Activity Report 2014

Project-Team AVALON

Algorithms and Software Architectures for Distributed and HPC Platforms

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

Table of contents

Keywords: Distributed System, Parallel Programming Model, Scheduling, Energy Consumption, Middleware

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

1. Members

Research Scientists

Christian Perez [Team leader, Inria, Senior Researcher, HdR]
Frédéric Desprez [Inria, Senior Researcher, HdR]
Gilles Fedak [Inria, Researcher]
Laurent Lefèvre [Inria, Researcher, HdR]
Frédéric Suter [CNRS, Researcher, HdR]
Marcos Dias de Assunçao [Inria, since Jun 2014]

Faculty Members

Eddy Caron [ENS Lyon, Associate Professor, HdR]
Jean-Patrick Gelas [Univ. Lyon I, Associate Professor]
Olivier Glück [Univ. Lyon I, Associate Professor]

Engineers

Simon Delamare [CNRS]
Matthieu Imbert [Inria]
Jean-Christophe Mignot [CNRS]
Ghislain Landry Tsafack Chetsa [Inria, until Nov 2014, granted by Caisse des Dépôts et Consignations]
Nicolas Bard [CNRS, until Feb 2014]
Abderhaman Cheniour [CNRS, since Jul 2014]
Marc Pinhède [Inria]
Sylvain Bernard [ENS Lyon, until Jul 2014]
Amin Bsila [SysFera, provisional assignment until May 2014]
Laurent Pouilloux [Inria, granted by Caisse des Dépôts et Consignations]
François Rossigneux [Inria, until Dec 2014, granted by Caisse des Dépôts et Consignations]
Guillaume Verger [Inria, until Oct 2014, granted by CapRezo]
Huaxi Zhang [Inria, until Sep 2014, granted by Min. de l'Economie]

PhD Students

Sylvain Gault [Inria]
Pedro Silva [Inria, from Oct 2014]
Anthony Simonet [Inria]
Violaine Villebonnet [Inria]
Daniel Balouek Thomert [NewGeneration SR]
Radu Carpa [ENS Lyon, from Mar 2014]
Maurice Faye [Joint PhD with the University Gaston Berger, St Louis, Sénégal; UGB grant, CMIRA grant, and LIP grant]
Vincent Lanore [ENS Lyon]
Arnaud Lefray [INSA Centre Val de Loire]
Jérôme Richard [Inria, since Nov 2014]
Semen Marchuk [Inria, since Dec 2014]

Post-Doctoral Fellows

Jonathan Rouzaud-Cornabas [Inria, until Aug 2014]
Hélène Coullon [Inria, from Oct 2014]

Lamiel Toch [Inria, until Mar 2014]

Visiting Scientists
Julio Anjos [PhD Student, Univ. Fed. do Rio Grande do Sul, Brazil, from May 2014]
Asma Ben Cheikh Ahmed [PhD Student, Univ. Tunis, Tunisia, Sep 15 – Dec 15 2014]
Mircea Moca [ENS Lyon, Invited Professor, Nov 15 – Dec 15 2014]
Tchimou N'Takpé [Assistant Professor, Univ. Nangui Abrogoua, Abidjan, Oct 2014 – Nov 2014]
Qian Zhang [PhD Student, Australian National University, Australia, Sep – Nov 2014]

Administrative Assistant
Evelyne Blesle [Inria]

Other
David Loup [INSA Lyon / Inria part-time apprentice, from Sep 2014]

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. Those large scale platforms based on the aggregation of large clusters (Grids), huge datacenters (Clouds), collections of volunteer PC (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 different 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 not interested in the resources used during the execution. Instead, they are interested in how their application is going to be executed: in which duration, at which cost, what is 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 (maximing 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 Modelization*

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. The 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.

2.2.4. Application Mapping and Scheduling

Avalon will propose multi-criteria mapping and scheduling algorithms to meet the challenge of automatizing 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 testbed GRID'5000 will be our platform of choice for experiments.

3. Research Program

3.1. Energy Application Profiling and Modelization

International roadmaps schedule to build exascale systems by the 2018 time frame. According to the Top500 list published in November 2013, the most powerful supercomputer is the Tianhe-2 platform, a machine with more than 3,000,000 cores. It consumes more than 17 MW for a maximum performance of 33 PFlops while the Defense Advanced Research Projects Agency (DARPA) has set to 20 MW the maximum energy consumption of an exascale supercomputer [32].

Energy efficiency is therefore a major challenge for building next generation large scale platforms. The targeted platforms will gather hundreds of million 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 the measure, the understanding, and the analysis of the large-scale platform energy consumption. Unlike approaches [34] 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 resources 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 [37], phase detection for specific HPC applications [42], etc.). As a second step, we aim at designing a framework model that allows interactions, dialogues and decisions taken in cooperation between 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 of the time implicitly linked to "analytics" to refer to issues such as 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 infrastructure 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. Resourc-Agnostic Application Description Model

When programming in parallel, users expect to obtain performance improvement, whatever the cost is. For long, parallel machines have been simple enough to let a user program them given a minimal abstraction of their hardware. For example, MPI [36] exposes the number of nodes but hides the complexity of network topology behind a set of collective operations; OpenMP [40] simplifies the management of threads on top of a shared memory machine while OpenACC [39] 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 [35]. 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 [43] 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.

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 others research axis. We plan to consider application mapping and scheduling through 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 propose scheduling algorithms for dynamic and elastic platforms. As the amount 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 can not 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 [41] and several issues have been raised in Clouds. Nowadays virtualization is used as the sole mechanism to secure different users sharing resources on Clouds. But, due to improper virtualization of all the components of Clouds (such as micro-architectural components), 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 bring 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 [38] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [44] 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 comsunption, 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 volume of data. For instance, the Large Synoptic Survey Telescope (http://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 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 cooperates 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. These collaborations bring scientific applications that are both dynamic and data-intensive.

5. New Software and Platforms

5.1. BitDew/Active Data

Participants: Gilles Fedak [correspondant], Anthony Simonet.

BITDEW is an open source middleware implementing a set of distributed services for large scale data management on Desktop Grids and Clouds. BITDEW relies on five abstractions to manage the data : i) replication indicates how many occurrences of a data should be available at the same time on the network, ii) fault-tolerance controls the policy in presence of hardware failures, iii) lifetime is an attribute absolute or relative to the existence of other data, which decides of the life cycle of a data in the system, iv) affinity drives movement of data according to dependency rules, v) protocol gives the runtime environment hints about the protocol to distribute the data (http, ftp, or bittorrent). Programmers define for every data these simple criteria, and let the BITDEW runtime environment manage operations of data creation, deletion, movement, replication, and fault-tolerance operation.

BITDEW is distributed open source under the GPLv3 or Cecill licence at the user's choice. 10 releases were produced over the last two years, and it has been downloaded approximately 6,000 times on the Inria forge. Known users are Université Paris-XI, Université Paris-XIII, University of Florida (USA), Cardiff University (UK) and University of Sfax (Tunisia). In terms of support, the development of BitDew is partly funded by the Inria ADT BitDew and by the ANR MapReduce projects. Thanks to this support, we have developed and released the first prototype of the MapReduce programming model for Desktop Grids on top of BitDew. In 2012, 8 versions of the software have been released, including the version 1.2.0 considered as a stable release of BitDew with many advanced features. Our most current extension focuses on Active Data, which is a data-centric and event-driven programming model combined with a runtime environment, which allows to expose and manage data set life cycle. Active Data strength is to facilitate the development of applications that handle dynamic data sets distributed on heterogeneous systems and infrastructures.

