IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

Activity Report 2018

Project-Team AVALON

Algorithms and Software Architectures for Distributed and HPC Platforms

IN COLLABORATION WITH: Laboratoire de l'Informatique du Parallélisme (LIP)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Project-Team AVALON

Creation of the Team: 2012 February 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.1.13. - Virtualization
A1.2.1. - Dynamic reconfiguration
A1.3. - Distributed Systems
A1.3.4. - Peer to peer
A1.3.5. - Cloud
A1.3.6. - Fog, Edge
A1.6. - Green Computing
A2.1.6. - Concurrent programming
A2.1.7. - Distributed programming
A2.1.10. - Domain-specific languages
A2.5.2. - Component-based Design
A2.6.2. - Middleware
A2.6.4. - Ressource management
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.8. - Big data (production, storage, transfer)
A4.4. - Security of equipment and software
A6.2.7. - High performance computing
A7.1. - Algorithms
A7.1.1. - Distributed algorithms
A7.1.2. - Parallel algorithms
A8.2. - Optimization
A8.2.1. - Operations research
A8.2.2. - Evolutionary algorithms
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B1.1.7. - Bioinformatics
B3.2. - Climate and meteorology
B4.1. - Fossile energy production (oil, gas)
B4.2.2. - Fusion
B4.5. - Energy consumption
B4.5.1. - Green computing
B6.1.1. - Software engineering
B8.1.1. - Energy for smart buildings

B9.5.1. - Computer science
B9.7. - Knowledge dissemination
B9.7.1. - Open access
B9.7.2. - Open data
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Christian Perez [Team leader, Inria, Senior Researcher, HDR]
Marcos Dias de Assunção [Inria, Starting Research Position]
Thierry Gautier [Inria, Researcher]
Laurent Lefèvre [Inria, Researcher, HDR]
Frédéric Suter [CNRS, Researcher, HDR]

Faculty Members

Yves Caniou [Univ. Claude Bernard, Associate Professor]
Eddy Caron [École Normale Supérieure de Lyon, Associate Professor, HDR]
Jean-Patrick Gelas [Univ. Claude Bernard, Associate Professor]
Olivier Glück [Univ. Claude Bernard, Associate Professor]

Post-Doctoral Fellows

Jérôme Richard [Inria, until Mar 2018]
Cyril Seguin [Inria, from Apr 2018]

PhD Students

Dorra Boughzala [Inria]
Maverick Chardet [Inria]
Arthur Chevalier [Orange Labs]
Hadrien Croubois [École Normale Supérieure de Lyon, until Aug 2018]
Idriss Daoudi [Inria, from Oct 2018]
Alexandre Da Silva Veith [Univ de Lyon]
Jad Darrous [Inria]
Felipe Rodrigo de Souza [École Normale Supérieure de Lyon]
Zeina Houmani [École Normale Supérieure de Lyon, from Nov 2018]
Aurélie Kong-Win-Chang [Univ. Claude Bernard]
Valentin Lorentz [École Normale Supérieure de Lyon, until Sep 2018]
Issam Raïs [Inria, until Sep 2018]
Philippe Virouleau [Inria, until Feb 2018]

Technical staff

Mathilde Boutigny [Inria, Engineer, until Mar 2018, granted by BPIFRANCE FINANCEMENT SA]
Simon Delamare [CNRS, Engineer, 50%]
David Loup [Inria, Engineer]
Jean-Christophe Mignot [CNRS, Engineer, 20%]
Matthieu Imbert [Inria, Engineer, 40%]

Interns

Lucas Besnard [Inria, from Apr 2018 until Jun 2018]
Zeina Houmani [Inria, until Jul 2018]
Valentin Huguet [Univ. Claude Bernard, from Mar 2018 until Aug 2018]
Bastien Marsaud [CNRS, from Apr 2018 until Jun 2018]
Lukas Schmidt [Univ. Claude Bernard, until Apr 2018]

Administrative Assistants

Solene Audoux [Inria, Assistant, from Nov 2018]
Evelyne Blesle [Inria, Assistant]

2. Overall Objectives

2.1. Presentation

The fast evolution of hardware capabilities in terms of wide area communication, computation and machine virtualization leads to the requirement of another step in the abstraction of resources with respect to parallel and distributed applications. These large scale platforms based on the aggregation of large clusters (Grids), huge datacenters (Clouds) with IoT (Edge/Fog), collections of volunteer PCs (Desktop computing platforms), or high performance machines (Supercomputers) are now available to researchers of different fields of science as well as to private companies. This variety of platforms and the way they are accessed also have an important impact on how applications are designed (*i.e.,* the programming model used) as well as how applications are executed (*i.e.,* the runtime/middleware system used). The access to these platforms is driven through the use of multiple services providing mandatory features such as security, resource discovery, virtualization, load-balancing, monitoring, *etc.*

The goal of the AVALON team is to execute parallel and/or distributed applications on parallel and/or distributed resources while ensuring user and system objectives with respect to performance, cost, energy, security, *etc.* Users are generally not interested in the resources used during the execution. Instead, they are interested in how their application is going to be executed: the duration, its cost, the environmental footprint involved, *etc.* This vision of utility computing has been strengthened by the cloud concepts and by the short lifespan of supercomputers (around three years) compared to application lifespan (tens of years). Therefore a major issue is to design models, systems, and algorithms to execute applications on resources while ensuring user constraints (price, performance, *etc.*) as well as system administrator constraints (maximizing resource usage, minimizing energy consumption, *etc.*).

2.2. Objectives

To achieve the vision proposed in Section 2.1, the AVALON project aims at making progress to four complementary research axes: energy, data, component models, and application scheduling.

2.2.1. Energy Application Profiling and Modeling

AVALON will improve the profiling and modeling of scientific applications with respect to energy consumption. In particular, it will require to improve the tools that measure the energy consumption of applications, virtualized or not, at large scale, so as to build energy consumption models of applications.

2.2.2. Data-intensive Application Profiling, Modeling, and Management

AVALON will improve the profiling, modeling, and management of scientific applications with respect to CPU and data intensive applications. Challenges are to improve the performance prediction of parallel regular applications, to model and simulate (complex) intermediate storage components, and data-intensive applications, and last to deal with data management for hybrid computing infrastructures.

2.2.3. Resource-Agnostic Application Description Model

AVALON will design component-based models to capture the different facets of parallel and distributed applications while being resource agnostic, so that they can be optimized for a particular execution. In particular, the proposed component models will integrate energy and data modeling results. AVALON in particular targets OpenMP runtime as a specific use case.

2.2.4. Application Mapping and Scheduling

AVALON will propose multi-criteria mapping and scheduling algorithms to meet the challenge of automating the efficient utilization of resources taking into consideration criteria such as performance (CPU, network, and storage), energy consumption, and security. AVALON will in particular focus on application deployment, workflow applications, and security management in clouds.

All our theoretical results will be validated with software prototypes using applications from different fields of science such as bioinformatics, physics, cosmology, *etc.* The experimental testbeds GRID'5000 (cf Section 6.8) and Silecs (cf Section 6.10 will be our platforms of choice for experiments.

3. Research Program

3.1. Energy Application Profiling and Modeling

Despite recent improvements, there is still a long road to follow in order to obtain energy efficient, energy proportional and eco-responsible exascale systems by 2022. Energy efficiency is therefore a major challenge for building next generation large-scale platforms. The targeted platforms will gather hundreds of millions of cores, low power servers, or CPUs. Besides being very important, their power consumption will be dynamic and irregular.

Thus, to consume energy efficiently, we aim at investigating two research directions. First, we need to improve measurement, understanding, and analysis on how large-scale platforms consume energy. Unlike approaches [36] that mix the usage of internal and external wattmeters on a small set of resources, we target high frequency and precise internal and external energy measurements of each physical and virtual resource on large-scale distributed systems.

Secondly, we need to find new mechanisms that consume less and better on such platforms. Combined with hardware optimizations, several works based on shutdown or slowdown approaches aim at reducing energy consumption of distributed platforms and applications. To consume less, we first plan to explore the provision of accurate estimation of the energy consumed by applications without pre-executing and knowing them while most of the works try to do it based on in-depth application knowledge (code instrumentation [39], phase detection for specific HPC applications [44], *etc.*). As a second step, we aim at designing a framework model that allows interaction, dialogue and decisions taken in cooperation among the user/application, the administrator, the resource manager, and the energy supplier. While smart grid is one of the last killer scenarios for networks, electrical provisioning of next generation large IT infrastructures remains a challenge.

3.2. Data-intensive Application Profiling, Modeling, and Management

Recently, the term "Big Data" has emerged to design data sets or collections so large that they become intractable for classical tools. This term is most time implicitly linked to "analytics" to refer to issues such as data curation, storage, search, sharing, analysis, and visualization. However, the Big Data challenge is not limited to data-analytics, a field that is well covered by programming languages and run-time systems such as Map-Reduce. It also encompasses data-intensive applications. These applications can be sorted into two categories. In High Performance Computing (HPC), data-intensive applications leverage post-petascale infrastructures to perform highly parallel computations on large amount of data, while in High Throughput Computing (HTC), a large amount of independent and sequential computations are performed on huge data collections.

These two types of data-intensive applications (HTC and HPC) raise challenges related to profiling and modeling that the AVALON team proposes to address. While the characteristics of data-intensive applications are very different, our work will remain coherent and focused. Indeed, a common goal will be to acquire a better understanding of both the applications and the underlying infrastructures running them to propose the best match between application requirements and infrastructure capacities. To achieve this objective, we will extensively rely on logging and profiling in order to design sound, accurate, and validated models. Then, the proposed models will be integrated and consolidated within a single simulation framework (SIMGRID). This will allow us to explore various potential "what-if?" scenarios and offer objective indicators to select interesting infrastructure configurations that match application specificities.

Another challenge is the ability to mix several heterogeneous infrastructures that scientists have at their disposal (*e.g.,* Grids, Clouds, and Desktop Grids) to execute data-intensive applications. Leveraging the aforementioned results, we will design strategies for efficient data management service for hybrid computing infrastructures.

3.3. Resource-Agnostic Application Description Model

With parallel programming, users expect to obtain performance improvement, regardless its cost. For long, parallel machines have been simple enough to let a user program use them given a minimal abstraction of their hardware. For example, MPI [38] exposes the number of nodes but hides the complexity of network topology behind a set of collective operations; OpenMP [42] simplifies the management of threads on top of a shared memory machine while OpenACC [41] aims at simplifying the use of GPGPU.

However, machines and applications are getting more and more complex so that the cost of manually handling an application is becoming very high [37]. Hardware complexity also stems from the unclear path towards next generations of hardware coming from the frequency wall: multi-core CPU, many-core CPU, GPGPUs, deep memory hierarchy, *etc.* have a strong impact on parallel algorithms. Hence, even though an abstract enough parallel language (UPC, Fortress, X10, *etc.*) succeeds, it will still face the challenge of supporting distinct codes corresponding to different algorithms corresponding to distinct hardware capacities.

