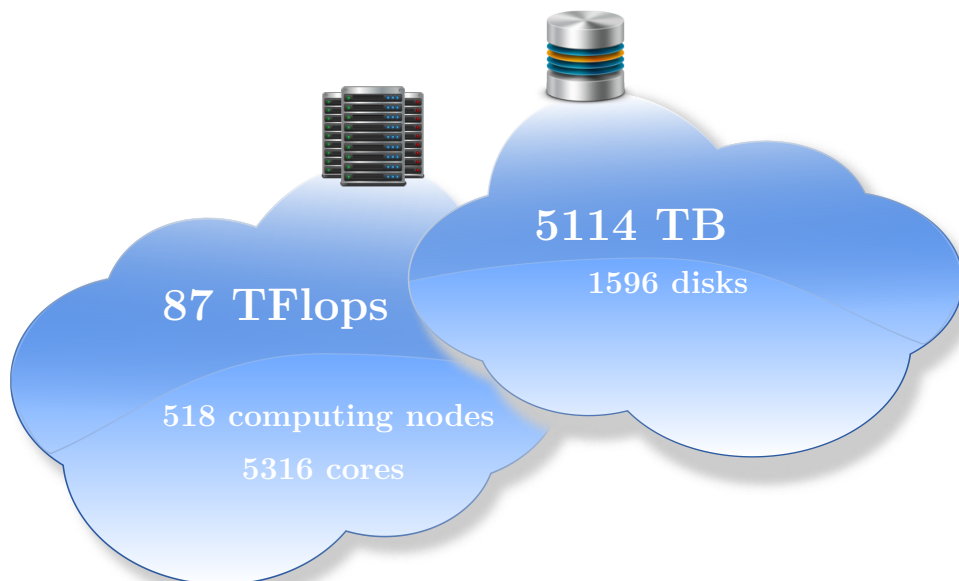
Finally, you have probably noticed the recent interest around the HPC and Big Data thematic at the national and international level (through articles in Wort.lu around the ICPEI initiatives for instance). The UL remains deeply involved at this level and we will keep you updated in the coming months as regards the developments on these topics, including our incoming implication in PRACE.

Enjoy the reading of this issue, we welcome your comments and questions by mail, on Twitter or GitHub.

Dr. S. Varrette and Prof. P. Bouvry

💬 UL HPC at a glance

87 TFlops

5114 TB

1596 disks

518 computing nodes

5316 cores

➡ http://hpc.uni.lu

📊 Platform Statistics

One of the UL HPC datacenters (in Belval Biotech 1), operating at full capacity.

The UL HPC platform has continuously evolved since the initial computing servers ordered in 2006. The HPC equipment is based on both classical Commercial off-the shelf (COTS) and specialized products.

> The UL HPC features as of January 2016, a computational power of **87 TFlops** (5316 computing cores) and **5.1 PBytes** for storage (incl. 1.7 PB for backups).

The bridge between computing and storage is a specificity of the UL service, with reasons both legal (certain data must remain in the country) and performance related.

As of January 2016, **335 users** are registered and active on the platform. The used computing time in 2015 by the registered users is measured at **2255 CPUYears**, thus **performed in a single year** period. This confirms the increasing usage of the platform over the years (see below) and demonstrates our commitment to excellence and continuous improvement to accelerate scientific discovery and innovation within the University (see also the list of publications generated thanks to the UL HPC Platform).
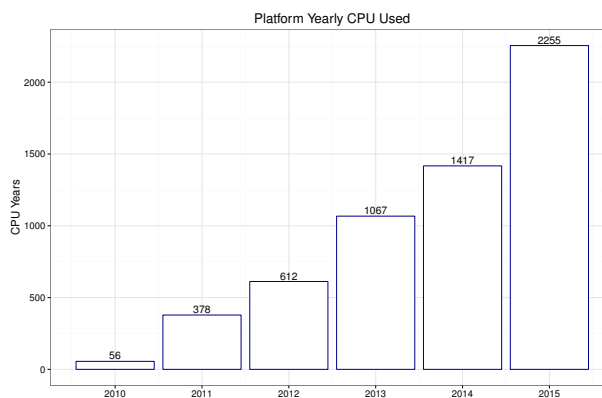
In terms of cumulative hardware investments since 2007 (excluding server rooms costs), we now reach a total of **5.244 M€**. This continuous effort makes the UL HPC platform **the largest facility of this type in Luxembourg** (after GoodYear industrial R&D Center) and is competitive with similar centers in the Grande-Region, as shown in the next table.