5.2. DIET

Participants: Daniel Balouek Thomert, Eddy Caron [correspondant], Frédéric Desprez, Maurice Faye, Arnaud Lefray, Guillaume Verger, Jonathan Rouzaud-Cornabas, Lamiel Toch, Huaxi Zhang.

Huge problems can now be processed over the Internet thanks to Grid and Cloud middleware systems. The use of on-the-shelf applications is needed by scientists of other disciplines. Moreover, the computational power and memory needs of such applications may of course not be met by every workstation. Thus, the RPC paradigm seems to be a good candidate to build Problem Solving Environments on the Grid or Cloud. The aim of the DIET project (http://graal.ens-lyon.fr/DIET) is to develop a set of tools to build computational servers accessible through a GridRPC API.

Moreover, the aim of a middleware system such as DIET is to provide a transparent access to a pool of computational servers. DIET focuses on offering such a service at a very large scale. A client which has a problem to solve should be able to obtain a reference to the server that is best suited for it. DIET is designed to take into account the data location when scheduling jobs. Data are kept as long as possible on (or near to) the computational servers in order to minimize transfer times. This kind of optimization is mandatory when performing job scheduling on a wide-area network. DIET is built upon *Server Daemons*. The scheduler is scattered across a hierarchy of *Local Agents* and *Master Agents*. Applications targeted for the DIET platform are now able to exert a degree of control over the scheduling subsystem via *plug-in schedulers*. As the applications that are to be deployed on the Grid vary greatly in terms of performance demands, the DIET plug-in scheduler facility permits the application designer to express application needs and features in order that they be taken into account when application tasks are scheduled. These features are invoked at runtime after a user has submitted a service request to the MA, which broadcasts the request to its agent hierarchy.

DIET provide a support for Cloud architecture. and it takes benefits from virtualized resources. As cloud resources are dynamic, we have on-going research in the field of automatic and elastic deployment for middleware systems. DIET will be able to extend and reduce the amount on aggregated resources and adjust itself when resources fail.

In the context of the Seed4C project, we have studied how secured our platform, authenticated and secured interactions between the different parts of our middleware and between our middleware and its users. By the way, we have added the SSL support into the DIET communication layer. We have worked to show how to securely use public cloud storage without taking the risk of losing confidentiality of data stored on them.

We have started a work to design a plug-in schedulers into DIET to deal with energy management. Using this scheduler we have obtain a significatif gain close to 25% with a minor weakening of performance (6%). Moreover we have experimented some dynamic resources management through DIET based on the energy criteria.

5.3. Sam4c

Participants: Eddy Caron, Arnaud Lefray [correspondant], Jonathan Rouzaud-Cornabas.

Sam4C (https://gforge.inria.fr/projects/sam4c/) -Security-Aware Models for Clouds- is a graphical and textual editor to model Cloud applications (as virtual machines, processes, files and communications) and describe its security policy. Sam4C is suitable to represent any static application without deadline or execution time such as n-tiers or parallel applications. This editor is generated in Java from an EMF -Eclipse Modeling Framework-metamodel to simplify any modifications or extensions. The application model and the associated security policy are compiled in a single XML file which serves as input for an external Cloud security-aware scheduler. Alongside with this editor, Cloud architecture models and provisioning algorithms are provided for simulation (in the current version) or real deployments (in future versions). During this step of development this software is private and available only for Seed4C project members. The design of Sam4c is a joint effort with INSA Centre Val de Loire.

5.4. SimGrid

Participants: Jonathan Rouzaud-Cornabas, Frédéric Suter [correspondant].

SIMGRID is a toolkit for the simulation of distributed applications in heterogeneous distributed environments. The specific goal of the project is to facilitate research in the area of parallel and distributed large scale systems, such as Grids, P2P systems and clouds. Its use cases encompass heuristic evaluation, application prototyping or even real application development and tuning. SIMGRID has an active user community of more than one hundred members, and is available under GPLv3 from http://simgrid.gforge.inria.fr/.

5.5. HLCMi, L²C, & Gluon++

Participants: Hélène Coullon, Vincent Lanore, Christian Perez [correspondant], Jérôme Richard.

HLCMI (http://hlcm.gforge.inria.fr) is an implementation of the HLCM component model. HLCM is a generic extensible component model with respect to component implementations and interaction concerns. Moreover, HLCM is abstract; it is its specialization—such as HLCM/L²C—that defines the primitive elements of the model, such as the primitive components and the primitive interactions.

HLCMI is making use of Model-driven Engineering (MDE) methodology to generate a concrete assembly from an high level description. It is based on the Eclipse Modeling Framework (EMF). HLCMI contains 700 Emfatic lines to describe its models and 7000 JAVA lines for utility and model transformation purposes. HLCMI is a general framework that supports several HLCM specializations: HLCM/CCM, HLCM/JAVA, HLCM/L²C and HLCM/Charm++ (known as Gluon++).

L^{2}C (http://hlcm.gforge.inria.fr) is a *Low Level Component* model implementation targeting at use-cases where overhead matters such as High-Performance Computing. L^{2}C does not offer network transparency neither language transparency. Instead, L^{2}C lets the user choose between various kinds of interactions between components, some with ultra low overhead and others that support network transport. L^{2}C is extensible as additional interaction kinds can be added quite easily. L^{2}C currently supports C++, FORTRAN 2013, MPI and CORBA interactions.

Gluon++ (http://hlcm.gforge.inria.fr) is a thin component model layer added on top of Charm++ (http://charm.cs.uiuc.edu/). It defines chare components as a Charm++ chare with minimal metadata, C++ components as a C++ class with minimal metadata, (asynchronous) entry method calls between components, and plain C++ method calls between components.

L^{2}C and Gluon++ are implemented in the LLCMc++ framework (http://hlcm.gforge.inria.fr). It is distributed under a LGPL licence and represents 6400 lines of C++.

5.6. Execo

Participants: Matthieu Imbert [correspondant], Laurent Pouilloux.

Execo(http://execo.gforge.inria.fr) is a Python library designed for rapid prototyping of experiments on distributed systems, automatization of system administration tasks (such as deployment and configuration of distributed middleware), and creation of reproducible experiments scripts. It allows easy and asynchronous management of thousands of local or remote unix processes and offers tools for easy usage of the Grid'5000 platform services.

Execo currently has more than 20 users in and outside the AVALON team, who rely on it to automate experimental workflows. It was used to develop one of the two contenders who won the 2014 Grid'5000 Large Scale Deployment Challenge. It is used as a building block in the Grid'5000 metrology service and has been used to produce experimental results involved in numerous papers and reports.

It is distributed under GPLv3 and it is made of 7200 lines of code.

5.7. Kwapi

Participants: Laurent Lefèvre [correspondant], François Rossigneux, Jean-Patrick Gelas, Laurent Pouilloux.