Therefore, the challenge we aim to address is to define a model, for describing the structure of parallel and distributed applications that enables code variations but also efficient executions on parallel and distributed infrastructures. Indeed, this issue appears for HPC applications but also for cloud oriented applications. The challenge is to adapt an application to user constraints such as performance, energy, security, *etc.*

Our approach is to consider component based models [45] as they offer the ability to manipulate the software architecture of an application. To achieve our goal, we consider a "compilation" approach that transforms a resource-agnostic application description into a resource-specific description. The challenge is thus to determine a component based model that enables to efficiently compute application mapping while being tractable. In particular, it has to provide an efficient support with respect to application and resource elasticity, energy consumption and data management. OpenMP runtime is a specific use case that we target.

3.4. Application Mapping and Scheduling

This research axis is at the crossroad of the AVALON team. In particular, it gathers results of the three other research axis. We plan to consider application mapping and scheduling addressing the following three issues.

3.4.1. Application Mapping and Software Deployment

Application mapping and software deployment consist in the process of assigning distributed pieces of software to a set of resources. Resources can be selected according to different criteria such as performance, cost, energy consumption, security management, *etc.* A first issue is to select resources at application launch time. With the wide adoption of elastic platforms, *i.e.,* platforms that let the number of resources allocated to an application to be increased or decreased during its execution, the issue is also to handle resource selection at runtime.

The challenge in this context corresponds to the mapping of applications onto distributed resources. It will consist in designing algorithms that in particular take into consideration application profiling, modeling, and description.

A particular facet of this challenge is to propose scheduling algorithms for dynamic and elastic platforms. As the number of elements can vary, some kind of control of the platforms must be used accordingly to the scheduling.

3.4.2. Non-Deterministic Workflow Scheduling

Many scientific applications are described through workflow structures. Due to the increasing level of parallelism offered by modern computing infrastructures, workflow applications now have to be composed not only of sequential programs, but also of parallel ones. New applications are now built upon workflows with conditionals and loops (also called non-deterministic workflows).

These workflows cannot be scheduled beforehand. Moreover cloud platforms bring on-demand resource provisioning and pay-as-you-go billing models. Therefore, there is a problem of resource allocation for non-deterministic workflows under budget constraints and using such an elastic management of resources.

Another important issue is data management. We need to schedule the data movements and replications while taking job scheduling into account. If possible, data management and job scheduling should be done at the same time in a closely coupled interaction.

3.4.3. Security Management in Cloud Infrastructure

Security has been proven to be sometimes difficult to obtain [43] and several issues have been raised in Clouds. Nowadays virtualization is used as the sole mechanism to allow multiple users to share resources on Clouds, but since not all components of Clouds (such as micro-architectural components) can be properly virtualized, data leak and modification can occur. Accordingly, next-generation protection mechanisms are required to enforce security on Clouds and provide a way to cope with the current limitation of virtualization mechanisms.

As we are dealing with parallel and distributed applications, security mechanisms must be able to cope with multiple machines. Our approach is to combine a set of existing and novel security mechanisms that are spread in the different layers and components of Clouds in order to provide an in-depth and end-to-end security on Clouds. To do it, our first challenge is to define a generic model to express security policies.

Our second challenge is to work on security-aware resource allocation algorithms. The goal of such algorithms is to find a good trade-off between security and unshared resources. Consequently, they can limit resources sharing to increase security. It leads to complex trade-off between infrastructure consolidation, performance, and security.

4. Application Domains

4.1. Overview

The AVALON team targets applications with large computing and/or data storage needs, which are still difficult to program, maintain, and deploy. Those applications can be parallel and/or distributed applications, such as large scale simulation applications or code coupling applications. Applications can also be workflow-based as commonly found in distributed systems such as grids or clouds.

The team aims at not being restricted to a particular application field, thus avoiding any spotlight. The team targets different HPC and distributed application fields, which brings use cases with different issues. This will be eased by our various collaborations: the team participates to the INRIA-Illinois Joint Laboratory for Petascale Computing, the Physics, Radiobiology, Medical Imaging, and Simulation French laboratory of excellence, the E-Biothon project, the INRIA large scale initiative Computer and Computational Sciences at Exascale (C2S@Exa), and to BioSyL, a federative research structure about Systems Biology of the University of Lyon. Moreover, the team members have a long tradition of cooperation with application developers such as CERFACS and EDF R&D. Last but not least, the team has a privileged connection with CC IN2P3 that opens up collaborations, in particular in the astrophysics field.

In the following, some examples of representative applications we are targeting are presented. In addition to highlighting some application needs, they also constitute some of the use cases we will use to valide our theoretical results.

4.2. Climatology

The world's climate is currently changing due to the increase of the greenhouse gases in the atmosphere. Climate fluctuations are forecasted for the years to come. For a proper study of the incoming changes, numerical simulations are needed, using general circulation models of a climate system. Simulations can be of different types: HPC applications (*e.g.,* the NEMO framework [40] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [46] for global climate modeling), or workflows (long term global climate modeling).

As for most applications the team is targeting, the challenge is to thoroughly analyze climate-forecasting applications to model their needs in terms of programing model, execution model, energy consumption, data access pattern, and computing needs. Once a proper model of an application has been set up, appropriate scheduling heuristics could be designed, tested, and compared. The team has a long tradition of working with CERFACS on this topic, for example in the LEGO (2006-09) and SPADES (2009-12) French ANR projects.

4.3. Astrophysics

Astrophysics is a major field to produce large volumes of data. For instance, the Large Synoptic Survey Telescope (https://www.lsst.org/lsst/) will produce 15 TB of data every night, with the goals of discovering thousands of exoplanets and of uncovering the nature of dark matter and dark energy in the universe. The Square Kilometer Array (http://www.skatelescope.org/) produces 9 Tbits/s of raw data. One of the scientific projects related to this instrument called Evolutionary Map of the Universe is working on more than 100 TB of images. The Euclid Imaging Consortium (https://www.euclid-ec.org/) will generate 1 PB data per year.

AVALON collaborates with the *Institut de Physique Nucléaire de Lyon* (IPNL) laboratory on large scale numerical simulations in astronomy and astrophysics. Contributions of the AVALON members have been related to algorithmic skeletons to demonstrate large scale connectivity, the development of procedures for the generation of realistic mock catalogs, and the development of a web interface to launch large cosmological simulations on GRID'5000.

This collaboration, that continues around the topics addressed by the CLUES project (http://www.clues-project.org), has been extended thanks to the tight links with the CC-IN2P3. Major astrophysics projects execute part of their computing, and store part of their data on the resources provided by the CC-IN2P3. Among them, we can mention SNFactory, Euclid, or LSST. These applications constitute typical use cases for the research developed in the AVALON team: they are generally structured as workflows and a huge amount of data (from TB to PB) is involved.

4.4. Bioinformatics

Large-scale data management is certainly one of the most important applications of distributed systems in the future. Bioinformatics is a field producing such kinds of applications. For example, DNA sequencing applications make use of MapReduce skeletons.

The AVALON team is a member of BioSyL (http://www.biosyl.org), a Federative Research Structure attached to University of Lyon. It gathers about 50 local research teams working on systems biology. Moreover, the team cooperated with the French Institute of Biology and Chemistry of Proteins (IBCP http://www.ibcp.fr) in particular through the ANR MapReduce project where the team focuses on a bio-chemistry application dealing with protein structure analysis. AVALON has also started working with the Inria Beagle team (https://team.inria.fr/beagle/) on artificial evolution and computational biology as the challenges are around high performance computation and data management.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Henri Casanova, Arnaud Legrand, Martin Quinson and Frédéric Suter. "SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation" received the "Best Paper Award" of the Workshop on Education for High-Performance Computing (EduHPC-18).
- Anchen Chai, Sorina Camarasu-Pop, Tristan Glatard, Hugues Benoit-Cattin and Frédéric Suter. "Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems" received the "Best Workshop Paper on Heterogenous Systems" of the 24th International European Conference on Parallel and Distributed Computing (EuroPar'2018).

BEST PAPERS AWARDS:

[15]
H. CASANOVA, A. LEGRAND, M. QUINSON, F. SUTER. *SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation*, in "EduHPC-18 - Workshop on Education for High-Performance Computing", Dallas, United States, November 2018, pp. 1-10, https://hal.inria.fr/hal-01891513

[17]
A. CHAI, S. CAMARASU-POP, T. GLATARD, H. BENOIT-CATTIN, F. SUTER. *Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems*, in "Europar 2018 - 24th International European Conference on Parallel and Distributed Computing ; Workshop HeteroPar 2018", Turin, Italy, Lecture Notes in Computer Science (LNCS), August 2018, forthcoming, https://hal.archives-ouvertes.fr/hal-01887369

6. New Software and Platforms

6.1. MAD

Madeus Application Deployer
KEYWORDS: Automatic deployment - Distributed Software - Component models - Cloud computing
SCIENTIFIC DESCRIPTION: MAD is a Python implementation of the Madeus deployment model for multi-component distributed software. Precisely, it allows to: 1. describe the deployment process and the dependencies of distributed software components in accordance with the Madeus model, 2. describe an assembly of components, resulting in a functional distributed software, 3. automatically deploy the component assembly of distributed software following the operational semantics of Madeus.
RELEASE FUNCTIONAL DESCRIPTION: Initial submission with basic functionalities of MAD

NEWS OF THE YEAR: Operational prototype.

- Participants: Christian Pérez, Dimitri Pertin, Hélène Coullon and Maverick Chardet
- Partners: IMT Atlantique - LS2N - LIP
- Contact: Hélène Coullon
- Publications: Madeus: A formal deployment model - Behavioral interfaces for reconfiguration of component models

6.2. DIET

Distributed Interactive Engineering Toolbox
KEYWORDS: Scheduling - Clusters - Grid - Cloud - HPC - Middleware - Data management.
FUNCTIONAL DESCRIPTION: Middleware for grids and clouds. Toolbox for the use and porting of intensive computing applications on heterogeneous architectures.

RELEASE FUNCTIONAL DESCRIPTION: - Upgrade to support Cmake 3.3 and later - Update workflow unit tests to take the results of the execution into account - DIET workflow engine was improved
NEWS OF THE YEAR: Work on the next DIET release (DIET 2.11) New DIET Webboard based on Angular Two news biological application platform based on DIET (Aevol and Wasabi) Rutgers University (NJ, USA) Collaboration

- Participants: Joel Faubert, Hadrien Croubois, Abdelkader Amar, Arnaud Lefray, Aurélien Bouteiller, Benjamin Isnard, Daniel Balouek, Eddy Caron, Eric Bois, Frédéric Desprez, Frédéric Lombart, Gaël Le Mahec, Guillaume Verger, Huaxi Zhang, Jean-Marc Nicod, Jonathan Rouzaud-Cornabas, Lamiel Toch, Maurice Faye, Peter Frauenkron, Philippe Combes, Philippe Laurent, Raphaël Bolze and Yves Caniou
- Partners: CNRS - ENS Lyon - UCBL Lyon 1 - Sysfera
- Contact: Eddy Caron
- URL: http://graal.ens-lyon.fr/diet/

6.3. SimGrid

KEYWORDS: Large-scale Emulators - Grid Computing - Distributed Applications
SCIENTIFIC DESCRIPTION: SimGrid is a toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. The simulation engine uses algorithmic and implementation techniques toward the fast simulation of large systems on a single machine. The models are theoretically grounded and experimentally validated. The results are reproducible, enabling better scientific practices.