➡ http://hpc.uni.lu

UL HPC Newsletter

Editors:	Dr. S. Varrette and Prof. P. Bouvry
Authors:	**The UL HPC Team** *aka:*
	↪ S. Varrette, P. Bouvry (Managers)
	↪ H. Cartiaux, V. Plugaru (FSTC)
	↪ S. Diehl (LCSB)
Date:	2016-02-02
Version:	v1.1
	http://twitter.com/ULHPC
	http://github.com/ULHPC
	hpc-sysadmins@uni.lu

Country	Institute	#Nodes	(CPU) #Cores	TFlops R_{peak}	TB (Shared) Storage
Luxembourg	**UL HPC** (Uni.lu)	**518**	**5316**	**87**	**5114**
	LIST	58	800	6.21	144
France	LORIA (G5K), Nancy	320	2520	26.98	82
	ROMEO, Reims	174	3136	49.26	245
Belgium	NIC4, University of Liège	128	2048	32.00	20
	Université Catholique de Louvain	112	1344	13.28	120
	UGent / VSC, Gent	440	8768	275.30	1122
Germany	bwGrid, Heidelberg	140	1120	12.38	32
	bwForCluster, Ulm	444	7104	266.40	400
	bwHPC MLS&WISO, Mannheim	604	9728	371.60	420

📢 What happened in 2015?

HPC as part of the UL Digital Strategy

Since his arrival at the head of the University, our new president Rainer Klump aimed at defining an ambitious Digital Strategy for the university that would make it a model European Research University of the 21st century. In this context, the president confirms the UL HPC as a key supporting infrastructure for this strategy. To this end:

- Prof. Pascal Bouvry has been officially appointed Senior advisor for the president (*Chargé de mission auprès du recteur*) as regards the HPC strategy;
- The UL HPC budget has been significantly increased to match the proposed strategy (approved by the rectorate in April 2015). The responsibility for this budget has been transferred from Prof. Franck Leprévost to Prof. Pascal Bouvry.

This definitively rewards our continuous efforts, and especially the ones of the former vice-president Prof. Franck Leprévost, to give to HPC the necessary recognition and means (*i.e* budget) at the highest levels. We are thus very thankful for his restless support, vision and steering over the last years (that permitted to bring the effort of the University to develop the HPC facility from 250 K€ to 2 M€ per year) and wish him the best in the next challenges of his career.

The coming months will bring additional changes, especially as regards the HPC governance. All changes in the management structure will be communicated immediately.

New Computing HW

2015 has seen several increases of the computing capacity of the UL HPC platform with several additions to the Gaia cluster which **more than doubled its computing power**.

System Type	Vendor	Model	#Nodes	TFlops R_{peak}	GB RAM
Large memory system	Delta	D88x-M8-BI	1	2.4	3072
Large memory system	SGI	UV-2000	1	3	4096
General purpose HPC	Dell	FX2s/FC430	24	23.04	3072
Big Data analytics	HP	Moonshot 1500/m710	90	10.39	2880
Accelerator systems	Dell	C4130 (Tesla K80 GPU)	4 (+16)	3.84 (+46.56)	n/a
Accelerator upgrade	NVIDIA	Tesla K40m GPU	(+10)	(+ 11.7)	n/a
Total computing additions: 120 (+26 GPUs)				**100.93**	

➡ **http://hpc.uni.lu**

UL HPC Newsletter

Editors:	Dr. S. Varrette and Prof. P. Bouvry
Authors:	**The UL HPC Team** *aka:*
	↪ S. Varrette, P. Bouvry (Managers)
	↪ H. Cartiaux, V. Plugaru (FSTC)
	↪ S. Diehl (LCSB)
Date:	2016-02-02
Version:	v1.1
	http://twitter.com/ULHPC
	http://github.com/ULHPC
	hpc-sysadmins@uni.lu