Kwapi (https://launchpad.net/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.

5.8. Platforms

5.8.1. Grid'5000

Participants: Frédéric Desprez, Simon Delamare, Laurent Lefèvre, David Loup, Christian Perez, Marc Pinhède, Laurent Pouilloux.

The GRID'5000 experimental platform (http://www.grid5000.fr) 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.

Not only GRID'5000 is heavily used for Avalon research, but several team members are also involved in GRID'5000 direction:

- Frédéric Desprez is leading the "Groupement d'Intérêt Scientifique Groupement Grille 5K" which drives GRID'5000.

- Laurent Lefèvre is responsible of the GRID'5000 Lyon platform and member of the GRID'5000 direction committee.

- Christian Perez is leading the Hemera initiative (https://www.grid5000.fr/Hemera) and he is a member of the GRID'5000 direction committee.

- Simon Delamare is the plateform's operational manager.

Avalon also provides an important effort for Grid'5000 operation and development by hosting several enginneers belonging to Grid'5000 technical team (Marc Pinhède, David Loup) or HEMERA IPL (Laurent Pouilloux).

6. New Results

6.1. Energy efficiency of large scale distributed systems

Participants: Laurent Lefèvre, Daniel Balouek Thomert, Eddy Caron, Radu Carpa, Ghislain Landry Tsafack Chetsa, Marcos Dias de Assunçao, Jean-Patrick Gelas, Olivier Glück, Jean-Christophe Mignot, François Rossigneux, Violaine Villebonnet.

6.1.1. *Improving Energy Efficiency of Large Scale Systems without a priori Knowledge of Applications and Services*

Unlike their hardware counterpart, software solutions to the energy reduction problem in large scale and distributed infrastructures hardly result in real deployments. At the one hand, this can be justified by the fact that they are application oriented. At the other hand, their failure can be attributed to their complex nature which often requires vast technical knowledge behind proposed solutions and/or thorough understanding of applications at hand. This restricts their use to a limited number of experts, because users usually lack adequate skills. In addition, although subsystems including the memory and the storage are becoming more and more power hungry, current software energy reduction techniques fail to take them into account. We propose a methodology for reducing the energy consumption of large scale and distributed infrastructures. Broken into three steps known as (i) phase identification, (ii) phase characterization, and (iii) phase identification and system reconfiguration; our methodology abstracts away from any individual applications as it focuses on the infrastructure, which it analyses the runtime behaviour and takes reconfiguration decisions accordingly.

The proposed methodology is implemented and evaluated in high performance computing (HPC) clusters of varied sizes through a Multi-Resource Energy Efficient Framework (MREEF). MREEF implements the proposed energy reduction methodology so as to leave users with the choice of implementing their own system reconfiguration decisions depending on their needs. Experimental results show that our methodology reduces the energy consumption of the overall infrastructure of up to 24% with less than 7% performance degradation. By taking into account all subsystems, our experiments demonstrate that the energy reduction problem in large scale and distributed infrastructures can benefit from more than "the traditional" processor frequency scaling. Experiments in clusters of varied sizes demonstrate that MREEF and therefore our methodology can easily be extended to a large number of energy aware clusters. The extension of MREEF to virtualized environments like cloud shows that the proposed methodology goes beyond HPC systems and can be used in many other computing environments.

6.1.2. Reservation based Usage for Energy Efficient Clouds: the Climate/Blazar Architecture

The FSN XLcloud project (cf Section 8.1) strives to establish the demonstration of a High Performance Cloud Computing (HPCC) platform based on OpenStask, that is designed to run a representative set of compute intensive workloads, including more specifically interactive games, interactive simulations and 3D graphics. XLcloud is based on OpenStack, and Avalon is contributing to the energy efficiency part of this project. We have proposed and brought our contribution to Climate, a new resource reservation framework for OpenStack, developed in collaboration with Bull, Mirantis and other OpenStack contributors. Climate allows the reservation of both physical and virtual resources, in order to provide a mono-tenancy environment suitable for HPC applications. Climate chooses the most efficient hosts (flop/W). This metric is computed from the CPU / GPU informations, mixed with real power consumption measurements provided by the Kwapi framework. The user requirements may be loose, allowing Climate to choose the best time slot to place the reservation. Climate has been improved with standby mode features, to shut down automatically the unused hosts. The first release of Climate was done in January 2014. Through the OpenStack process, Climate is now named Blazar.

6.1.3. Clustered Virtual Home Gateway (vHGW)

This result is a joint work between Avalon team (J.P. Gelas, L. Lefevre) and Addis Abeba University (M. Tsibie and T. Assefa). The customer premises equipment (CPE), which provides the interworking functions between the access network and the home network, consumes more than 80% of the total power in a wireline access network. In the GreenTouch initiative (cf Section 8.3), we aim at a drastic reduction of the power consumption by means of a passive or quasi-passive CPE. Such approach requires that typical home gateway functions, such as routing, security, and home network management, are moved to a virtual home gateway (vHGW) server in the network. In our first prototype virtual home gateways of the subscribers were put in LXC containers on a unique GNU/Linux server. The container approach is more scalable than separating subscribers by virtual machines. We demonstrated a sharing factor of 500 to 1000 virtual home gateways on one server, which consumes about 150 W, or 150 to 300 mW per subscriber. Comparing this power consumption with the power of about 2 W for the processor in a thick client home gateway, we achieved an efficiency gain of 5-10x. The prototype was integrated and demonstrated at TIA 2012 in Dallas. In our current work, we propose the Clustered vHGWs Data center architecture to yield optimal energy conservation through virtual machine's migration among physical nodes based on the current subscriber's service access state, while ensuring SLA respective subscribers. Thus, optimized energy utilization of the data center is assured without compromising the availability of service connectivity and QoS preferences of respective subscribers. The last prototype including those new features was integrated and demonstrated recently to the GreenTouch consortium members at Melbourne University.

6.1.4. Energy proportionality with heterogeneous computing resources

This work [16] focuses on improving energy proportionality of large scale virtualized environments. The main problem of such infrastructures is their high static costs due to high idle power consumption of idle servers. Our goal is to reach an infrastructure able to adapt its energy consumption to the current working load. Therefore we propose an original infrastructure composed of heterogeneous computing resources. We consider the heterogeneity at the level of the architecture, and we gather in our platform low power ARM processors together with powerful x86 servers. Around this infrastructure, we are developing a decisional framework to schedule applications on the architecture, or combination of architectures, most suitable to their current needs. The framework reacts dynamically to the resource needs evolutions by migrating the applications to the chosen destinations, and switching off unused nodes to save energy. We validate our scheduling policies by building a simulator based on a set of experimental inputs about power and performance hardware profiles and applications load profiles. This work is jointly done with IRIT Lab. (Toulouse) under the support of Inria Large Scale Initiative Hemera.

