UL HPC Newsletter

Issue 1 – March 2015

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

In this issue...

Welcome to issue #1 of the UL HPC Newsletter.

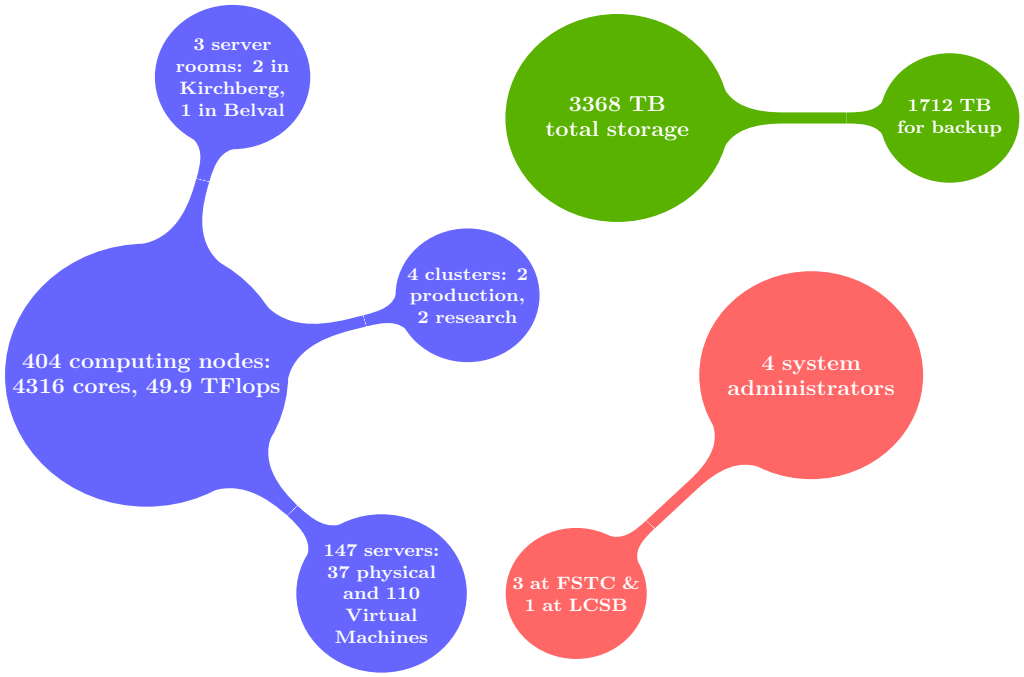
As part of the UL HPC quality plan, we are proud to announce the release of a UL HPC newsletter. The objective is to keep you informed of the latest advances and news as regards the UL HPC platform.
So you hold the very first issue where we will:

- review 2014 activities and achievements;
- outline the HPC platform upgrade plans for the coming weeks and months, whether in terms of **storage** (GPFS, Isilon, GlusterFS), **computing** (introducing new systems from SGI, Delta and HP) or **services** such as the new Galaxy portal or the RESIF framework able to facilitate the management of the software built on top of the platform and available as Environment modules.

Finally, the 2015 editions of the UL HPC school will be detailed.
We hope that this initial newsletter is informative and interesting for you, and we welcome your comments and questions by mail, on Twitter or GitHub.

Dr. S. Varrette and Prof. P. Bouvry

The UL HPC Platform at a glance



http://hpc.uni.lu

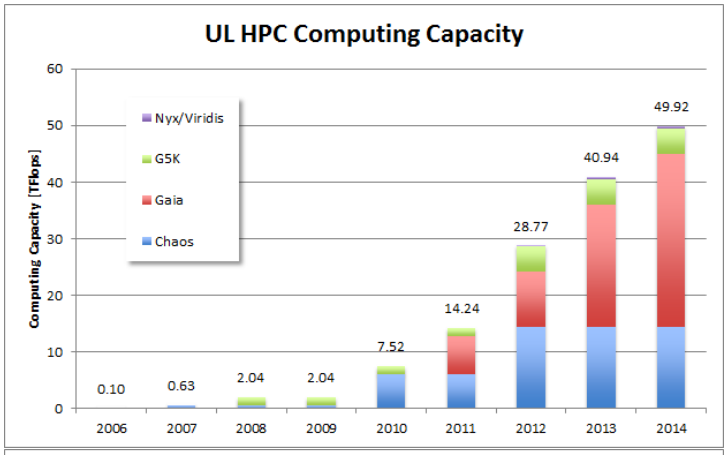
Editors: Dr. S. Varrette and Prof. P. Bouvry
Authors: **The UL HPC Team** *aka*:
↪ S. Varrette, P. Bouvry (Managers)
↪ H. Cartiaux, V. Plugaru (FSTC)
↪ S. Diehl (LCSB)
Date: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