Large memory (3 and 4 TB RAM) Delta system (above, dedicated to the RUES research unit) and SGI UV 2000 (right, dedicated to LCSB) allowing in-memory processing of huge datasets.
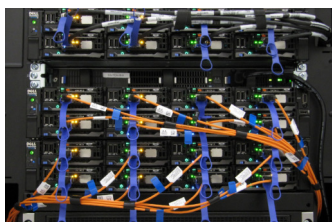
Three ultra-dense Dell FX2s enclosures (left) with state-of-the-art Haswell processors, highly improve the time to solution of tightly coupled calculations of the PHYMS research unit. Scalable software such as CP2K, ABINIT, QuantumESPRESSO, YAMBO, VASP and many others are able to use the performant distributed resources and speed up the computing time.

We also added the following components:

- Two dense, energy efficient enclosures from HP (bottom, left), based on the Haswell architecture completed with the Crystalwell package (a large 128MB L4 cache) are ideal for a variety of BigData analytics workloads performed by LCSB or SnT and economics analyses of the LSF.
- New ultra-dense Dell C4130 GPU nodes (bottom, right) containing 16 NVIDIA Tesla K80 accelerators (32 GK210 GPUs), in addition to an upgrade with 10 Tesla K40m cards for the Dell R720 systems.

http://hpc.uni.lu

UL HPC Newsletter

GPFS & Isilon Storage

In spring 2015, the main storage infrastructure of the Gaia cluster has been migrated from NFS to GPFS.

In technical terms, four NetApp E5400 disk enclosures providing **960 TB** of raw storage were attached to four pairs of storage nodes (shown right) that granted redundancy and failover, all connected to Gaia's 40Gbps QDR InfiniBand network. The transition of user home and work directories (totalling **43 TB**) and group projects (totalling **336 TB**), was successfully performed at the end of April 2015, with no changes required in the user workflows. The **performance** obtained on the new storage setup **increased by a factor of 2** – GPFS sustaining **more than 6GB/s** for parallel operations.

The new Dell/EMC Isilon system (shown left), commonly funded from SIU, LCSB and HPC budget lines, also came online last spring and now provides a large **1850 TB** shared storage space for all members of the UL. The OneFS based system integrates 16 storage nodes in two performance tiers and is designed for scalability, resiliency and operational flexibility. Projects requiring BigData storage, currently available on GPFS with redundant copies on the Isilon, can migrate between the two systems based on performance and access requirements. All in all, the current storage capacities available on the sole Gaia cluster are summarized in the table below.

Filesystem	Usage	#Enclosures	#Disks	Raw Capacity (TB)
Gaia GPFS	Home/Work	4	240	960 TB
Gaia Lustre	Scratch	4	200	480 TB
Isilon (OneFS)	Projects	16	768	1944 TB
Total:		**24**	**1462**	**3384 TB**

OS/Software Upgrades

As part of our commitment to constantly improve the services offered by the HPC platform, in 2015 we have performed the most comprehensive software upgrades in the 8 years existence of UL HPC.

A staged upgrade of the OS on 280 computing systems from Debian 6 to Debian 7 has been done on both Chaos and Gaia clusters. 55 management systems and 101 hosted Virtual Machines have also undergone OS upgrades, allowing for increased stability and uninterrupted security support of these critical systems.
The Subnet Manager of Gaia, providing routing across the InfiniBand fabric, has been configured in redundant mode for increased availability of this high speed network.

➦ http://hpc.uni.lu

UL HPC Newsletter

Editors: Dr. S. Varrette and Prof. P. Bouvry
Authors: **The UL HPC Team** *aka:*
↪ S. Varrette, P. Bouvry (Managers)
↪ H. Cartiaux, V. Plugaru (FSTC)
↪ S. Diehl (LCSB)
Date: 2016-02-02
Version: v1.1
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

The databases of the OAR Resource and Job Management System on Chaos and Gaia, containing the complete traces of **5 million user jobs** , have been migrated from MySQL to PostgreSQL.