Its models of networks, cpus and disks are adapted to (Data)Grids, P2P, Clouds, Clusters and HPC, allowing multi-domain studies. It can be used either to simulate algorithms and prototypes of applications, or to emulate real MPI applications through the virtualization of their communication, or to formally assess algorithms and applications that can run in the framework.

The formal verification module explores all possible message interleavings in the application, searching for states violating the provided properties. We recently added the ability to assess liveness properties over arbitrary and legacy codes, thanks to a system-level introspection tool that provides a finely detailed view of the running application to the model checker. This can for example be leveraged to verify both safety or liveness properties, on arbitrary MPI code written in C/C++/Fortran.

NEWS OF THE YEAR: There were 3 major releases in 2018: The public API was sanitized (with compatibility wrappers in place). Th documentation was completely overhauled. Our continuous integration was greatly improved (45 Proxy Apps + BigDFT + StarPU + BatSim now tested nightly). Some kernel headers are now installed, allowing external plugins. Allow dynamic replay of MPI apps, controlled by S4U actors. Port the MPI trace replay engine to C++, fix visualization (+ the classical bug fixes and doc improvement).

- Participants: Adrien Lèbre, Arnaud Legrand, Augustin Degomme, Florence Perronnin, Frédéric Suter, Jean-Marc Vincent, Jonathan Pastor, Luka Stanisic and Martin Quinson
- Partners: CNRS - ENS Rennes
- Contact: Martin Quinson
- URL: https://simgrid.org/

6.4. Kwapi

KiloWatt API
KEYWORD: Power monitoring
FUNCTIONAL DESCRIPTION: Kwapi is a software framework dealing with energy monitoring of large scale infrastructures through heterogeneous energy sensors. Kwapi has been designed inside the FSN XLCloud project for Openstack infrastructures. Through the support of Hemera Inria project, kwapi has been extended and deployed in production mode to support easy and large scale energy profiling of the Grid5000 resources. Kwapi now supports high frequency powermeters of the Grid5000 Lyon platform.

RELEASE FUNCTIONAL DESCRIPTION: - many bugfixes - multiprocessing instead of multithreading - many optimizations

- Participants: François Rossigneux, Jean-Patrick Gelas, Laurent Lefèvre, Laurent Pouilloux, Simon Delamare and Matthieu Imbert
- Contact: Laurent Lefèvre
- URL: https://launchpad.net/kwapi

6.5. execo

KEYWORDS: Toolbox - Deployment - Orchestration - Python
FUNCTIONAL DESCRIPTION: Execo offers a Python API for asynchronous control of local or remote, standalone or parallel, unix processes. It is especially well suited for quickly and easily scripting workflows of parallel/distributed operations on local or remote hosts: automate a scientific workflow, conduct computer science experiments, perform automated tests, etc. The core python package is execo. The execo_g5k package provides a set of tools and extensions for the Grid5000 testbed. The execo_engine package provides tools to ease the development of computer sciences experiments.

RELEASE FUNCTIONAL DESCRIPTION: - misc python3 support fixes - basic documentation for wheezy compatible package build - remove some debug outputs - fix crash in processes conductor in some situations - improve/fix process stdout/stderr handlers - fix ge_cluste_networ_equipments - add a FAQ

- Participants: Florent Chuffart, Laurent Pouilloux and Matthieu Imbert
- Contact: Matthieu Imbert
- URL: http://execo.gforge.inria.fr

6.6. SeeDep

Seed based Deployment
KEYWORDS: Reproducibility - Deployment - Cloud

SCIENTIFIC DESCRIPTION: SeeDep aims at devising a new way where researchers can communicate in a comprehensive and accurate way the experimentation set-up used in their work. It lies on two components: (i) a public algorithm that generates experimentation networks, and (ii) a generation key (i.e. a seed) that can be shared which specifies the said network. Therefore, researchers only need to share (in their paper for instance) the “generation key” that corresponds to their experimentation network. With such key, any other researcher/professional will be able to re-generate a comprehensive and accurate model of the same network.
FUNCTIONAL DESCRIPTION: SeeDep is a framework aiming at generating, reproducing and deploying experiments set-up on different Cloud platforms.

- Participants: Cyril Seguin and Eddy Caron
- Partner: Nokia Bell Labs
- Contact: Eddy Caron

6.7. libkomp

Runtime system libkomp
KEYWORDS: HPC - Multicore - OpenMP
FUNCTIONAL DESCRIPTION: libKOMP is a runtime support for OpenMP compatible with différent compiler: GNU gcc/gfortran, Intel icc/ifort or clang/llvm. It is based on source code initially developed by Intel for its own OpenMP runtime, with extensions from Kaapi software (task representation, task scheduling). Moreover it contains an OMPT module for recording trace of execution.

RELEASE FUNCTIONAL DESCRIPTION: Initial version

- Contact: Thierry Gautier
- URL: http://gitlab.inria.fr/openmp/libkomp

6.8. Platform: Grid’5000

Participants: Laurent Lefèvre, Simon Delamare, David Loup, Christian Perez.

FUNCTIONAL DESCRIPTION

The Grid’5000 experimental platform is a scientific instrument to support computer science research related to distributed systems, including parallel processing, high performance computing, cloud computing, operating systems, peer-to-peer systems and networks. It is distributed on 10 sites in France and Luxembourg, including Lyon. Grid’5000 is a unique platform as it offers to researchers many and varied hardware resources and a complete software stack to conduct complex experiments, ensure reproducibility and ease understanding of results. In 2018, a new generation of high speed wattmeters has been deployed on the Lyon site. They allow energy monitoring with up to 50 measurements per second. In parallel, a new version of kwapi (software stack for energy monitoring) has been proposed and redesigned.

- Contact: Laurent Lefèvre
- URL: https://www.grid5000.fr/

6.9. Platform: Leco

Participants: Thierry Gautier, Laurent Lefèvre, Christian Perez.

FUNCTIONAL DESCRIPTION

The LECO experimental platform is a new medium size scientific instrument funded by DRRT to investigate research related to BigData and HPC. It is located in Grenoble as part of the the HPCDA computer managed by UMS GRICAD. The platform has been deployed in 2018 and was available for experiment since the summer. All the nodes of the platform are instrumented to capture the energy consumption and data are available through the Kwapi software.

- Contact: Thierry Gautier

6.10. Platform: SILECS

Participants: Laurent Lefèvre, Simon Delamare, Christian Perez.

The SILECS infrastructure (IR ministère) aims at providing an experimental platform for experimental computer Science (Internet of things, clouds, hpc, big data, *etc.*). This new infrastructure is based on two existing infrastructures, Grid'5000 and FIT.

- Contact: Christian Perez
- URL: https://www.silecs.net/

7. New Results

7.1. Energy Efficiency in HPC and Large Scale Distributed Systems

Participants: Laurent Lefèvre, Dorra Boughzala, Christian Perez, Issam Raïs, Mathilde Boutigny.

7.1.1. Building and Exploiting the Table of Leverages in Large Scale HPC Systems

Large scale distributed systems and supercomputers consume huge amounts of energy. To address this issue, an heterogeneous set of capabilities and techniques that we call leverages exist to modify power and energy consumption in large scale systems. This includes hardware related leverages (such as Dynamic Voltage and Frequency Scaling), middleware (such as scheduling policies) and application (such as the precision of computation) energy leverages. Discovering such leverages, benchmarking and orchestrating them, remains a real challenge for most of the users. We have formally defined energy leverages, and we proposed a solution to automatically build the table of leverages associated with a large set of independent computing resources. We have shown that the construction of the table can be parallelized at very large scale with a set of independent nodes in order to reduce its execution time while maintaining precision of observed knowledge [22], [25].

7.1.2. Automatic Energy Efficient HPC Programming: A Case Study

Energy consumption is one of the major challenges of modern datacenters and supercomputers. By applying Green Programming techniques, developers have to iteratively implement and test new versions of their software, thus evaluating the impact of each code version on their energy, power and performance objectives. This approach is manual and can be long, challenging and complicated, especially for High Performance Computing applications. In [24], we formally introduces the definition of the Code Version Variability (CVV) leverage and present a first approach to automate Green Programming (*i.e.,* CVV usage) by studying the specific use-case of an HPC stencil-based numerical code, used in production. This approach is based on the automatic generation of code versions thanks to a Domain Specific Language (DSL), and on the automatic choice of code version through a set of actors. Moreover, a real case study is introduced and evaluated though a set of benchmarks to show that several trade-offs are introduced by CVV 1. Finally, different kinds of production scenarios are evaluated through simulation to illustrate possible benefits of applying various actors on top of the CVV automation.

7.1.3. Performance and Energy Analysis of OpenMP Runtime Systems with Dense Linear Algebra Algorithms

In the article [9], we analyze performance and energy consumption of five OpenMP runtime systems over a non-uniform memory access (NUMA) platform. We also selected three CPU-level optimizations or techniques to evaluate their impact on the runtime systems: processors features Turbo Boost and C-States, and CPU Dynamic Voltage and Frequency Scaling through Linux CPUFreq governors. We present an experimental study to characterize OpenMP runtime systems on the three main kernels in dense linear algebra algorithms (Cholesky, LU, and QR) in terms of performance and energy consumption. Our experimental results suggest that OpenMP runtime systems can be considered as a new energy leverage, and Turbo Boost, as well as C-States, impacted significantly performance and energy. CPUFreq governors had more impact with Turbo Boost disabled, since both optimizations reduced performance due to CPU thermal limits. An LU factorization with concurrent-write extension from libKOMP achieved up to 63% of performance gain and 29% of energy decrease over original PLASMA algorithm using GNU C compiler (GCC) libGOMP runtime.

7.1.4. Energy Simulation of GPU based Infrastructures

Through the IPL Hac-Specis and the PhD of Dorra Boughzala we begin to explore the modeling and calibrating of energy consumption of GPU architectures. We use the SimGrid simulation framework for the integration and validation on large scale systems.

7.2. HPC Component Models and Runtimes

Participants: Thierry Gautier, Christian Perez, Jérôme Richard.

7.2.1. On the Impact of OpenMP Task Granularity

Tasks are a good support for composition. During the development of a high-level component model for HPC, we have experimented to manage parallelism from components using OpenMP tasks. Since version 4-0, the standard proposes a model with dependent tasks that seems very attractive because it enables the description of dependencies between tasks generated by different components without breaking maintainability constraints such as separation of concerns. In [20], we present our feedback on using OpenMP in our context. We discover that our main issues are a too coarse task granularity for our expected performance on classical OpenMP runtimes, and a harmful task throttling heuristic counter-productive for our applications. We present a completion time breakdown of task management in the Intel OpenMP runtime and propose extensions evaluated on a testbed application coming from the Gysela application in plasma physics.

7.2.2. Building and Auto-Tuning Computing Kernels: Experimenting with BOAST and StarPU in the GYSELA Code

Modeling turbulent transport is a major goal in order to predict confinement performance in a tokamak plasma. The gyrokinetic framework considers a computational domain in five dimensions to look at kinetic issues in a plasma; this leads to huge computational needs. Therefore, optimization of the code is an especially important aspect, especially since coprocessors and complex manycore architectures are foreseen as building blocks for Exascale systems. This project [6] aims to evaluate the applicability of two auto-tuning approaches with the BOAST and StarPU tools on the gysela code in order to circumvent performance portability issues. A specific computation intensive kernel is considered in order to evaluate the benefit of these methods. StarPU enables to match the performance and even sometimes outperform the hand-optimized version of the code while leaving scheduling choices to an automated process. BOAST on the other hand reveals to be well suited to get a gain in terms of execution time on four architectures. Speedups in-between 1.9 and 5.7 are obtained on a cornerstone computation intensive kernel.