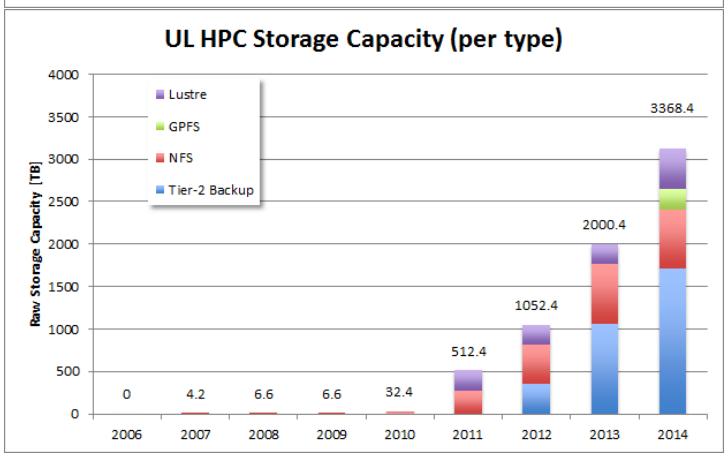
UL HPC Statistics

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 march 2015, **49.9 TFlops** for computations (over 4316 computing cores) and **3 PetaBytes** for storage (incl. 1.7 PB for backups).

The evolution of both computing and storage capacity, is depicted in the below figures.





As of March 2015, **281 users** are registered and active on the platform. The used computing effort in 2014 by the registered users is measured as **1407 CPUYears** , thus performed in a single year period.

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

What's new?

OS/Software Upgrades

As part of our commitment to constantly improve the services offered by the HPC platform, we will proceed with an upgrade of the underlying software running the platform during Q1 and Q2 2015.

This includes the upgrade of the Operating System (OS) of all components to Debian 7 (Wheezy), but also the upgrade of the software offered via modules/Lmod with the new RESIF framework.
To be as seamless as possible, these upgrades will be performed in a step-wise approach:

1. half of gaia (`gaia-83` to `gaia-154`) will be migrated to Debian 7 on **April 6th**, to benefit from the day off. This phase is intended to last **three weeks**, giving enough time for everybody to adapt and test their workflow on the future configuration;
2. the rest of gaia will be migrated to Debian 7 on **April, 27th**. Also, the home directories of Gaia will be migrated from the current NFS storage server to a distributed GPFS setup;
3. the chaos cluster will be upgraded following gaia, *i.e.* on **April 28th**.

The HPC team will strive to minimize the downtime of the clusters, with access always being kept open to either Gaia or Chaos.

GPFS & Isilon Storage

The main storage system of the Gaia cluster is under reorganization, with a planned change of the underlying filesystem of the home directories from NFS to GPFS. Also, most of the projects will be migrated to the newly acquired EMC Isilon 1PB system, based on their size and access requirements.
For users these transitions will be seamless, as their files will be synchronized with the new storage systems on the date of the migration (planned on **April 27th**).

Indeed, the main NFS server **Galactus** has been under high load since 2014, with the doubling of the number of computing nodes accessing it. As the limitations of this technology became more pronounced, a distributed setup with IBM's GPFS was chosen to replace it. The migration to GPFS will allow for better stability, speed and future storage extensions, thus improving the productivity of the users of this cluster.

As for the EMC Isilon system, it is the result of a public tender released in 2014 for the acquisition of a scalable and modular Network Accessible Storage (NAS) to serve both SIU and HPC needs. Commonly funded from SIU, LCSB and HPC budget lines, the chosen solution is now deployed and validated with the expectation of a production release to all UL users by June, 2015.