The software environment available through the Environment Modules system has been regenerated for the new underlying OS and embedded in the new RESIF framework.
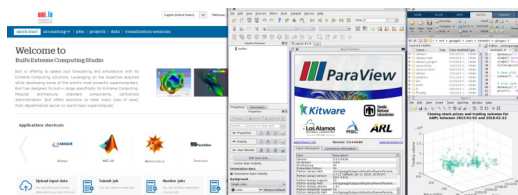
Finally, the XCS web portal for visualisation (see left, running remote instances of ParaView and Matlab) underwent an upgrade to the latest version available, with improved usability and security. Access to the portal, which uses the Gaia cluster's GPGPUs for 3D acceleration, is available to all UL HPC users upon request.

Complementary services

In the second half of 2015, we have opened access to a new online collaboration service: GitLab.uni.lu (right). Like Github, Gitlab features **advanced Git repository management**. The GitLab service is immediately available to UL HPC platform users with their account and to their external collaborators that have a GitHub account.
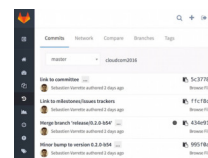
Our long-running GForge.uni.lu collaboration system (featuring static web hosting for projects, Git or Subversion repositories etc.) has recently been upgraded. Access to GForge is through dedicated accounts, (requested by mail: `admin@gforge.uni.lu`), for both UL members and external partners.

For easy, **Dropbox-like**, file sharing (and also since we were not convinced by the DropIT system) we have deployed an OwnCloud.uni.lu instance (right). OwnCloud is available on UL premises (or through the VPN) for UL HPC users.
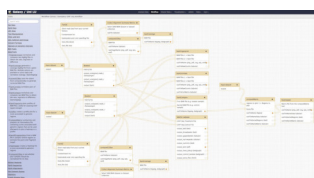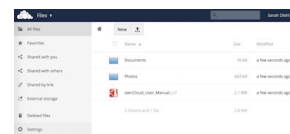
The Galaxy portal (`http://galaxy-server.uni.lu`) was also updated (left). This web-based platform provides a simplified interface to many popular **bioinformatics tools and generation of reproducible workflows**. Many more tools are now available, especially for the analysis of ChIP and RNA sequencing and variant data.

2015 HPC Schools

Two HPC schools have been organized in 2015 by the HPC team, comprising tutorials and practical sessions meant to promote best practices in the usage of the platform.

➥ `http://hpc.uni.lu`

UL HPC Newsletter

In **March**, 35 UL staff and students registered to the newcomer sessions meant to cover foundation topics. In **June**, the yearly two-day workshop has been organized together with leading computational scientists of the UL and HPC technologists. The workshop offered a broad set of sessions focusing on both foundation and advanced topics:

- Access to and interaction with the UL HPC infrastructures;
- HPC challenges, especially as regards (big)data and storage management;
- HPC workflow management for sequential and parallel tasks;
- HPC programming and usage of the main software and tools available on the platform (Matlab, R, physics, chemistry, bioinformatics tools, MPI implementations) and services (Galaxy) using the platform.
- Scientific visualization and software environment management.

48 UL staff and students registered to this summer school, with the participants entitled to 1 ECTS credit offered by the Doctoral School of Computer Science and Computer Engineering (DS-CSCE) upon the successful completion of the sessions.

🕒 What's next?

Centre de Calcul (CDC) in Belval

By the end of 2016, a new supercomputing site will be implemented at the Belval campus, in the House of Knowledge's Computing Center (*Maison du Savoir - Centre de Calcul*, shown left). Located underground (CDC S-02) below the floor already deployed to host Restena, SIU and research equipment, the new site will feature up to 5 server rooms dedicated to the hosting of the future UL HPC platform (2 for pure HPC equipment, 2 for storage and 1 hybrid). In line with the continuous efforts to build a green and sustainable campus, the future UL HPC site will feature cutting-edge cooling technologies based on Direct Liquid Cooling (DLC). Designed in a step-wise approach, the implementation of the new UL HPC platform will allow to reach a capacity of **3 PFlops** (for computing) and **20 PB** (for storage) by 2020, thus supporting and advancing the next generation science in Luxembourg and abroad.

IPCEI-HPC-BDA

You might have noticed a recent change in the global perception of HPC around you. Too often neglected in the past, EU acknowledged recently the "*necessity to act and invest in the development and deployment of HPC technology, Big Data and Applications. Failure to do so will seriously undermine European competitiveness and Europe will miss an important opportunity to ensure its industries compete on a global level*" (see Blog post of EU Commissioner on Digital Economy & Society G. H. Oettinger).