7.3. Modeling and Simulation of Parallel Applications and Distributed Infrastructures

Participants: Eddy Caron, Zeina Houmani, Frédéric Suter.

7.3.1. SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation

It is typical in High Performance Computing (HPC) courses to give students access to HPC platforms so that they can benefit from hands-on learning opportunities. Using such platforms, however, comes with logistical and pedagogical challenges. For instance, a logistical challenge is that access to representative platforms must be granted to students, which can be difficult for some institutions or course modalities; and a pedagogical challenge is that hands-on learning opportunities are constrained by the configurations of these platforms. A way to address these challenges is to instead simulate program executions on arbitrary HPC platform configurations. In [15] we focus on simulation in the specific context of distributed-memory computing and MPI programming education. While using simulation in this context has been explored in previous works, our approach offers two crucial advantages. First, students write standard MPI programs and can both debug and analyze the performance of their programs in simulation mode. Second, large-scale executions can be simulated in short amounts of time on a single standard laptop computer. This is possible thanks to SMPI, an

MPI simulator provided as part of SimGrid. After detailing the challenges involved when using HPC platforms for HPC education and providing background information about SMPI, we present SMPI Courseware. SMPI Courseware is a set of in-simulation assignments that can be incorporated into HPC courses to provide students with hands-on experience for distributed-memory computing and MPI programming learning objectives. We describe some these assignments, highlighting how simulation with SMPI enhances the student learning experience.

7.3.2. Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems

File replication is widely used to reduce file transfer times and improve data availability in large distributed systems. Replication techniques are often evaluated through simulations, however, most simulation platform models are oversimplified, which questions the applicability of the findings to real systems. In [17], we investigate how platform models influence the performance of file replication strategies on large heterogeneous distributed systems, based on common existing techniques such as prestaging and dynamic replication. The novelty of our study resides in our evaluation using a realistic simulator. We consider two platform models: a simple hierarchical model and a detailed model built from execution traces. Our results show that conclusions depend on the modeling of the platform and its capacity to capture the characteristics of the targeted production infrastructure. We also derive recommendations for the implementation of an optimized data management strategy in a scientific gateway for medical image analysis.

7.3.3. WRENCH: Workflow Management System Simulation Workbench

Scientific workflows are used routinely in numerous scientific domains, and Workflow Management Systems (WMSs) have been developed to orchestrate and optimize workflow executions on distributed platforms. WMSs are complex software systems that interact with complex software infrastructures. Most WMS research and development activities rely on empirical experiments conducted with full-fledged software stacks on actual hardware platforms. Such experiments, however, are limited to hardware and software infrastructures at hand and can be labor- and/or time-intensive. As a result, relying solely on real-world experiments impedes WMS research and development. An alternative is to conduct experiments in simulation.

In [16] we presented WRENCH, a WMS simulation framework, whose objectives are (i) accurate and scalable simulations; and (ii) easy simulation software development. WRENCH achieves its first objective by building on the SimGrid framework. While SimGrid is recognized for the accuracy and scalability of its simulation models, it only provides low-level simulation abstractions and thus large software development efforts are required when implementing simulators of complex systems. WRENCH thus achieves its second objective by providing high- level and directly re-usable simulation abstractions on top of SimGrid. After describing and giving rationales for WRENCH's software architecture and APIs, we present a case study in which we apply WRENCH to simulate the Pegasus production WMS. We report on ease of implementation, simulation accuracy, and simulation scalability so as to determine to which extent WRENCH achieves its two above objectives. We also draw both qualitative and quantitative comparisons with a previously proposed workflow simulator.

7.3.4. A Microservices Architectures for Data-Driven Service Discovery

Usual microservices discovery mechanisms are normally based on a specific user need *(Goal-based approaches)*. However, in today's evolving architectures, users need to discover what features they can take advantage of before looking for the available microservices. In collaboration with RDI2 (Rutgers University) we developped a data-driven microservices architecture that allows users to discover, from specific objects, the features that can be exerted on these objects as well as all the microservices dedicated to them [28]. This architecture, based on the main components of the usual microservices architectures, adopts a particular communication strategy between clients and registry to achieve the goal. This article contains a representation of a microservice data model and a P2P model that transforms our architecture into a robust and scalable system. Also, we designed a prototype to validate our approach using Istio library.

7.4. Cloud Resource Management

Participants: Eddy Caron, Hadrien Croubois, Jad Darrous, Christian Perez.

7.4.1. Nitro: Network-Aware Virtual Machine Image Management in Geo-Distributed Clouds

Recently, most large cloud providers, like Amazon and Microsoft, replicate their Virtual Machine Images (VMIs) on multiple geographically distributed data centers to offer fast service provisioning. Provisioning a service may require to transfer a VMI over the wide-area network (WAN) and therefore is dictated by the distribution of VMIs and the network bandwidth in-between sites. Nevertheless, existing methods to facilitate VMI management (*i.e.,* retrieving VMIs) overlook network heterogeneity in geo-distributed clouds. In [19], we design, implement and evaluate Nitro, a novel VMI management system that helps to minimize the transfer time of VMIs over a heterogeneous WAN. To achieve this goal, Nitro incorporates two complementary features. First, it makes use of deduplication to reduce the amount of data which will be transferred due to the high similarities within an image and in-between images. Second, Nitro is equipped with a network-aware data transfer strategy to effectively exploit links with high bandwidth when acquiring data and thus expedites the provisioning time. Experimental results show that our network-aware data transfer strategy offers the optimal solution when acquiring VMIs while introducing minimal overhead. Moreover, Nitro outperforms state-of-the-art VMI storage systems (*e.g.,* OpenStack Swift) by up to 77%.

7.4.2. Toward an Autonomic Engine for Scientific Workflows and Elastic Cloud Infrastructure

The constant development of scientific and industrial computation infrastructures requires the concurrent development of scheduling and deployment mechanisms to manage such infrastructures. Throughout the last decade, the emergence of the Cloud paradigm raised many hopes, but achieving full platform autonomicity is still an ongoing challenge. We built a workflow engine that integrated the logic needed to manage workflow execution and Cloud deployment on its own. More precisely, we focus on Cloud solutions with a dedicated Data as a Service (DaaS) data management component. Our objective was to automate the execution of workflows submitted by many users on elastic Cloud resources. This contribution proposes a modular middleware infrastructure and details the implementation of the underlying modules:

- A workflow clustering algorithm that optimises data locality in the context of DaaS-centered communications;
- A dynamic scheduler that executes clustered workflows on Cloud resources;
- A deployment manager that handles the allocation and deallocation of Cloud resources according to the workload characteristics and users' requirements.

All these modules have been implemented in a simulator to analyse their behaviour and measure their effectiveness when running both synthetic and real scientific workflows. We also implemented these modules in the Diet middleware to give it new features and prove the versatility of this approach. Simulation running the WASABI workflow (waves analysis based inference, a framework for the reconstruction of gene regulatory networks) showed that our approach can decrease the deployment cost by up to 44% while meeting the required deadlines [13].

7.4.3. Madeus: A Formal Deployment Model

Distributed software architecture is composed of multiple interacting modules, or components. Deploying such software consists in installing them on a given infrastructure and leading them to a functional state. However, since each module has its own life cycle and might have various dependencies with other modules, deploying such software is a very tedious task, particularly on massively distributed and heterogeneous infrastructures. To address this problem, many solutions have been designed to automate the deployment process. In [18], we introduce Madeus, a component-based deployment model for complex distributed software. Madeus accurately describes the life cycle of each component by a Petri net structure, and is able to finely express the dependencies between components. The overall dependency graph it produces is then used to reduce deployment time by parallelizing deployment actions. While this increases the precision and performance of the model, it also increases its complexity. For this reason, the operational semantics need to be clearly defined to prove results such as the termination of a deployment. In this paper, we formally describe the operational semantics of Madeus, and show how it can be used in a use-case: the deployment of a real and large distributed software (*i.e.,* , OpenStack).

In [18], we have proposed an extension based on component behavioral interfaces to the Aeolus component model to better separate the concerns of component users (*e.g.,* application architect) from component developers.

7.5. Data Stream Processing on Edge Computing

Participants: Eddy Caron, Felipe Rodrigo de Souza, Marcos Dias de Assunção, Laurent Lefèvre, Alexandre Da Silva Veith.

7.5.1. Latency-Aware Placement of Data Stream Analytics on Edge Computing

The interest in processing data events under stringent time constraints as they arrive has led to the emergence of architecture and engines for data stream processing. Edge computing, initially designed to minimize the latency of content delivered to mobile devices, can be used for executing certain stream processing operations. Moving opera- tors from cloud to edge, however, is challenging as operator-placement decisions must consider the application requirements and the network capabilities. We introduce strategies to create placement configurations for data stream processing applications whose operator topologies follow series parallel graphs[35]. We consider the operator characteristics and requirements to improve the response time of such appli- cations. Results show that our strategies can improve the response time in up to 50% for application graphs comprising multiple forks and joins while transferring less data and better using the resources.

7.5.2. Estimating Throughput of Stream Processing Applications in FoG Computing

Recent trends exploit decentralized infrastructures (e.g.. Fog computing) to deploy DSP (Data Stream Processing) applications and leverage the computational power. Fog computing overlaps some features of Cloud computing and includes others, for instance, location awareness. The operator placement problem consists of determining, within a set of distributed computing resources, the computing resources that should host and execute each operator of the DSP application, with the goal of optimizing QoS requirements of the application. The QoS requirements of the application refer to processing time, costs, throughput, etc. We propose a model to estimate the application throughput at each layer of Fog computing (Devices, Edge and Cloud) by considering a given placement solution. The estimated throughput provides a useful insight to determine the amount of physical resources to meet the QoS requirements. The model allows to identify the application bottleneck, when facing data rate variations, and provides information to self-scale in or out the DSP application.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. IFPEN

We have a collaboration with IFPEN (http://ifpenergiesnouvelles.com/). IFPEN develops numerical codes to solve PDE with specific adaption of the preconditioning step to fit the requirement of their problems. With a PhD student (Adrien Roussel) we have studied the parallel implementation of multi-level decomposition domains on many-core architecture and KNL processor.

8.1.2. Nokia Bell Labs

AVALON has been actively collaborating with Nokia, formerly Alcatel-Lucent Bell Labs, in the framework of the Nokia/Alcatel-Lucent Inria Joint Laboratory. We was involved in the following Research Actions (Actions de Recherche (ADR) in French) of this laboratory. ADR Nokia Bell Labs /Inria: Procedural Generation of Networks for Security Research & Experimentations. The objective of this project is to address such challenge. We aim at devising a new way where researchers can communicate in a comprehensive and accurate way the experimentation set-up used in their work. The main objective would be to research and develop the procedural generation of credible network topologies and test beds resembling real operational infrastructures of various kinds (e.g. classical ICT, virtualized Cloud or SDN based, SCADA infrastructures etc.), as a method of creating data algorithmically as opposed to manually. This work is done with a postdoc position: Cyril Seguin.