6.1.5. Energy efficient Core Networks

This work [11] seeks to improve the energy efficiency of backbone networks by providing an intra-domain Software Defined Network (SDN) approach to selectively turn off a subset of links. To do this, we change

the status of router ports and transponders on the two extremities of a link. The status of these components is set to sleep mode whenever a link is not required to transfer data, and brought back to operational state when needed. We have analyzed the implementation issues of an energy-efficient SDN-based traffic engineering in core networks. We propose the STREETE framework (SegmenT Routing based Energy Efficient Traffic Engineering) that represents an online method to switch some links off/on dynamically according to the network load. We have implemented our proposed algorithms in the OMNET++ packet-based discrete event simulator. Experiments considering real network topologies (Germany50 and Ge´ant) and real dynamic traffic matrices allowed us to quantify the trade-off between energy saving and impact of our solution on network performance. As mean to reroute the traffic we use a promising new protocol, SPRING. This comes in contrast with other works, which use classical IP link weights changes or MPLS+RSVP-TE for this purpose. SPRING proved itself well suited for dynamic reconfiguration of the network. Experimental results show that the consumption of 44% of links can be reduced while preserving good quality of service.

6.1.6. *Energy aware scheduling for multi data centers clouds*

Our work tackles the challenge of improving the energy efficiency of server provisioning and workload management [17]. It introduces a metric allowing infrastructure administrators to specify their preferences between performance and energy savings. We describe a framework for resource management which provides control for informed and automated provisioning at the scheduler level while providing developers (administrator or end-user) with an abstract layer to implement aggregation and resource ranking based on contextual information such as infrastructure status, users' preferences and energy- related external events occurring over time. We integrate our solution in DIET which allows for managing heterogeneous nodes at the middleware layer. The evaluation is performed by means of simulations and real-life experiments on the GRID'5000 testbed. Results show improvements in energy efficiency with minimal impact on application and system performance. Implementation has been used within the industrial project Nu@ge in the context of a federation of modular datacenters.

6.2. Simulation of Large Scale Distributed Systems

Participants: Frédéric Desprez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.2.1. *Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*

The study of parallel and distributed applications and platforms, whether in the cluster, grid, peer-to-peer, volunteer, or cloud computing domain, often mandates empirical evaluation of proposed algorithmic and system solutions via *simulation*. Unlike direct experimentation via an application deployment on a real-world testbed, simulation enables fully repeatable and configurable experiments for arbitrary hypothetical scenarios. Two key concerns are accuracy (so that simulation results are scientifically sound) and scalability (so that simulation experiments can be fast and memory-efficient). While the scalability of a simulator is easily measured, the accuracy of many state-of-the-art simulators is largely unknown because they have not been sufficiently validated. In this work we describe recent accuracy and scalability advances made in the context of the SIMGRID simulation framework. A design goal of SIMGRID is that it should be versatile, i.e., applicable across all aforementioned domains. We present quantitative results that show that SIMGRID compares favorably to state-of-the-art domain-specific simulators in terms of scalability, accuracy, or the trade-off between the two. An important implication is that, contrary to popular wisdom, striving for versatility in a simulator is not an impediment but instead is conducive to improving both accuracy and scalability.

6.2.2. *Simulation of MPI Applications with Time-Independent Traces*

Analyzing and understanding the performance behavior of parallel applications on parallel computing platforms is a long-standing concern in the High Performance Computing community. When the targeted platforms are not available, simulation is a reasonable approach to obtain objective performance indicators and explore various hypothetical scenarios. In the context of applications implemented with the Message Passing Interface, two simulation methods have been proposed, on-line simulation and off-line simulation, both with their own drawbacks and advantages. In this work we present an off-line simulation framework, i.e., one that

simulates the execution of an application based on event traces obtained from an actual execution. The main novelty of this work, when compared to previously proposed off-line simulators, is that traces that drive the simulation can be acquired on large, distributed, heterogeneous, and non-dedicated platforms. As a result the scalability of trace acquisition is increased, which is achieved by enforcing that traces contain no time-related information. Moreover, our framework is based on an state-of-the-art scalable, fast, and validated simulation kernel.

6.2.3. Adding Storage Simulation Capacities to the SimGrid Toolkit

For each kind of distributed computing infrastructures, i.e., clusters, grids, clouds, data centers or supercomputers, storage is a essential component to cope with the tremendous increase in scientific data production and the ever-growing need for data analysis and preservation. Understanding the performance of a storage subsystem or dimensioning it properly is an important concern for which simulation can help by allowing for fast, fully repeatable, and configurable experiments for arbitrary hypothetical scenarios. However, most simulation frameworks tailored for the study of distributed systems offer no or little abstractions or models of storage resources.

In this work we extend SimGrid, a versatile toolkit for the simulation of large-scale distributed computing systems, with storage simulation capacities. We define the required abstractions and propose a new API to handle storage components and their contents in SimGrid-based simulators. Then we characterize the performance of the fundamental storage component that are disks and derive models of these resources. Finally we list several concrete use cases of storage simulations in clusters, grids, clouds, and data centers for which the proposed extension would be beneficial.

6.3. MapReduce Computations on Hybrid Distributed Computations Infrastructures

Participants: Gilles Fedak, Julio Anjos, Asma Ben Cheikh Ahmed.

In this section we report on our efforts to provide MapReduce Computing environments on Hybrid infrastructures, i.e composed of Desktop Grids and Cloud computing environments.

6.3.1. BIGhybrid - A Toolkit for Simulating MapReduce in Hybrid Infrastructures

Cloud computing has increasingly been used as a platform for running large business and data processing applications. Although clouds have become extremely popular, when it comes to data processing, their use incurs high costs. Conversely, Desktop Grids, have been used in a wide range of projects, and are able to take advantage of the large number of resources provided by volunteers, free of charge. Merging cloud computing and desktop grids into a hybrid infrastructure can provide a feasible low-cost solution for big data analysis. Although frameworks like MapReduce have been devised to exploit commodity hardware, their use in a hybrid infrastructure raise some challenges due to their large resource heterogeneity and high churn rate. This study introduces BIGhybrid, a toolkit that is used to simulate MapReduce in hybrid environments. Its main goal is to provide a framework for developers and system designers that can enable them to address the issues of Hybrid MapReduce. In this paper, we describe the framework which simulates the assembly of two existing middleware: BitDew- MapReduce for Desktop Grids and Hadoop-BlobSeer for Cloud Computing. The experimental results that are included in this work demonstrate the feasibility of our approach.

6.3.2. Parallel Data Processing in Dynamic Hybrid Computing Environment Using MapReduce

In this work, we propose a novel MapReduce computation model in hybrid computing environment called HybridMR is proposed. Using this model, high performance cluster nodes and heterogeneous desktop PCs in Internet or Intranet can be integrated to form a hybrid computing environment. In this way, the computation and storage capability of large-scale desktop PCs can be fully utilized to process large-scale datasets. HybridMR relies on a hybrid distributed file system called HybridDFS, and a time-out method has been used in HybridDFS to prevent volatility of desktop PCs, and file replication mechanism is used

to realize reliable storage. A new node priority-based fair scheduling (NPBFS) algorithm has been developed in HybridMR to achieve both data storage balance and job assignment balance by assigning each node a priority through quantifying CPU speed, memory size and I/O bandwidth. Performance evaluation results show that the proposed hybrid computation model not only achieves reliable MapReduce computation, reduces task response time and improves the performance of MapReduce, but also reduces the computation cost and achieves a greener computing mode.