As a consequence, A European Strategic Positioning Paper was released to define an *Important Project of Common European Interest (IPCEI)* on HPC and Big Data

UL HPC Newsletter

Editors: Dr. S. Varrette and Prof. P. Bouvry
Authors: **The UL HPC Team** *aka:*
↪ S. Varrette, P. Bouvry   (Managers)
↪ H. Cartiaux, V. Plugaru   (FSTC)
↪ S. Diehl   (LCSB)
Date: 2016-02-02
Version: v1.1
 http://twitter.com/ULHPC
 http://github.com/ULHPC
 hpc-sysadmins@uni.lu

Enabled Applications (IPCEI-HPC-BDA) by the Luxembourg government. The project was officially launched on November 17th 2015 at the European Data Forum together with France, Italy and Spain – an article was published in Wort at this occasion.

In this context, the UL, together with the LIST and Luxinnovation will support the Ministry of Economy to provide an HPC and Big Data Enabled Applications implementation roadmap to the European Council and European Commission by September 2016.

Selected Use Cases
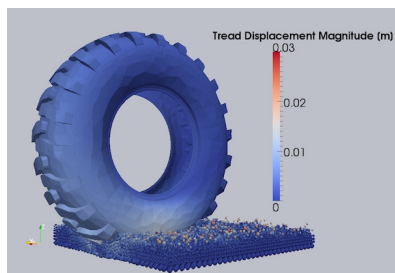
Snow-tire interaction - LuXDEM [2]

The LuXDEM research group is developing simulation approaches that can describe the interaction of a tire surface on a snow-covered road. To predict the viscoelastic deformation of the tire when it comes into contact with snow a Finite Element analysis is applied. The coupling between the discrete approach to characterize snow and the FEM approach for the tire represents an accurate model to assess the traction. The predicted results obtained by the simulation tool are compared to experimental data for validation.

The Soft Matter Theory group, part of PhyMS research unit, continues to explore means to understand and control the processes which allow matter of all kinds (incl. biological) to respond to stimuli and provide new material properties. For instance, it has been known since the seventies that high salt causes DNA to reverse its twist, from a right-handed spiral (called *B*-form) to a left-handed spiral (called *Z*-form) – see right picture. However the subtlety of DNA's interaction with the surrounding water and salt is such that a quantitative theory of this transition has remained elusive. Thanks to the UL HPC resources, it has been possible for the first time to run low-level models with a sufficient level of sampling to give an accurate description of the high-level process of salt-driven untwisting and re-twisting.
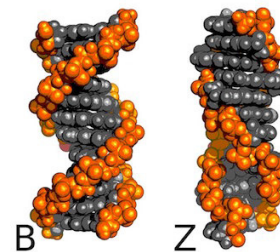
Atomistic DNA in High Salt - PhyMS [3]
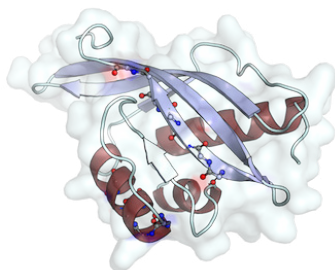
Evolutionary Algorithms for the Inv. Protein Folding Problem - PCOG/LCSB [4]

Protein structure prediction is an essential step in understanding the molecular mechanisms of living cells, with widespread application in biotechnology and health. The Inverse Folding Problem (IFP) of finding sequences that fold into a defined structure, an important research problem at the heart of rational protein design, is being tackled at the PCO Group in partnership with the LCSB.

 http://hpc.uni.lu

UL HPC Newsletter

Editors: Dr. S. Varrette and Prof. P. Bouvry
Authors: The UL HPC Team *aka*:
↪ S. Varrette, P. Bouvry (Managers)
↪ H. Cartiaux, V. Plugaru (FSTC)
↪ S. Diehl (LCSB)
Date: 2016-02-02
Version: v1.1
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

A multi-level UAV swarm simulator is under development at the SnT, with the objective of tackling three real application scenarios: surveillance, acquisition and tracking. The cooperation between levels is a new approach which allows to compensate weaknesses of a single level of swarms by taking advantage of the abilities of each level. The core of this simulator is based on Game Theoretical aspects and co-evolutionary optimization algorithms.