8.2. Bilateral Grants with Industry

8.2.1. Orange

We have a collaboration with Orange. This collaboration is sealed through a CIFRE Phd grant. The research of the Phd student (Arthur Chevalier) focuses on placement and compliance aspects of software licenses in a Cloud architecture. Today, the use of software is regulated by licenses, whether they are free, paid for and with or without access to its sources. The number of licenses required for specific software can be calculated with several metrics, each defined by the software vendor. Our goal is to propose a deployment algorithm that takes into account different metrics.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. CPER

Participants: Thierry Gautier, Laurent Lefèvre, Christian Perez.

The LECO experimental platform is a new medium size scientific instrument deployed in Grenoble in 2018. It was funded by the CPER 2015-2020 LECO++ to investigate research related to BigData and HPC.

9.2. National Initiatives

9.2.1. PIA

9.2.1.1. PIA ELCI, Environnement Logiciel pour le Calcul Intensif, 2014-2018

Participants: Mathilde Boutigny, Thierry Gautier, Laurent Lefèvre, Christian Perez, Issam Raïs, Jérôme Richard, Philippe Virouleau.

The ELCI PIA project is coordinated by BULL with several partners: CEA, Inria, SAFRAB, UVSQ.

This project aims to improve the support for numerical simulations and High Performance Computing (HPC) by providing a new generation software stack to control supercomputers, to improve numerical solvers, and pre- and post computing software, as well as programming and execution environment. It also aims to validate the relevance of these developments by demonstrating their capacity to deliver better scalability, resilience, modularity, abstraction, and interaction on some application use-cases. AVALON is involved in WP1 and WP3 ELCI Work Packages through the PhD of Issam Raïs and the postdoc of Hélène Coullon. Laurent Lefèvre is the Inria representative in the ELCI technical committee.

9.2.2. MRSEI

9.2.2.1. Fennec, FastEr NaNo-Characterisation, 24 months, 2018-2021

Participants: Eddy Caron, Christian Perez.

The goal of the ANR-MRSEI FENNEC project is to support the submission of a project to the European call DT-NMBP-08-2019 entitled “Real-time nano-characterisation technologies (RIA)”.

9.2.3. Inria Large Scale Initiative

9.2.3.1. DISCOVERY, DIStributed and COoperative management of Virtual EnviRonments autonomouslY, 4 years, 2015-2019

Participants: Maverick Chardet, Jad Darrous, Christian Perez.

To accommodate the ever-increasing demand for Utility Computing (UC) resources, while taking into account both energy and economical issues, the current trend consists in building larger and larger Data Centers in a few strategic locations. Although such an approach enables UC providers to cope with the actual demand while continuing to operate UC resources through centralized software system, it is far from delivering sustainable and efficient UC infrastructures for future needs.

The DISCOVERY initiative aims at exploring a new way of operating Utility Computing (UC) resources by leveraging any facilities available through the Internet in order to deliver widely distributed platforms that can better match the geographical dispersal of users as well as the ever increasing demand. Critical to the emergence of such locality-based UC (LUC) platforms is the availability of appropriate operating mechanisms. The main objective of DISCOVERY is to design, implement, demonstrate and promote the LUC Operating System (OS), a unified system in charge of turning a complex, extremely large-scale and widely distributed infrastructure into a collection of abstracted computing resources which is efficient, reliable, secure and at the same time friendly to operate and use.

To achieve this, the consortium is composed of experts in research areas such as large-scale infrastructure management systems, network and P2P algorithms. Moreover two key network operators, namely Orange and RENATER, are involved in the project.

By deploying and using such a LUC Operating System on backbones, our ultimate vision is to make possible to host/operate a large part of the Internet by its internal structure itself: A scalable set of resources delivered by any computing facilities forming the Internet, starting from the larger hubs operated by ISPs, government and academic institutions, to any idle resources that may be provided by end-users.

9.2.3.2. HAC SPECIS, High-performance Application and Computers, Studying PErformance and Correctness In Simulation, 4 years, 2016-2020

Participants: Dorra Boughzala, Idriss Daoudi, Thierry Gautier, Laurent Lefèvre, Frédéric Suter.

Over the last decades, both hardware and software of modern computers have become increasingly complex. Multi-core architectures comprising several accelerators (GPUs or the Intel Xeon Phi) and interconnected by high-speed networks have become mainstream in HPC. Obtaining the maximum performance of such heterogeneous machines requires to break the traditional uniform programming paradigm. To scale, application developers have to make their code as adaptive as possible and to release synchronizations as much as possible. They also have to resort to sophisticated and dynamic data management, load balancing, and scheduling strategies. This evolution has several consequences:

First, this increasing complexity and the release of synchronizations are even more error-prone than before. The resulting bugs may almost never occur at small scale but systematically occur at large scale and in a non deterministic way, which makes them particularly difficult to identify and eliminate.

Second, the dozen of software stacks and their interactions have become so complex that predicting the performance (in terms of time, resource usage, and energy) of the system as a whole is extremely difficult. Understanding and configuring such systems therefore becomes a key challenge.

These two challenges related to correctness and performance can be answered by gathering the skills from experts of formal verification, performance evaluation and high performance computing. The goal of the HAC SPECIS Inria Project Laboratory is to answer the methodological needs raised by the recent evolution of HPC architectures by allowing application and runtime developers to study such systems both from the correctness and performance point of view.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. POP

Participant: Frédéric Suter.

Program: H2020 Center of Excellence

Project acronym: POP

Project title: Performance Optimisation and Productivity

Duration: 2015-2018

Coordinator: Barcelona Supercomputing Centre (BSC)

Other partners: High Performance Computing Center Stuttgart of the University of Stuttgart (HLRS), Jülich Supercomputing Centre (JSC), Numerical Algorithm Group (NAG), Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), TERATEC (TERATEC).

Abstract: The Center of Excellence for Performance Optimisation and Productivity provides performance optimisation and productivity services for academic and industrial codes. European's leading experts from the High Performance Computing field will help application developers getting a precise understanding of application and system behaviour. This project is supported by the European Commission under H2020 Grant Agreement No. 676553.

Established codes, but especially codes never undergone any analysis or performance tuning, may profit from the expertise of the POP services which use latest state-of-the-art tools to detect and locate bottlenecks in applications, suggest possible code improvements, and may even help by Proof-of-Concept experiments and mock-up test for customer codes on their own platforms.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

9.3.2.1. COST IC1305 : Nesus

Participants: Marcos Dias de Assunção, Laurent Lefèvre.

Program: COST

Project acronym: IC1305

Project title: Network for Sustainable Ultrascale Computing (NESUS)

Duration: 2014-2018

Coordinator: Jesus Carretero (Univ. Madrid)

Abstract: Ultrascale systems are envisioned as large-scale complex systems joining parallel and distributed computing systems that will be two to three orders of magnitude larger that today's systems. The EU is already funding large scale computing systems research, but it is not coordinated across researchers, leading to duplications and inefficiencies. The goal of the NESUS Action is to establish an open European research network targeting sustainable solutions for ultrascale computing aiming at cross fertilization among HPC, large scale distributed systems, and big data management. The network will contribute to glue disparate researchers working across different areas and provide a meeting ground for researchers in these separate areas to exchange ideas, to identify synergies, and to pursue common activities in research topics such as sustainable software solutions (applications and system software stack), data management, energy efficiency, and resilience. In Nesus, Laurent Lefèvre is co-chairing the Working on Energy Efficiency (WG5).

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. Joint Laboratory for Extreme Scale Computing (JLESC) (2014-2018)

Participants: Thierry Gautier, Christian Perez, Jérôme Richard.

Partners: NCSA (US), ANL (US), Inria (FR), Jülich Supercomputing Centre (DE), BSC (SP), Riken (JP).
The purpose of the Joint Laboratory for Extreme Scale Computing (JLESC) is to be an international, virtual organization whose goal is to enhance the ability of member organizations and investigators to make the bridge between Petascale and Extreme computing. The founding partners of the JLESC are Inria and UIUC. Further members are ANL, BSC, JSC and RIKEN-AICS.

JLESC involves computer scientists, engineers and scientists from other disciplines as well as from industry, to ensure that the research facilitated by the Laboratory addresses science and engineering's most critical needs and takes advantage of the continuing evolution of computing technologies.

9.4.1.2. Associate Team DALHIS – Data Analysis on Large-scale Heterogeneous Infrastructures for Science (2013-2018)

Participant: Frédéric Suter.

Partners: EPC Myriads (Rennes, Bretagne Atlantique), AVALON (Grenoble, Rhône-Alpes), Data Science and Technology Department (LBNL,USA).

The goal of the Inria-LBL collaboration is to create a collaborative distributed software ecosystem to manage data lifecycle and enable data analytics on distributed data sets and resources. Specifically, our goal is to build a dynamic software stack that is user-friendly, scalable, energy-efficient and fault tolerant. We plan to approach the problem from two dimensions: (i) Research to determine appropriate execution environments that allow users to seamlessly execute their end-to-end dynamic data analysis workflows in various resource environments and scales while meeting energy-efficiency, performance and fault tolerance goals; (ii) Engagement in deep partnerships with scientific teams and use a mix of user research with system software R&D to address specific challenges that these communities face, and inform future research directions from acquired experience.

9.4.2. Inria Associate Teams Not Involved in an Inria International Labs

9.4.2.1. Associate Team SUSTAM – Sustainable Ultra Scale compuTing, dAta and energy Management (2017-2020)

Participants: Eddy Caron, Hadrien Croubois, Marcos Dias de Assunção, Alexandre Da Silva Veith, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, Valentin Lorentz, Christian Perez, Issam Raïs.

International Partners: Rutgers University (United States) - RDI2 - Manish Parashar

The SUSTAM associate team will focus on the joint design of a multi-criteria orchestration framework dealing with resources, data and energy management in an sustainable way. The SUSTAM associated team will enable a long-term collaboration between the Inria AVALON team and the Rutgers Discovery Informatics Institute (RDI2) from Rutgers University (USA). The SUSTAM associated team is leaded by Laurent Lefèvre.

9.4.3. Participation in Other International Programs

9.4.3.1. Joint Project CNRS/University of Melbourne – Algorithms for Placement and Reconfiguration of Data Stream Processing Applications (2017-2018)

Participants: Marcos Dias de Assunção, Alexandre Da Silva Veith, Laurent Lefèvre.

Partner: Clouds Lab (The University of Melbourne, Australia).

Much of the "big data" produced today is created as continuous data streams that are most valuable when processed quickly. Several data stream processing frameworks have been designed for running on clusters of homogeneous computers. Under most frameworks, an application is a Direct Acyclic Graph (DAG) whose vertices are operators that execute transformations over the incoming data and edges that define how the data flows between operators. While cloud computing is a key infrastructure for deploying such frameworks, more modern solutions leverage the edges of the Internet (e.g. edge computing) to offload some of the processing from the cloud and hence reduce the end-to-end latency. The placement and reconfiguration of stream processing DAGs onto highly distributed and heterogeneous infrastructure are, however, challenging endeavours. This project aims to investigate algorithms for the placement and dynamic reconfiguration of stream processing components considering multiple criteria.