6.3.3. *Ensuring Privacy for MapReduce on Hybrid Clouds Using Information Dispersal Algorithm*

MapReduce is a powerful model for parallel data processing. The motivation of this work is to allow running map-reduce jobs partially on untrusted infrastructures, such as public Clouds and Desktop Grid, while using a trusted infrastructure, such as private cloud, to ensure that no outsider could get the 'entire' information. Our idea is to break data into meaningless chunks and spread them on a combination of public and private clouds so that the compromise would not allow the attacker to reconstruct the whole data-set. To realize this, we use the Information Dispersion Algorithms (IDA), which allows to split a file into pieces so that, by carefully dispersing the pieces, there is no method for a single node to reconstruct the data if it cannot collaborate with other nodes. We propose a protocol that allows MapReduce computing nodes to exchange the data and perform IDA-aware MapReduce computation. We conduct experiments on the Grid'5000 testbed and report on performance evaluation of the prototype.

6.4. Using Active Data to Provide Smart Data Surveillance to E-Science Users

Participants: Gilles Fedak, Anthony Simonet.

Large scientific experiments drive scientists to use many storage and computing platforms as well as different applications, tools and analysis scripts. The resulting heterogeneous environments make data management operations challenging; the significant number of events and the absence of data integration makes it difficult to track data provenance, manage sophisticated analysis processes, and recover from unexpected situations. Current approaches often require costly human intervention and are inherently error prone. The difficulty managing and manipulating such large and highly distributed datasets also limits automated sharing and collaboration. In this collaboration with Kyle Chard and Ian Foster from Argonne National Lab and University of Chicago, we study a real world e-Science application involving terabytes of data, using three different analysis and storage platforms, and a number of applications and analysis processes. We demon- strate that using a specialized data life cycle and programming model—Active Data—we can easily implement global progress monitoring, sharing and recovery from unexpected events in heterogeneous environments and automate human tasks.

6.5. HPC Component Model

Participants: Hélène Coullon, Vincent Lanore, Christian Perez, Jérôme Richard.

6.5.1. *3D FFT and L^{2}C*

We have studied the relevance of dealing with 3D FFT parallel algorithms with the software component model L^{2}C [31]. We have implemented several existing 3D FFT algorithms, and we have evaluated their performance, their scalability, and their reuse rate. Experiments made on clusters of Grid'5000 and on the Curie supercomputer up to 8192 cores show that L2C based 3D assemblies are scalable and have the same kind of performance than existing 3D libraries such as FFTW or 2DECOMP. This work confirms than components can be used for optimized HPC applications

6.5.2. *Stencil Skeletons in L^{2}C*

Mesh-based scientific simulation is an important class of scientific application which could benefit from component models. Therefore, we have studied and designed a first adaptation of the SIPSim model [33] (Structured Implicit Parallelism for scientific Simulations) to handle HPC component models. The heat equation application has been implemented on top of L^{2}C following this adapted SIPSim model. First experiments on clusters of Grid'5000 and on the Curie supercomputer show promising results, of which a complete analysis is still ongoing. This work is a first step toward a complete implicit parallelism stencil skeleton using L^{2}C.

6.5.3. *Reconfigurable HPC component model*

High-performance applications whose structure changes dynamically during execution are extremely complex and costly to develop, maintain and adapt to new hardware. Such applications would greatly benefit from easy reuse and separation of concerns which are typical advantages of component models. Unfortunately, no existing component model is both HPC-ready (in terms of scalability and overhead) and able to easily handle dynamic reconfiguration.

We aim at addressing performance, scalability and programmability by separating locking and synchronization concerns from reconfiguration code. To this end, we have defined *directMOD*, a component model which provides on one hand a flexible mechanism to lock subassemblies with a very small overhead and high scalability, and on the other hand a set of well-defined mechanisms to easily plug various independently-written reconfiguration components to lockable subassemblies. We evaluate both the model itself and a C++/MPI implementation called *directL2C* based on L^{2}C.

6.6. Security for Virtualization and Clouds

Participants: Eddy Caron, Arnaud Lefray, Jonathan Rouzaud-Cornabas.

Our framework Security Aware Models for Clouds has two purposes. The first one is, for a client, to model an IaaS application composed of virtual machines, applications, datas and communications and specify the associated security requirements. The whole modelization is contained into a XML file. The second one is the scheduling. It takes as inputs application models (XML) and the infrastructure of the cloud (currently in XML) i.e. a hierarchical set of physical machines. The scheduler encapsulates applications into virtual machines when needed and then maps virtual machines onto physical machines. The result of this scheduling is a file with the mapping i.e. a list of (VM, PM) couples.

The scheduler, as a standalone engine, can be used as simulator. But it can be interfaced with a Cloud stack (e.g. OpenStack, OpenNebula) to act as a production scheduler. This inferfacing is achieved by dynamically inferring the infrastructure model from the Cloud database and applying the decision i.e the output mapping list. Furthermore, the security policies (as input) are splitted for local security enforcement on each physical machine.

Sam4C (Security-Aware Models For Clouds) is a twofold framework, namely Sam4C-Modeler and Sam4C-Scheduler. The first is dedicated to modeling an application with the tenant's virtual machines and network interconnection. The second is is a security-aware scheduler, meaning it overrides the basic default scheduler with mainly the following enhanced capabilities

We have designed a scheduling module called SPS. This module is designed to support all the operations concerning the Cloud. It is based on the OpenStack and extends OpenStack with security aspects to fulfil the requirements of Seed4C.

6.7. Locality-aware Cooperation for VM Scheduling in Distributed Clouds

Participant: Frédéric Desprez.

In collaboration with the Ascola team (A. Lèbre, J. Pastor), ASAP team (Marin Bertier), and the Myriads team (C. Tedeschi), we worked on the design of a distributed Cloud Computing infrastructure [23]. The promotion of such infrastructures as the next platform to deliver the Utility Computing paradigm, leads to new virtual machines (VMs) scheduling algorithms leveraging peer-to-peer approaches. Although these proposals considerably improve the scalability, leading to the management of hundreds of thousands of VMs over thousands of physical machines (PMs), they do not consider the network overhead introduced by multi-site infrastructures. This overhead can have a dramatic impact on the performance if there is no mechanism favoring intra-site versus inter-site manipulations.

In 2014, we designed a new building block designed on top of a network with Vivaldi coordinates maximizing the locality criterion (*i.e.,* efficient collaborations between PMs) [12]. We combined such a mechanism with DVMS, a large-scale virtual machine scheduler and showed its benefit by discussing several experiments performed on four distinct sites of the Grid'5000 testbed. With our proposal and without changing the scheduling decision algorithm, the number of inter-site operations has been reduced by 72%. This result provides a glimpse of the promising future of using locality properties to improve the performance of massive distributed Cloud platforms.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. *Animerique*

One of the goals of the CapRézo company is to provide an original tool to make 2D/3D animation films. This tool is an innovative and distributed numerical platform. This platform is built on software developed by Avalon like DIET. Technologies developed in collaboration between CapRézo and Inria are based on Cloud federation environment. The collaboration, started in 2014, is scheduled for the next 5 years.