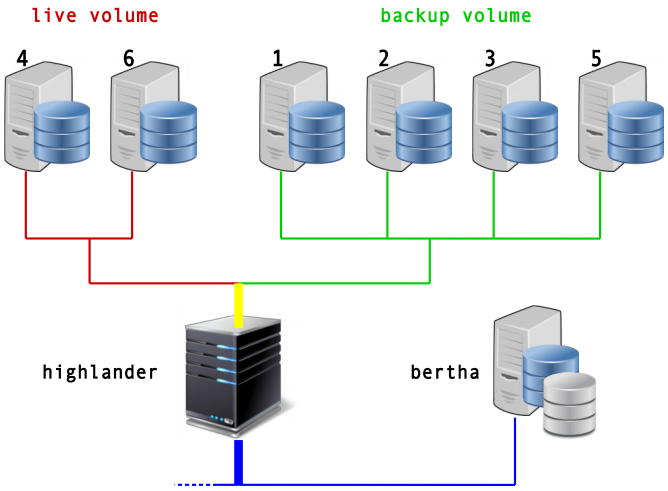
UL HPC Newsletter

Issue 1 – March 2015

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

LCSB storage consolidation

Biomedical research sees ever-increasing data storage requirements as both working datasets and old data (retained for reproducibility and legal reasons) grow. Data storage elements bought in the past years at the LCSB are currently being integrated in an uniform storage pool, allowing for easier usage, management and backup automation. In this context, Jean-François Le Fillâtre used to recycle before his departure the old Certon boxes into an unified storage solution merged under the GlusterFS filesystem according to the below architecture. In particular, the `highlander` server is designed to expose the storage areas (composed by six separate Certon systems totalling 944 TB raw storage, and the new `bertha` server featuring 648 TB raw storage) to the LCSB users. We need to follow up the setup if this system to make it officially released as soon as possible.



The Gluster filesystem (part of the RedHat Storage Server solution) has been chosen for this task following an indepth comparison of the modern distributed filesystems Ceph, OpenAFS, Hadoop HDFS, XtreemFS, Lustre, BeeGFS, GPFS and OCFS2.

New Computing HW

Delta server (left), SGI UV-2000 (center) and HP Moonshot (right)

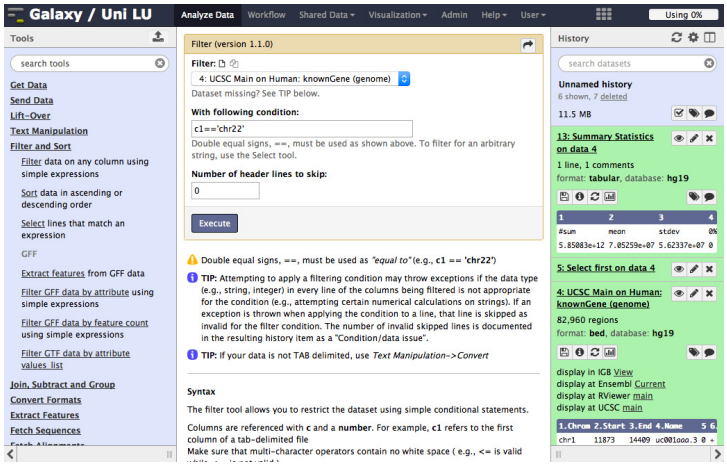
Editors: Dr. S. Varrette and Prof. P. Bouvry
Authors: **The UL HPC Team** *aka*:
↪ S. Varrette, P. Bouvry (Managers)
↪ H. Cartiaux, V. Plugaru (FSTC)
↪ S. Diehl (LCSB)
Date: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

New specialized computing systems will be part of the UL HPC platform in Q2 2015:

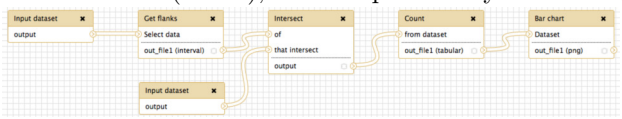
- Delta D88x-M8-BI (Gaia cluster: gaia-80) featuring 8 * Intel Xeon E7-8880v2 @ 2.5 GHz (120 cores: 2.4 TFlops) with 3TB of RAM and 12.6TB of local storage;
- SGI UV 2000 (Gaia cluster: gaia-81) featuring 16 * Intel Xeon E5-4650v2 @ 2.4 GHz (160 cores: 3 TFlops) with 4TB of RAM;
- HP Moonshot (Nyx cluster: moonshot-1): 30 blades each with an Intel Xeon E3-1284Lv3 @ 1.8 GHz and 32GB of RAM (120 cores: 3.4 TFlops and 960GB of RAM in total).

The Delta system will be used by RUES members for Computational Mechanics workloads, while the SGI UV-2000 system will be dedicated for Bioinformatics analyses of the LCSB. The HP Moonshot mini-cluster will be initially used by the HPC team as a testbed for new software solutions and for the evaluation of this low-power platform for HPC workloads.

The Galaxy Portal



Galaxy web interface (above), and sample Galaxy workflow (below)



Galaxy is an open, web-based platform for data intensive biomedical research. It allows users to easily perform, reproduce and share complete bioinformatics analyses. The Galaxy service has been integrated with the Gaia cluster, enabling many users to work in parallel and use the computing power of this cluster.

> UL users have access to Galaxy at `galaxy-server.uni.lu` with their HPC credentials. For more information on this service, contact us by mail.