9.5. International Research Visitors

9.5.1. Research Stays Abroad

Participant: Olivier Glück.

Olivier Glück has been invited professor by Concordia University (Faculty of Engineering & Computer Science, Department of Computer Science & Software Engineering). He has conducted researches with Dr. Brigitte Jaumard, Professor & Research Chair, Tier 1, on the Optimization of Communication Networks. He has worked on the following Virtual Machine (VM) migration optimization problem: find a scheduling of VM migration that minimizes the makespan *i.e.,* total duration of the migration assuming that the current VM placement and the target one are given. He has proposed a new sequence-based optimization model with a Mixed Integer Linear Program (MILP), which not only guarantees the finding of the best VM migration scheduling but also the migration of the largest possible number of VMs in the case of deadlocks. He also worked on the design of heuristic algorithms for VM migration and a generator of real VM migration instances to evaluate the models and heuristics proposed. He has also worked on the task offloading problem in edge computing.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Laurent Lefèvre was co-General chair of the conference SBAC-PAD : The International Symposium on Computer Architecture and High Performance Computing, Lyon, France (September 24-27, 2018). He was co-organizer of colloquium : "Digital society vs eco-responsability: impact on billions of connected objects, networks and clouds", with Centre Jacques Cartier, Concordia University, Inria, GDS CNRS EcoInfo, Ecole Normale Superieure of Lyon (November 12-13, 2018). He co-organized the E3-RSD school on Energy Efficiency in Networks and Distributed Systems, Dinard (October 1-4, 2018) and the GreenDays@Toulouse v2.0 (2018 Edition) "From IoT to Exascale, what about energy efficiency and carbon emission reduction ?", Toulouse (July 2-3, 2018)

10.1.1.2. Member of the Organizing Committees

Laurent Lefèvre was co Special Session Organizer of Special Session on High Performance Computing Benchmarking and Optimization (HPBench 2018), during HPCS conference, Orleans (July 16-20, 2018).

Christian Perez served on the Organizing Committee of the 2018 International Conference on High Performance Computing & Simulation (HPCS 2017) as Conference Awards and Recognitions Co-Chair, on the Organizing Comittee of the French Journées Calculs et Données (Lyon, October 24-26) and of the 1st Grid'5000-FIT school (Sophia Antipolis, April 3-6).

Frédéric Suter served in the Organization commitee of the session on Exascale Computing for High Energy Physics of the 14th IEEE International Conference on eScience and chaired the "Parallelism" track of the Conférence d'informatique en Parallélisme, Architecture et Système (Compas 2018).

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

Christian Perez was member of the program committees of the IEEE International Conference on Cluster Computing 2018 (Cluster 2018), the 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid (CCGRID 2018), the 26th annual High Performance Computing Symposium 2018 (HPC '18), the 2018 Supercomputing Asia conference, and of the 4th International Workshop on Autonomic High Performance Computing. He was a member of the Birds of Feather Committee of SuperComputing 2018 and of the Tutorial Committee of ISC High Performance 2018.

Olivier Glück was a member of the program committees of PDP 2018 (26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing) and ICA3PP-2018 (18th International Conference on Algorithms and Architectures for Parallel Processing).

Frédéric Suter was member of program committee of SBAC-PAD 2018 (30th International Symposium on Computer Architecture and High Performance Computing), VECPAR 2018 (13th International Meeting High Performance Computing for Computational Science) and ICA3PP-2018 (18th International Conference on Algorithms and Architectures for Parallel Processing)

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

Christian Perez reviewed articles for the IEEE's Transactions on Big Data journal and for Oil & Gas Science and Technology journal of IFP Energies nouvelles.

Thierry Gautier reviewed articles for the ACM Transaction on Architecture and Code Optimization.

Frédéric Suter reviewed articles for the Journal of Parallel and Distributed Computing and Concurrency and Computation: Practice and Experience.

10.1.4. Invited Talks

Christian Perez has been invited to give the following talks:

- "Towards Reconfigurable HPC Component Models", Invited talk, 4th International Workshop on Autonomic High Performance Computing, Orléans, France, July 17.
- "HPC Component Models", Keynote talk, HPCS, Orléans, France, July 19, 2018.

Thierry Gautier has been invited to give the following talk:

- "Tasks' management in OpenMP: what about performance guarantee?", Maison de la Simulation, November 20, 2018.

Frédéric Suter has been invited to give the following talk:

- "Simulating MPI applications : the SMPI Approach" at the "Modeling and Simulation of HPC Architectures and Applications" mini-symposium at the SIAM Conference on Parallel Processing for Scientific Applications.

10.1.5. Leadership within the Scientific Community

Eddy Caron is animator and co-chair for the FIL (Fédération Informatique de Lyon) on the theme IDCHP (Informatique Distribuée et Calcul Haute Performance).

Laurent Lefèvre is animator and chair of the transversal action on "Energy" of the French GDR RSD ("Réseaux et Systèmes Distribués").

Christian Perez is co-leader of the pole Distributed Systems of the French GDR RSD ("Réseaux et Systèmes Distribués").

10.1.6. Scientific Expertise

Christian Perez reviewed two projects for the Regional Program STIC-AmSud.

Olivier Glück is member of the CNU (Conseil National des Universités) section 27 (Computer Science). He participated to the 2018 "Qualifications" session and "Suivi de carriere" session.

10.1.7. Research Administration

Eddy Caron is Deputy Director in charge of call for projects, research transfert and international affairs for the LIP. He is a member of ”Conseil Technologique Logiciel MINALOGIC". He is member and reviewer for InriaHub.

Olivier Glück is member of the "Conseil Académique" of Lyon 1 University and Lyon University.

Christian Perez represents INRIA in the overview board of the France Grilles Scientific Interest Group. He is a member of the executive board and the sites committee of the Grid'5000 Scientific Interest Group and member of the executive board of the Silecs testbed. He is a member of the Inria Grenoble Rhône-Alpes Strategic Orientation Committee. He is in charge of exploring potential scientific collaborations between INRIA and SKA France.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Yves Caniou, Algorithmique programmation impérative initiation, 25h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Pratique d'Unix, 3h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Programmation Concurrente, 35h et responsabilité d'UE, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Projet Informatique, 3h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Réseaux, 45h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Système d'Exploitation, 32h and responsible of UE, niveau L2, Université d'Ho Chi Minh Ville, Vietnam.

Licence: Yves Caniou, Système d'Exploitation, 39h et co-responsible of UE, niveau L3, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Projet pour l'Orientation en Master, 3h, niveau M1, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Responsible of alternance students, 30h, niveau M1, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Responsible of Master SRIV (Systèmes, Réseaux et Infrastructures Virtuelles), 30h, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Projet Bibliographie et Certifications, 4h and responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Sécurité, 36h et responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Sécurité, 20h, niveau M2, IGA Casablanca, Maroc.

Master: Yves Caniou, Systèmes Avancés, 4.5h, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Approfondissement Scientifique, 6h, niveau M2, Université Claude Bernard Lyon 1, France.

Licence: Eddy Caron, Projet 1, 48h, L3, ENS de Lyon. France.

Master: Eddy Caron, Projet Intégré, 42h, M1, ENS de Lyon. France.

Master: Eddy Caron, Système distribués, 30h, M1, ENS de Lyon. France.

Master: Eddy Caron, Advanced Topics in Scalable Data Management, 20h, M2, ENS de Lyon. France.

Licence: Olivier Glück, Licence pedagogical advisor, 30h, niveaux L1, L2, L3, Université Lyon 1, France.

Licence: Olivier Glück, Introduction Réseaux et Web, 32h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Bases de l'architecture pour la programmation, 62h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Algorithmique programmation impérative initiation, 50h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Réseaux, 2x70h, niveau L3, Université Lyon 1, France.

Master: Olivier Glück, Réseaux par la pratique, 8h, niveau M1, Université Lyon 1, France.

Master: Olivier Glück, Responsible of Master SRIV (Systèmes, Réseaux et Infrastructures Virtuelles) located at IGA Casablanca, 20h, niveau M2, IGA Casablanca, Maroc.

Master: Olivier Glück, Applications systèmes et réseaux, 30h, niveau M2, Université Lyon 1, France.

Master: Olivier Glück, Applications systèmes et réseaux, 24h, niveau M2, IGA Casablanca, Maroc.

Master: Olivier Glück, Administration des Systèmes et des Réseaux, 16h, niveau M2, Université Lyon 1, France.

Master: Laurent Lefèvre, Parallélisme, 12h, niveau M1, Université Lyon 1, France.

Master: Laurent Lefèvre, Réseaux avancés, 24h, niveau M2, IGA Casablanca, Maroc.

Licence: Frédéric Suter, Programmation Concurrente, 32h, niveau L3, Université Lyon 1, France.

Master: Jean-Patrick Gelas, Programmation embarquée et mobile des objets, 30h, niveau M1, Université Lyon 1, France

Master: Jean-Patrick Gelas, Introduction au Cloud Computing, 21h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Système d'exploitation, 45h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Projet en Informatique en Anglais, 15h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Réseaux Avancés, 24h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Sécurité et Admin des infra résx, 30h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Technologies embarquées, 18h, niveau M2 (Image), Université Lyon 1, France

Master: Jean-Patrick Gelas, Routage (BGP), Routeurs et IPv6, 12.5h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Systèmes embarqués (GNU/Linux, Android, ARM, Arduino), 24h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Analyse de performance, 3h, niveau M2 (TIW), Université Lyon 1, France

Master: Jean-Patrick Gelas, Cloud Computing, 15h, niveau M2 (TIW), Université Lyon 1, France

Master: Jean-Patrick Gelas, Stockage, Cloud et Virtualisation, 9.5h, niveau M2 (SRIV), Université Lyon 1, France

Master: Jean-Patrick Gelas, Développement informatique, 23h, niveau M2 (Data science), Université Lyon 1, France

10.2.2. Supervision

PhD: Issam Raïs, *Discover, model and combine energy leverages for large scale energy efficient infrastructures*, 29 sept. 2018, Laurent Lefèvre (dir), Anne Benoit (Roma Team, LIP, ENS Lyon, co-dir) and Anne-Cécile Orgerie (CNRS, Myriads team, Irisa Rennes, co-dir).

PhD: Hadrien Croubois, *Toward an autonomic engine for scientific workflows and elastic Cloud infrastructure*, 16 oct. 2018, Eddy Caron (dir).

PhD: Philippe Virouleau,*Studying and improving the use of NUMA architectures through runtime systems*, 5 jun. 2018, Fabrice Rastello (Inria, Corse team, dir), Thierry Gautier (AVALON team, LIP, ENS Lyon, co-dir) and François Broquedis (UGA, Corse team, Grenoble, co-dir).

PhD: Adrien Roussel, *Parallelization of iterative methods to solve sparse linear systems using task based runtime systems on multi and many-core architectures: application to Multi-Level Domain Decomposition methods*, 6 fev. 2018, Thierry Gautier (dir), Jean-Marc Gratien (IFPen).