7.2. Bilateral Grants with Industry

7.2.1. *NewGeneration-SR*

We have a collaboration with the company NewGeneration-SR. The aim of this company is to reduce the energy impact through solutions on each layer of the energy consumption (from the data- center design and the production to usage). NewGeneration-SR improve the life cycle (design, production, recycling) in order to reduce the environmental impact of it. NewGeneration-SR was member of the Nu@ge consortium: one of five national Cloud Computing projects with "emprunts d'avenir" funding. With a CIFRE PhD student (Daniel Balouek), we are developing models to reduce the energy consumption for the benefit of data-center

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. *French National Fund for the Digital Society Project (FSN)*

8.1.1.1. FSN XLcloud, 2012-2014
Participants: Jean-Patrick Gelas, Laurent Lefèvre, François Rossigneux.

Focused on high-performance computing, the XLcloud collaborative project sets out to define and demonstrate a cloud platform based on *HPC-as-a-Service*. This is designed for computational intensive workloads, with interactive remote visualisation capabilities, thus allowing different users to work on a common platform. XLcloud project's members design, develop and integrate the software elements of a High Performance Cloud Computing (HPCC) System.

Expected results of the projects include : Functional and technical specification of the XLcloud platform architecture, open source API of the XLcloud platform, implementation of algorithms for 3D and video streaming display, prototype of the XLcloud platform including the support of on-demand virtual clusters and remote visualisation service, use cases for validation, illustrating the performance and suggesting future improvements.

XLcloud aims at overcoming some of the most important challenges of implementing operationally high performance applications in the Cloud. The goal is to allow partners of the project to take leadership position in the market, as cloud service providers, or as technology providers. XLcloud relies on a consortium of various partners (BULL (project leader), TSP, Silkan, EISTI, Ateme, Inria, CEA List, OW2, AMG.Lab).

In this project, the Avalon team investigates the issue of energy awareness and energy efficiency in OpenStack Cloud based platforms.

8.1.2. *French National Research Agency Projects (ANR)*

8.1.2.1. *ANR EMERGENCE CloudPower, Cloud Service providing HPC on-demand to innovative SME's, 35 months, ANR-12-EMMA-0038*
Participants: Gilles Fedak, Sylvain Bernard.

High performance computing (HPC) allows scientists and industries to run large numerical application on huge data volumes. The HPC is a key factor in knowledge and innovation in many fields of industry and service, with high economic and social issues: aerospace, finance and business intelligence, energy and environment, chemicals and materials, medicine and biology , digital art and games, Web and social networks, ... Today, acquiring HPC supercomputer is very expensive, making HPC unreachable to SMIs / SMEs for their research and development. The CloudPower project results from the XtremWeb research and development project. Its goal is to offer a low cost Cloud HPC service for small and medium-sized innovative companies. With CloudPower, companies and scientists will run their simulations to design and develop new products on a powerful, scalable, economical, reliable and secure infrastructure.

The project will lead the creation of a new and innovative company operating the platform implemented in the framework of the ANR Emergence. CloudPower will implement SaaS / PaaS portal for customers and develop extensions to allow commercial exploitation of unused resources. Building on the network of SMIs from the competitiveness clusters System@tic and LyonBiopole, we will implement scenarios and/or demonstrators which illustrate the ability of CloudPower to increase competitiveness, research and marketing of innovative SMEs.

8.1.2.2. *ANR INFRA MOEBUS, Multi-objective scheduling for large computing platforms, 4 years, ANR-13-INFR-000, 2013-2016*
Participants: Christian Perez, Laurent Lefèvre, Frédéric Suter.

The ever growing evolution of computing platforms leads to a highly diversified and dynamic landscape. The most significant classes of parallel and distributed systems are supercomputers, grids, clouds and large hierarchical multi-core machines. They are all characterized by an increasing complexity for managing the jobs and the resources. Such complexity stems from the various hardware characteristics and from the applications characteristics. The MOEBUS project focuses on the efficient execution of parallel applications submitted by various users and sharing resources in large-scale high-performance computing environments.

We propose to investigate new functionalities to add at low cost in actual large scale schedulers and programming standards, for a better use of the resources according to various objectives and criteria. We propose to revisit the principles of existing schedulers after studying the main factors impacted by job submissions. Then, we will propose novel efficient algorithms for optimizing the schedule for unconventional objectives like energy consumption and to design provable approximation multi-objective optimization algorithms for some relevant combinations of objectives. An important characteristic of the project is its right balance between theoretical analysis and practical implementation. The most promising ideas will lead to integration in reference systems such as SLURM and OAR as well as new features in programming standards implementations such as MPI or OpenMP.

8.1.2.3. ANR ARPEGE MapReduce, Scalable data management for Map-Reduce-based data-intensive applications on cloud and hybrid infrastructures, 4 years, ANR-09-JCJC-0056-01, 2010-2014
Participants: Frédéric Desprez, Gilles Fedak, Sylvain Gault, Christian Perez, Anthony Simonet.

MapReduce is a parallel programming paradigm successfully used by large Internet service providers to perform computations on massive amounts of data. After being strongly promoted by Google, it has also been implemented by the open source community through the Hadoop project, maintained by the Apache Foundation and supported by Yahoo! and even by Google itself. This model is currently getting more and more popular as a solution for rapid implementation of distributed data-intensive applications. The key strength of the MapReduce model is its inherently high degree of potential parallelism.

In this project, the AVALON team participates to several work packages which address key issues such as efficient scheduling of several MapReduce applications, integration using components on large infrastructures, security and dependability, and MapReduce for Desktop Grid.

8.1.2.4. ANR INFRA SONGS, Simulation Of Next Generation Systems, 4 years, ANR-12-INFRA-11, 2012-2015
Participants: Frédéric Desprez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

The last decade has brought tremendous changes to the characteristics of large scale distributed computing platforms. Large grids processing terabytes of information a day and the peer-to-peer technology have become common even though understanding how to efficiently such platforms still raises many challenges. As demonstrated by the USS SIMGRID project, simulation has proved to be a very effective approach for studying such platforms. Although even more challenging, we think the issues raised by petaflop/exaflop computers and emerging cloud infrastructures can be addressed using similar simulation methodology.

The goal of the SONGS project is to extend the applicability of the SIMGRID simulation framework from Grids and Peer-to-Peer systems to Clouds and High Performance Computation systems. Each type of large-scale computing system will be addressed through a set of use cases and lead by researchers recognized as experts in this area.

Any sound study of such systems through simulations relies on the following pillars of simulation methodology: Efficient simulation kernel; Sound and validated models; Simulation analysis tools; Campaign simulation management.

8.1.3. Inria Large Scale Initiative

8.1.3.1. HEMERA, 4 years, 2010-2014
Participants: Christian Perez, Laurent Pouilloux, Laurent Lefèvre.

Hemera deals with the scientific animation of the GRID'5000 community. It aims at making progress in the understanding and management of large scale infrastructure by leveraging competences distributed in various French teams. Hemera contains several scientific challenges and working groups. The project involves around 24 teams located in all around France.