Featured research: UL HPC big data ranking records

In the context of the GDRI-Algodec *Algorithmic Decision Theory* supported by the CNRS and the FNR, Prof. R. Bisdorff is developing multicriteria ranking algorithms for large sets of potential decision alternatives: up to several thousand of alternatives evaluated on multiple incommensurable ordinal performance criteria.

This research is motivated by the development of a visualization tool - a *heat map* - for performance tables showing the decision alternatives linearly ordered form the best to the worst, and the individual performances colored by quantiles equivalence classes. By using Python3.5 multiprocessing resources and the Digraph3 multicriteria software library, it is possible, on the ULHPC `gaia-80` machine using

Heat map example [1]

120 single threaded cores and 1.5 TB of memory, to linearly rank (without ties) in less than an hour a huge set of 250000 alternatives evaluated on 21 performance criteria by balancing economic, ecological and societal decision objectives. Data input is, on the one side, a 250000x21 performance tableau of 825 MB, and on the other side, a theoretical outranking space consisting of 62 000 000 000 (billions) of pairwise comparisons. A "small" set of 1000 decision alternatives, in a similar setting, may thus be ranked typically in less than 5 seconds.

[1] S. Wagle, M. Guzek, P. Bouvry, R. Bisdorff. An Evaluation Model for Selecting Cloud Services from Commercially Available Cloud Providers. In *Proc. of the 7th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2015)*, Vancouver, Canada, December 2015.
[2] M. Michael, F. Vogel, B. Peters. DEM-FEM coupling simulations of the interactions between a tire tread and granular terrain. *J. Comput. Methods Appl. Mech. Engrg. 2015*.
[3] J. Berryman, T. Schilling. Free Energies by Thermodynamic Integration Relative to an Exact Solution, Used to Find the Handedness-Switching Salt Concentration for DNA. *J. Chem. Theory Comput., 2013*.
[4] S. Nielsen, G. Danoy, W. Jurkowski, J. L. J. Laredo, R. Schneider, E.-G. Talbi, P. Bouvry. A Novel Multi-objectivisation Approach for Optimising the Protein Inverse Folding Problem. In *Proc. of the 18th European Conference on Applications of Evolutionary Computation (EvoApplications 2015)*, Springer.

http://hpc.uni.lu

UL HPC Newsletter

Issue 2 – January 2016

👤 Meet the team

Management

Pascal Bouvry is a full professor of the FSTC and the head of the ILIAS laboratory and the DS-CSCE doctoral school. He is also Senior advisor for the president (*Chargé de mission auprès du recteur*) as regards the HPC strategy. His team (PCOG) is composed of 25 researchers working on Parallel computing and Optimization applied to Cloud Computing and HPC (scheduling, energy-efficiency, security), Ad-Hoc Networks (Vanets simulation and service optimization) and Biology (gene sequencing, regulatory networks, protein folding).

Sébastien Varrette, PhD, is a Research Associate in Prof. Bouvry's team since 2007. Along with Prof. Bouvry, he defined and set up the global HPC initiative of the UL in 2007. In this context, he is managing the sysadmin team that maintains and extends the platform. In parallel, his research work focuses on Distributed Computing Platforms (clusters, grids or clouds), with a particular interest on the security and performance evaluation of distributed or parallel executions.

FSTC

Hyacinthe Cartiaux joined the HPC team in 2011 to set up the Grid'5000 Luxembourg site and has since been involved with all the HPC infrastructure of the UL, and other external services such as the Gforge. His interests cover IT automation and devops techniques, HPC & Grid Computing.

Valentin Plugaru is an HPC engineer part of the HPC team since 2014. Beginning with 2012 he has collaborated with Prof. Bouvry's team on research in Energy Efficiency and Performance Evaluation of HPC/Cloud environments. His general interests span R&D in High Performance Computing, Grid and Cloud Computing.

LCSB

Sarah Diehl is a bioinformatician and joined the LCSB BioCore in 2015 as an HPC systems administrator. Her goal is to bridge the gap between researchers and IT specialists. She is experienced in data management, next-generation sequencing analysis and development of analysis pipelines.

➡ http://hpc.uni.lu