PhD in progress: Dorra Boughzala, *Simulating Energy Consumption of Continuum Computing between Heterogeneous Numerical Infrastructures in HPC*, IPL Hac-Specis Inria, Laurent Lefèvre (dir) , Martin Quinson and Anne-Cécile Orgerie (Myriads, Rennes, co-dir) (since december 2017).

PhD in progress: Anchen Chai: *Simulation of the Distributed Execution of a Medical Imaging Simulator*, Hugues Benoit-Cattin (co-dir, CREATIS, INSA Lyon), Frédéric Suter (co-dir).

PhD in progress: Arthur Chevalier, *Optimisation du placement des licences logiciel des fonctions réseau dans le Cloud pour un déploiement économique et efficace*, Eddy Caron (dir), Noëlle Baillon (co-dir, Orange) (since october 2017).

PhD in progress: Jad Darrous, *Geo-distributed storage for distributed Cloud*, Gilles Fedak (dir) until Aug. 2017 then Christian Perez (dir), Shadi Ibrahim (co-dir).

PhD in progress: Idris Daoudi, *Simulation of OpenMP programs*, Olivier Aumage (dir, Storm team, Bordeaux), Thierry Gautier (co-dir) (since oct. 2018).

PhD in progress: Zeina Houmani, *A Data-driven microservices architecture for Deep Learning applications*, Eddy Caron (dir), Daniel Balouek-Thomert (Rutgers University) (since oct. 2018).

PhD pended: Aurélie Kong-Win-Chang: *Techniques de résilience pour l'ordonnancement de workflows sur plates-formes décentralisées (cloud computing) avec contraintes de sécurité*, Yves Robert (dir, ROMA, ÉNS-Lyon), Eddy Caron (co-dir) et Yves Caniou (co-dir).

PhD in progress: Felipe Rodrigo De Souza, *Networking Provisioning Algorithms for Highly Distributed Data Stream Processing*, École Doctorale, Eddy Caron (dir), Marcos Dias de Assunção (co-dir) (2017-2020).

PhD in progress: Alexandre Da Silva Veith: *Elastic Mechanisms for Big-Data Stream Analytics*, Labex MiLyon, Laurent Lefèvre (dir), Marcos Dias de Assunção (co-dir) (2016-2019).

PhD cancelled: Valentin Lorentz : *Energy traceability of data*, Gilles Fedak (dir), Laurent Lefèvre (co-dir) (2016-2018).

10.2.3. Juries

Eddy Caron was reviewer and examiner of the PhD defense committe of Bassirou Ngom. UPMC (Co-tutelle avec l'UCAD). (July 13, 2018). Hajer Salhi., University of Tunis El Manar (Phd defense will be in 2019). and examiner of the PhD defense committe of Anne-Lucie Vion, Université de Grenoble and Orange (March 29, 2018).

Laurent Lefèvre was PhD Opponent and Jury Member of the PhD defense committee of Selome Kostentinos Tesfatsion, Umea University, Sweden (April 16, 2018), reviewer and examiner of the PhD defense committee of Wilfried Yoro, Telecom SudParis (March 8, 2018), reviewer of the PhD of Jungmin Son, University of Melbourne, Australia (May 2018) and examiner of the PhD defense committee of David Guyon, IRISA, University of Rennes 1 (December 7, 2018).

Christian Perez was examiner of the HdR defense committe of Samuel Thibault, University of Bordeaux (December 12, 2018), reviewer and examiner of the PhD defense committe of Pierre Matri, University Polytenics of Madrid, Spain (June 10, 2018), Hugo Taboada, University of Bordeaux (December 11, 2018), Mohamed Abderrahim, University of Bretagne Loire (December 19, 2018), Stéphanie Challita, University of Lille (December 20, 2018), and examiner of the PhD defense committe of Yacine Taleb, European University of Brittany (October 2, 2018).

10.3. Popularization

10.3.1. Articles and contents

Laurent Lefèvre has been interviewed for the articles:

- "Le Groupe Casino remplace ses radiateurs par des data centers", Le Point Journal (December 23, 2018)
- "Comment cliquer sans trop polluer", Laurent Lefèvre, Le Progrès Journal, November 21, 2018
- "Google utilise-t-il réellement 100% d'énergie verte pour alimenter ses services ?", Libération, July 13, 2018
- "J'ai le (dé)clic écolo pour réduire mon empreinte", Journal La Vie, February 15, 2018

He was interviewed and filmed for "Pollution numérique/Digital pollution", Journal 19/20, France3 National Channel, December 9, 2018.

10.3.2. Education

Yves Caniou is responsible of the LPI Certification at Université Claude Bernard Lyon 1. Further discussions are in progress concerning the proposed tools and redaction of courses/questions.

In the contexte of the CHEL[s] (Collège des Hautes Etudes Lyon Science[s]) Eddy Caron is mentor for the MedTech project building the start-up VETARIA.

Jean-Patrick Gelas manage and build several modules of education about blockchain technologies for an erasmus+ european project. The BLISS project (Blockchain skills for ICT professionals), aims to improve the skills and competencies of ICT professionals by developing and making available educational resources and materials to address existing occupational needs and mismatches, resulting from the dynamic penetration of blockchain technology across all sectors of the EU economy (including banking, accounting, auditing and government services).

10.3.3. Interventions

Laurent Lefèvre has given several invited talks for :

- "Green-It and Video Games infrastructures - Jeu vidéo et collapsologie #2 : Meetup - Le dématérialisé, ça sauve les pandas ?", Meet-up Game Impact, Paris, November 29, 2018
- "Vers des Data Centers éco-responsables", Data Center World, Paris, November 28, 2018
- "Agile (d'accord !) mais Green (d'abord !)", Agile Tour 2018, Epitech, Bordeaux, November 2-3, 2018
- "Numérique et Environnement", with Françoise Berthoud, Groupe de travail développement durable, Centre Inria Bordeaux, October 19, 2018
- "CES 2018 : l'humain numérique de demain...", Journée Réseau Aramis, University Lyon2, Lyon, May 24, 2018
- "Le futur peut-il être numérique ?", with Françoise Berthoud and Jonathan Schaeffer, Journée Réseau Aramis, University Lyon2, Lyon, May 24, 2018
- "EE@G5K : the story continues...", Grid5000 Scientific Advisory Board, Sophia Antipolis, April 4, 2018
- "Rendre l'informatique plus verte", Conférence de formation des professeurs du secondaire en science informatique, Inria Montbonnot, March 21, 2018
- "WG5 : Energy efficiency final results", Network for Sustainable Computing (Nesus), COST European Action IC1305, Final Action Meeting, with Ariel Oleksiak, Madrid, Spain, March 15-16, 2018
- "CES 2018 - so little Green", Rencontres Ecoinfo, Lyon, January 30, 2018

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] H. CROUBOIS. *Toward an autonomic engine for scientific workflows and elastic Cloud infrastructure*, Université de Lyon, October 2018, https://tel.archives-ouvertes.fr/tel-01988995

[2] I. RAÏS. *Discover, model and combine energy leverages for large scale energy efficient infrastructures*, Université de Lyon, September 2018, https://tel.archives-ouvertes.fr/tel-01892387

[3] A. ROUSSEL. *Parallelization of iterative methods to solve sparse linear systems using task based runtime systems on multi and many-core architectures: application to Multi-Level Domain Decomposition methods*, Université Grenoble Alpes, February 2018, https://tel.archives-ouvertes.fr/tel-01753992

Articles in International Peer-Reviewed Journals

[4] F. ALMEIDA, M. DIAS DE ASSUNCAO, J. BARBOSA, V. BLANCO, I. BRANDIC, G. DA COSTA, M. F. DOLZ, A. C. ELSTER, M. JARUS, H. KARATZA, L. LEFÈVRE, I. MAVRIDIS, A. OLEKSIAK, A.-C. ORGERIE, J.-M. PIERSON. *Energy Monitoring as an Essential Building Block Towards Sustainable Ultrascale Systems*, in "Sustainable Computing : Informatics and Systems", March 2018, vol. 17, pp. 27-42 [*DOI :* 10.1016/J.SUSCOM.2017.10.013], https://hal.inria.fr/hal-01627757

[5] A. BENOIT, L. LEFÈVRE, A.-C. ORGERIE, I. RAÏS. *Reducing the energy consumption of large scale computing systems through combined shutdown policies with multiple constraints*, in "International Journal of High Performance Computing Applications", January 2018, vol. 32, n° 1, pp. 176-188 [*DOI :* 10.1177/1094342017714530], https://hal.inria.fr/hal-01557025

[6] J. BIGOT, V. GRANDGIRARD, G. LATU, J.-F. MÉHAUT, L. F. MILLANI, C. PASSERON, S. Q. MASNADA, J. RICHARD, B. VIDEAU. *Building and Auto-Tuning Computing Kernels: Experimenting with BOAST and StarPU in the GYSELA Code*, in "ESAIM: Proceedings and Surveys", October 2018, vol. 63 (2018), pp. 152 - 178 [*DOI :* 10.1051/PROC/201863152], https://hal.inria.fr/hal-01909325

[7] M. DIAS DE ASSUNCAO, A. DA SILVA VEITH, R. BUYYA. *Distributed Data Stream Processing and Edge Computing: A Survey on Resource Elasticity and Future Directions*, in "Journal of Network and Computer Applications", February 2018, vol. 103, pp. 1–17 [*DOI :* 10.1016/J.JNCA.2017.12.001], https://hal.inria.fr/hal-01653842

[8] M. DIAS DE ASSUNCAO, L. LEFÈVRE. *Bare-Metal Reservation for Cloud: an Analysis of the Trade Off between Reactivity and Energy Efficiency*, in "Cluster Computing", 2018, pp. 1-12 [*DOI :* 10.1007/S10586-017-1094-Y], https://hal.inria.fr/hal-01571288

[9] J. V. FERREIRA LIMA, I. RAÏS, L. LEFÈVRE, T. GAUTIER. *Performance and Energy Analysis of OpenMP Runtime Systems with Dense Linear Algebra Algorithms*, in "International Journal of High Performance Computing Applications", August 2018, pp. 1-17 [*DOI :* 10.1177//1094342018792079], https://hal.inria.fr/hal-01957220

[10] I. Raïs, A.-C. Orgerie, M. Quinson, L. Lefèvre. *Quantifying the Impact of Shutdown Techniques for Energy-Efficient Data Centers*, in "Concurrency and Computation: Practice and Experience", 2018, vol. 30, n° 17, pp. 1-13 [*DOI :* 10.1002/cpe.4471], https://hal.archives-ouvertes.fr/hal-01711812

International Conferences with Proceedings

[11] H. Acar, H. Benfenatki, J.-P. Gelas, C. Ferreira Da Silva, G. I. Alptekin, A.-N. Benharkat, P. Ghodous. *Software Greenability: A Case Study of Cloud-Based Business Applications Provisioning*, in "CLOUD 2018 - IEEE 11th International Conference on Cloud Computing", San Francisco, CA, United States, IEEE, July 2018, pp. 875-878 [*DOI :* 10.1109/CLOUD.2018.00125], https://hal.archives-ouvertes.fr/hal-01887065

[12] J. Anjos, K. J. Matteussi, P. R. R. De Souza, C. Geyer, A. da Silva Veith, G. Fedak, J. L. Victoria Barbosa. *Enabling Strategies for Big Data Analytics in Hybrid Infrastructures*, in "2018 International Conference on High Performance Computing Simulation (HPCS)", Orléans, France, July 2018, pp. 869-876, https://hal.inria.fr/hal-01875952

[13] A. Bonnaffoux, E. Caron, H. Croubois, O. Gandrillon. *A Cloud-aware autonomous workflow engine and its application to Gene Regulatory Networks inference*, in "CLOSER 2018 - 8th International conference on Cloud computing and Service Science", Funchal, Madeira, Portugal, March 2018, pp. 1-8, https://hal.inria.fr/hal-01808764

[14] Y. Caniou, E. Caron, A. Kong Win Chang, Y. Robert. *Budget-aware scheduling algorithms for scientific workflows with stochastic task weights on heterogeneous IaaS Cloud platforms*, in "IPDPSW 2018 - IEEE International Parallel and Distributed Processing Symposium Workshops", Vancouver, Canada, IEEE, May 2018, pp. 15-26 [*DOI :* 10.1109/IPDPSW.2018.00014], https://hal.inria.fr/hal-01808831

[15] *Best Paper*
H. Casanova, A. Legrand, M. Quinson, F. Suter. *SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation*, in "EduHPC-18 - Workshop on Education for High-Performance Computing", Dallas, United States, November 2018, pp. 1-10, https://hal.inria.fr/hal-01891513.