C. Pérez is leading the project; L. Lefevre and L. Pouilloux are managing scientific challenges on GRID'5000.

8.1.3.2. C2S@Exa, Computer and Computational Sciences at Exascale, 4 years, 2013-2017
Participants: Frédéric Desprez, Christian Perez, Laurent Lefèvre, Jérôme Richard.

Since January 2013, the team is participating to the C2S@Exa Inria Project Lab (IPL). This national initiative aims at the development of numerical modeling methodologies that fully exploit the processing capabilities of modern massively parallel architectures in the context of a number of selected applications related to important scientific and technological challenges for the quality and the security of life in our society. At the current state of the art in technologies and methodologies, a multidisciplinary approach is required to overcome the challenges raised by the development of highly scalable numerical simulation software that can exploit computing platforms offering several hundreds of thousands of cores. Hence, the main objective of C2S@Exa is the establishment of a continuum of expertise in the computer science and numerical mathematics domains, by gathering researchers from Inria project-teams whose research and development activities are tightly linked

to high performance computing issues in these domains. More precisely, this collaborative effort involves computer scientists that are experts of programming models, environments and tools for harnessing massively parallel systems, algorithmists that propose algorithms and contribute to generic libraries and core solvers in order to take benefit from all the parallelism levels with the main goal of optimal scaling on very large numbers of computing entities and, numerical mathematicians that are studying numerical schemes and scalable solvers for systems of partial differential equations in view of the simulation of very large-scale problems.

8.1.4. Inria ADT

8.1.4.1. Inria ADT Aladdin, 4 years, 2008-2014
Participants: Simon Delamare, Frédéric Desprez, Matthieu Imbert, Laurent Lefèvre, Christian Perez.

ADT ALADDIN is an Inria support action of technological development which supports the GRID'5000 instrument. Frédéric Desprez is leading this action (with David Margery from Rennes as the Technical Director). More information at Section 5.8.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. PRACE 2IP
Participants: Vincent Lanore, Christian Perez, Jérôme Richard.

Title: PRACE – Second Implementation Phase Project

Type: Integrated Infrastructure Initiative Project (I3)

Instrument: Combination of Collaborative projects and Coordination and support action

Duration: September 2011 - August 2014

Coordinator: Thomas Lippert (Germany)

Others partners: Jülich GmbH, GCS, GENCI, EPSRC, BSC, CSC, ETHZ, NCF, JKU, Vetenskapsradet, CINECA, PSNC, SIGMA, GRNET, UC-LCA, NUI Galway, UYBHM, CaSToRC, NCSA, Technical Univ. of Ostrava, IPB, NIIF

See also: http://prace-ri.eu

Abstract: The purpose of the PRACE RI is to provide a sustainable high-quality infrastructure for Europe that can meet the most demanding needs of European HPC user communities through the provision of user access to the most powerful HPC systems available worldwide at any given time. In tandem with access to Tier-0 systems, the PRACE-2IP project will foster the coordination between national HPC resources (Tier-1 systems) to best meet the needs of the European HPC user community. To ensure that European scientific and engineering communities have access to leading edge supercomputers in the future, the PRACE-2IP project evaluates novel architectures, technologies, systems, and software. Optimizing and scaling of application for Tier-0 and Tier-1 systems is a core service of PRACE.

Inria participates to Work Package 12 which is about novel programming techniques.

8.2.1.2. PaaSage
Participants: Christian Perez, Jonathan Rouzaud-Cornabas.

Title: PaaSage: Model-based Cloud Platform Upperware

Type: Seventh Framework Programme

Instrument: Collaborative project

Duration: October 2012 - September 2016 (48 months)

Coordinator: Pierre Guisset (GEIE ERCIM)

Others partners: SINTEF, STFC, HLRS, University of Stuttgart, Inria, CETIC, FORTH, be.wan, EVRY, SysFera, Flexiant, Lufthansa Systems, AG GWDG, Automotive Simulation Center Stuttgart e.V.

See also: http://paasage.eu

Abstract: PaaSage will deliver an open and integrated platform, to support both deployment and design of Cloud applications, together with an accompanying methodology that allows model-based development, configuration, optimization, and deployment of existing and new applications independently of the existing underlying Cloud infrastructures. Specifically it will deliver an IDE (Integrated Development Environment) incorporating modules for design time and execution time optimizations of applications specified in the CLOUD Modeling Language (CLOUD ML), execution-level mappers and interfaces and a metadata database.

8.2.2. Collaborations in European Programs, except FP7 & H2020

8.2.2.1. CHIST-ERA STAR

Participants: Laurent Lefèvre, Olivier Glück.

Title: SwiTching And tRansmission project

Type: CHIST-ERA (European Coordinated Research on Long-term Challenges in Information and Communication Sciences & Technologies ERA-Net)

Duration: 2013-2015

Coordinator: Jaafar Elmirghani (University of Leeds - UK)

Others partners: Inria ,University of Cambridge (UK), University of Leeds (UK), AGH University of Science and Technology Department of Telecommunications (Poland)

See also: http://www.chistera.eu/projects/star

Abstract: The Internet power consumption has continued to increase over the last decade as a result of a bandwidth growth of at least 50 to 100 times. Further bandwidth growth between 40% and 300% is predicted in the next 3 years as a result of the growing popularity of bandwidth intensive applications. Energy efficiency is therefore increasingly becoming a key priority for ICT organizations given the obvious ecological and economic drivers. In this project we adopt the GreenTouch energy saving target of a factor of a 100 for Core Switching and Routing and believe this ambitious target is achievable should the research in this proposal prove successful. A key observation in core networks is that most of the power is consumed in the IP layer while optical transmission and optical switching are power efficient in comparison, hence the inspiration for this project. Initial studies by the applicants show that physical topology choices in networks have the potential to significantly reduce the power consumption, however network optimization and the consideration of traffic and the opportunities afforded by large, low power photonic switch architectures will lead to further power savings. Networks are typically over provisioned at present to maintain quality of service. We will study optimum resource allocation to reduce the overprovisioning factor while maintaining the quality of service. Protection is currently provided in networks through the allocation of redundant paths and resources, and for full protection there is a protection route for every working route. Avalon is contributing to STAR in terms of software network protocols and services optimizations which will be combined with more efficient photonic switches in order to obtain a factor of 100 power saving in core networks can be realised through this project with significant potential for resulting impact on how core photonic networks are designed and implemented.

8.2.2.2. COST IC1305 : Nesus

Participants: Laurent Lefèvre, Marcos Dias de Assunçao.

Program: COST

Project acronym: IC1305

Project title: Network for Sustainable Ultrascale Computing (NESUS)

Duration: 2014-2019

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 Lefevre is co-chairing the Working on Energy Efficiency (WG5).

8.2.2.3. SEED4C

Program: Celtic-Plus

Project acronym: SEED4C

Project title: Security Embedded Element and Data privacy for the Cloud.