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

RESIF: Software / Modules Management

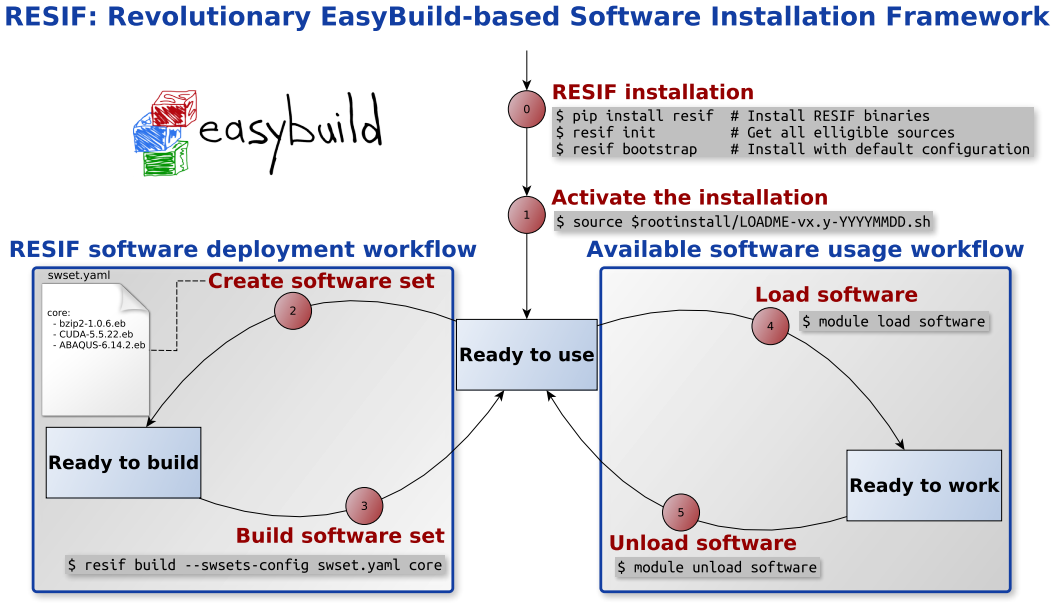
RESIF is an internally developed tool that aims to make the software management on the UL HPC platform more accessible to its users.
Its main feature is the automation of the deployment of a new software environment, while also making it easy to add software to an existing stack. The tool can be used both on the HPC platform and users' workstations, making it easy to have local installations of the same software in use on the clusters.

RESIF builds on the EasyBuild framework yet does not require an indepth knowlege of it, and provides users a simple workflow allowing:

- easy installation and initialization of RESIF and EasyBuild
- creation of a new software stack
- addition of new software to an existing software stack

The software environment available through the Modules system on the UL HPC platform will be regenerated with RESIF upon completion of the upgrade to Debian 7. This environment will contain a `core` set of applications used by the UL research groups and will be organized by theme, making it easy to find and use the required packages.

The typical workflow of the operations performed within RESIF is depicted in the below figure.



UL HPC Newsletter

Issue 1 – March 2015

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

2015 HPC Schools

In May 2014 the HPC Team successfully held the first edition of the UL HPC School which covered both basic and advanced usage of the HPC platform.

More than 60 participants from the FSTC, LCSB, SnT, FDEF and FLSHASE took part in the 8 practical sessions which spanned two days. For 2015, we are happy to announce the follow-up of this event.

For a better dissemination of the best practices among all UL users (including newcomers), the HPC Schools will be reproduced in two versions:

- training for newcomers, focusing on the basics, provided in half-day practical sessions twice a year (typically in March and October), to cover the arrival of new hires;
- the regular HPC School (a 2 days event covering all aspects including advanced ones) will be held once a year in June.

The first Newcomer Training Day took place on *March 13th, 2015* in the Kirchberg campus – see the official webpage for details.

The second edition of the regular HPC School will take place on **June, 18th-19th**, and will feature new and updated sessions compared with the previous event, based on its feedback. The full programme of this event will be announced in May.

Publications

In an attempt to promote the management operated on the UL HPC platform, an article has been published in the IEEE HPCS'2014 [1] – see the corresponding ORBilu entry.

It is also the occasion to recall that you are asked through the AUP to **acknowledge** your usage of the UL HPC platform in your research articles, and to cite the above mentioned paper if it is possible.

More importantly, you should mark your own ORBilu entries with the `ULHPC` tag using the guidelines proposed on the UL HPC website.

References

[1] S. Varrette, P. Bouvry, H. Cartiaux, and F. Georgatos. Management of an Academic HPC Cluster: The UL Experience. In *Proc. of the 2014 Intl. Conf. on High Performance Computing & Simulation (HPCS 2014)*, pages 959–967, Bologna, Italy, July 2014. IEEE.

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: 2016-02-01
Version: v1.2
http://twitter.com/ULHPC
http://github.com/ULHPC
hpc-sysadmins@uni.lu

Meet the team

Management

Pascal Bouvry is a full professor of the FSTC and the head of the IL-IAS research unit and the DS-CSCE doctoral school. 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 maintain and extend 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.