[16] H. Casanova, S. Pandey, J. Oeth, R. Tanaka, F. Suter, R. Ferreira Da Silva. *WRENCH: A Framework for Simulating Workflow Management Systems*, in "WORKS 2018 - 13th Workshop on Workflows in Support of Large-Scale Science", Dallas, United States, November 2018, pp. 1-12, https://hal.inria.fr/hal-01948162

[17] *Best Paper*
A. Chai, S. Camarasu-Pop, T. Glatard, H. Benoit-Cattin, F. Suter. *Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems*, in "Europar 2018 - 24th International European Conference on Parallel and Distributed Computing ; Workshop HeteroPar 2018", Turin, Italy, Lecture Notes in Computer Science (LNCS), August 2018, forthcoming, https://hal.archives-ouvertes.fr/hal-01887369.

[18] M. Chardet, H. Coullon, D. Pertin, C. Pérez. *Madeus: A formal deployment model*, in "4PAD 2018 - 5th International Symposium on Formal Approaches to Parallel and Distributed Systems (hosted at HPCS 2018)", Orléans, France, July 2018, pp. 1-8, https://hal.inria.fr/hal-01858150

[19] J. DARROUS, S. IBRAHIM, A. C. ZHOU, C. PÉREZ. *Nitro: Network-Aware Virtual Machine Image Management in Geo-Distributed Clouds*, in "CCGrid 2018 - 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Washington D.C., United States, IEEE, May 2018, pp. 553-562 [*DOI :* 10.1109/CCGRID.2018.00082], https://hal.inria.fr/hal-01745405

[20] T. GAUTIER, C. PÉREZ, J. RICHARD. *On the Impact of OpenMP Task Granularity*, in "IWOMP 2018 - 14th International Workshop on OpenMP for Evolving Architectures", Barcelone, Spain, Springer, September 2018, pp. 205-221 [*DOI :* 10.1007/978-3-319-98521-3_14], https://hal.inria.fr/hal-01901806

[21] B. NOËLLE, E. CARON, A. CHEVALIER, A.-L. VION. *Towards economic and compliant deployment of licenses in a Cloud architecture*, in "Workshop: Cloud Management and Operations. In conjunction with CLOUD 2018 - IEEE International Conference on Cloud Computing", San Franciso, United States, July 2018, https://hal.inria.fr/hal-01808751

[22] I. RAÏS, M. BOUTIGNY, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT. *Building the Table of Energy and Power Leverages for Energy Efficient Large Scale Systems*, in "HPCS: International Conference on High Performance Computing & Simulation", Orléans, France, July 2018, pp. 284-291 [*DOI :* 10.1109/HPCS.2018.00056], https://hal.archives-ouvertes.fr/hal-01845970

[23] I. RAÏS, E. CARON, L. LEFÈVRE. *Thermal aware scheduling on distributed computing water heaters*, in "CCNC 2018 - IEEE Consumer Communications & Networking Conference", Las Vegas, United States, January 2018, pp. 1-4 [*DOI :* 10.1109/CCNC.2018.8319253], https://hal.inria.fr/hal-01671728

[24] I. RAÏS, H. COULLON, L. LEFÈVRE, C. PÉREZ. *Automatic Energy Efficient HPC Programming: A Case Study*, in "ISPA 2018", Melbourne, Australia, December 2018, https://hal.inria.fr/hal-01922301

[25] I. RAÏS, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT. *Exploiting the Table of Energy and Power Leverages*, in "ICA3PP 2018 - 18th International Conference on Algorithms and Architectures for Parallel Processing", Guangzhou, China, November 2018, pp. 1-10, https://hal.archives-ouvertes.fr/hal-01927829

[26] A. DA SILVA VEITH, M. DIAS DE ASSUNCAO, L. LEFÈVRE. *Latency-Aware Placement of Data Stream Analytics on Edge Computing*, in "Service-Oriented Computing", Hangzhou, Zhejiang, China, Springer International Publishing, November 2018, pp. 215-229, https://hal.inria.fr/hal-01875936

[27] P. DE SOUZA JUNIOR, K. J. MATTEUSSI, J. C. S. DOS ANJOS, J. D. D. D. SANTOS, A. DA SILVA VEITH, C. R. GEYER. *Aten: A Dispatcher for Big Data Applications in Heterogeneous Systems*, in "2018 International Conference on High Performance Computing Simulation (HPCS)", Orléans, France, July 2018, pp. 585-592 [*DOI :* 10.1109/HPCS.2018.00098], https://hal.inria.fr/hal-01876973

National Conferences with Proceedings

[28] E. CARON, Z. HOUMANI. *Architecture Microservices pilotée par les données*, in "CAL 2018 - Conférence francophone sur les Architectures Logicielles", Grenoble, France, June 2018, pp. 1-6, https://hal.inria.fr/hal-01808758

Conferences without Proceedings

[29] M. CHARDET, H. COULLON, C. PÉREZ. *Behavioral interfaces for reconfiguration of component models*, in "Compas 2018 - Conférence d'informatique en Parallélisme, Architecture et Système", Toulouse, France, July 2018, pp. 1-8, https://hal.inria.fr/hal-01897803

[30] A. Chevalier. *Vers un déploiement conforme et économique des licences dans une architecture Cloud*, in "COMPAS 2018", Toulouse, France, July 2018, https://hal.archives-ouvertes.fr/hal-01905417

[31] T. Pellissier Tanon, M. Dias de Assuncao, E. Caron, F. M. Suchanek. *Demoing Platypus – A Multilingual Question Answering Platform for Wikidata*, in "ESWC 2018 - Extended Semantic Web Conference", Heracklion, Greece, June 2018, pp. 1-5 [*DOI :* 10.1007/978-3-319-98192-5_21], https://hal.archives-ouvertes.fr/hal-01824972

Research Reports

[32] T. P. Tanon, M. Dias de Assuncao, E. Caron, F. Suchanek. *Platypus – A Multilingual Question Answering Platform for Wikidata*, LIP - ENS Lyon, 2018, https://hal.archives-ouvertes.fr/hal-01730479

Scientific Popularization

[33] G. Amarasinghe, M. Dias de Assuncao, A. Harwood, S. Karunasekera. *A Data Stream Processing Optimisation Framework for Edge Computing Applications*, in "ISORC 2018 - IEEE 21st International Symposium on Real-Time Distributed Computing", Singapore, Singapore, IEEE, May 2018, pp. 91-98 [*DOI :* 10.1109/ISORC.2018.00020], https://hal.inria.fr/hal-01862063

[34] L. L. Nesi, M. Pillon, M. Dias de Assuncao, G. P. Koslovski. *GPU-Accelerated Algorithms for Allocating Virtual Infrastructure in Cloud Data Centers*, in "CCGRID 2018 - 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Washington, France, IEEE, May 2018, pp. 364-365 [*DOI :* 10.1109/CCGRID.2018.00057], https://hal.inria.fr/hal-01862065

[35] A. da Silva Veith, M. Dias de Assuncao, L. Lefèvre. *Latency-Aware Strategies for Placing Data Stream Analytics onto Edge Computing*, July 2018, 1 p. , HotEdge'18 - USENIX Workshop on Hot Topics in Edge Computing, Poster, https://hal.inria.fr/hal-01875999

References in notes

[36] R. Ge, X. Feng, S. Song, H.-C. Chang, D. Li, K. W. Cameron. *PowerPack: Energy Profiling and Analysis of High-Performance Systems and Applications*, in "IEEE Trans. Parallel Distrib. Syst.", May 2010, vol. 21, n^o^ 5, pp. 658–671 [*DOI :* 10.1109/TPDS.2009.76], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4906989

[37] A. Geist, S. Dosanjh. *IESP Exascale Challenge: Co-Design of Architectures and Algorithms*, in "Int. J. High Perform. Comput. Appl.", November 2009, vol. 23, n^o^ 4, pp. 401–402, http://dx.doi.org/10.1177/1094342009347766

[38] W. Gropp, S. Huss-Lederman, A. Lumsdaine, E. Lusk, B. Nitzberg, W. Saphir, M. Snir. *MPI: The Complete Reference – The MPI-2 Extensions*, 2, The MIT Press, September 1998, vol. 2, ISBN 0-262-57123-4

[39] H. Kimura, T. Imada, M. Sato. *Runtime Energy Adaptation with Low-Impact Instrumented Code in a Power-Scalable Cluster System*, in "Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing", Washington, DC, USA, CCGRID '10, IEEE Computer Society, 2010, pp. 378–387

[40] G. MADEC. *NEMO ocean engine*, Institut Pierre-Simon Laplace (IPSL), France, 2008, n^{o} 27, ISSN No 1288-1619

[41] OPENACC. *The OpenACC Application Programming Interface*, November 2011, Version 1.0, http://www.openacc-standard.org

[42] OPENMP ARCHITECTURE REVIEW BOARD. *OpenMP Application Program Interface*, July 2011, Version 3.1, http://www.openmp.org

[43] T. RISTENPART, E. TROMER, H. SHACHAM, S. SAVAGE. *Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds*, in "Proceedings of the 16th ACM conference on Computer and communications security", New York, NY, USA, CCS '09, ACM, 2009, pp. 199–212, http://doi.acm.org/10.1145/1653662.1653687

[44] B. ROUNTREE, D. K. LOWNENTHAL, B. R. DE SUPINSKI, M. SCHULZ, V. W. FREEH, T. BLETSCH. *Adagio: Making DVS Practical for Complex HPC Applications*, in "Proceedings of the 23rd international conference on Supercomputing", New York, NY, USA, ICS '09, ACM, 2009, pp. 460–469

[45] C. SZYPERSKI. *Component Software - Beyond Object-Oriented Programming*, 2, Addison-Wesley / ACM Press, 2002, 608 p.

[46] S. VALCKE. *The OASIS3 coupler: a European climate modelling community software*, in "Geoscientific Model Development", 2013, vol. 6, pp. 373-388, doi:10.5194/gmd-6-373-2013