Duration: 2012-2015

Coordinator: Stéphane Betge-Brezetz (Alcatel-Lucent lab)

Other partners: Gemalto, ENSI Bourges, Inria, Wallix, VTT Technical Research centre of Finland, Mikkelin Puhelin Oyj, Cygate, Nokia Siemens Networks, Finceptum OY (Novell), Solacia, Innovalia Association, Nextel, Software Quality Systems, Ikusi, Vicomtech, Biscaytik

Abstract: SEED4C is a Celtic-Plus project: an industry-driven European research initiative to define, perform and finance through public and private funding common research projects in the area of telecommunications, new media, future Internet, and applications and services focusing on a new "Smart Connected World" paradigm. Celtic-Plus is a EUREKA ICT cluster and is part of the inter-governmental EUREKA network.

The cloud security challenge not only reflects on the secure running of software on one single machine, but rather on managing and guaranteeing security of a computer group or cluster seen as a single entity. Seed4C focus is to evolve from cloud security with an isolated point or centralized points of enforcement for security to cloud security with cooperative points of enforcement for security.

8.3. International Initiatives

8.3.1. Inria International Labs

8.3.1.1. Inria-UIUC-NCSA Joint Laboratory for Petascale Computing
Participants: Eddy Caron, Frédéric Desprez, Olivier Glück, Vincent Lanore, Laurent Lefèvre, Christian Perez, Jonathan Rouzaud-Cornabas.

The Joint Laboratory for Petascale Computing focuses on software challenges found in complex high-performance computers. The Joint Laboratory is based at the University of Illinois at Urbana-Champaign and includes researchers from the French national computer science institute called Inria, Illinois' Center for Extreme-Scale Computation, and the National Center for Supercomputing Applications. Much of the Joint Laboratory's work will focus on algorithms and software that will run on Blue Waters and other petascale computers.

8.3.2. Participation In other International Programs

8.3.2.1. HPC visibility and strategy Workshop Algeria- Inria – Bull

Under high patronage of his Excellency the Minister for Higher Education and Scientific Research in Association of the Head Office of Scientific Research and Technological Development, this meeting comes within the framework of the partnership between Algeria, Inria and the Bull company, to set up an ambitious program, based on a great show of a material and software infrastructure for the digital simulation which will allow major steps forward in various scientific fields and important progress in term of industrial competitiveness and innovation.

Facing the growing evolution of the complexity of the feigned systems and the used volumes of data, the supercomputing becomes so major. This meeting which will gather about 150 participants, of whom persons in charge of supercomputing Algerian centers, will allow to exchange on problems related to modeling, simulation and HPC. The meeting will be organized around three main points: education research, bridge industry - research and ecosystem.

Eddy Caron (Avalon team) is an expert in the steering committee of the Ecosystem group.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

> Tchimou N'Takpé, Assistant Professor
> > Date: Oct 2014 - Nov 2014
> > Institution: Université Nangui Abrogoua, Abidjan (Cote d'Ivoire)
> Mircea Moca, Assistant Professor
> > Date: Nov 15th, 2014 - Dec 15th, 2014
> > Institution: Babes-Bolyai University (Roumania)
> Mircea Moca, Assistant Professor
> > Date: Nov 15th, 2014 - Dec 15th, 2014
> > Institution: Babes-Bolyai University (Roumania)
> Asma Ben Cheikh Ahmed, PhD Student
> > Date: Sep 15th, 2014 - Dec 15th, 2014
> > Institution: Faculté des Sciences de Tunis (Tunisia)
> Miranda Qian Zhang, PhD Student
> > Date: Sep 8th, 2014 - Oct 9th, 2014
> > Institution: Australian National University (Australia)
> Julio Anjos, PhD student
> > Date: May 4th, 2014 - May 4th, 2015
> > Institution: Universidade Federal do Rio Grande do Sul (Brazil)

8.4.1.1. Internships

> Anshul Gupta
> > Date: May 2014 - Jul 2014
> > Institution: LNM Institute of Information Technology (India)

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific events organisation

9.1.1.1. general chair, scientific chair

Laurent Lefevre and Frédéric Desprez were general chairs of the Grid 5000 Spring School 2014, Lyon, June 17-19, 2014.

F. Desprez: BigDataCloud (workshop within EuroPAR)

L. Lefevre :

- Co-Workshop organizer of TASUS 2014: Techniques and Applications for Sustainable Ultrascale Computing Systems, held during EUROPAR 2014, Porto, Portugal, August 25-26, 2014
- Co-organizer of the Green Days @ Rennes event: "Avalanche of data, virtual machines and connected objects : energy efficiency is mandatory - Impact on architectures, systems and networks", Rennes, July 1-2, 2014
- Co-Workshop chair of ExtremeGreen 2014 : Extreme Green & Energy Efficiency in Large Scale Distributed Systems , Chicago, IL, USA, May 2014

9.1.1.2. member of the organizing committee

G. Fedak: Publicity Chair for SBAC-PAD'14

Eddy Caron and Laurent Lefevre: co-organizer for the Inria booth at SC'14.

9.1.2. Scientific events selection

9.1.2.1. member of the conference program committee

E. Caron: HCW 2014, IPDPS 2014, CCA-2014

F. Desprez: Closer, IPDPS, CCGRID, VECPAR, Supercomputing, ComPAS, HotCloud, BigData, Cluster.

G. Fedak: AINA, EuroPar, CloudCom, HPDC

JP. Gelas: ExtremeGreen.

O. Glück: ExtremeGreen, PDP.

L. Lefevre : IEEE CCGrid15, AusPDC15, CLOUDNET14, EuroMPI/ASIA 2014, IEEE HPCC14, Greencom14, ICA3PP14, ENA-HPC14, Globecom14 SAC Green Communication Systems and Networks. ICCCN14 (DCBC), ICT4S14

C. Perez: Cloud Control Workshop, CCGRID'15, Cluster, EuroPar, VECPAR.

F. Suter: EuroPar, EuroMPI/Asia, VECPAR, VDTC.

9.1.2.2. reviewer

E. Caron: Hotcloud'14

JP. Gelas: ExtremeGreen.

O. Glück: EuroPar, Cluster.

F. Suter: CCGrid, Cluster.

9.1.3. Journal

9.1.3.1. member of the editorial board

- F. Desprez: Scalable Computing: Practice and Experience
- G. Fedak: Journal of Cluster Computing
- L. Lefevre : Associate editor of IEEE Transactions on Cloud Computing (TCC)

9.1.3.2. reviewer

E. Caron: JPDC, TPDS, Cluster Computing, Engineering Applications of Artificial Intelligence

F. Desprez: Parallel Computing

G. Fedak: Computer

C. Perez: Parallel Computing.

F. Suter: Parallel Computing, CCPE, JPDC, FGCS, GPAA.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence: Olivier Glück, Initiation Réseaux, 2x9h, niveau L2, Université Lyon 1, France

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

Master: Olivier Glück, Services et Protocoles Applicatifs sur Internet, 40h, niveau M2, Université Lyon 1, France

Master: Olivier Glück, Services et Protocoles Applicatifs sur Internet, 24h, niveau M2, IGA Casablanca, Maroc

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

Master: Jean-Patrick Gelas, Analyse de performance, 6h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Réseaux, 48h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Système d'exploitation, 30h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Architecture des routeurs, 6h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Systèmes embarqués (Linux, Android), 85h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Routage et IPv6, 45h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Routage et IPv6, 45h, niveau M2, Université Lyon 1, France

Master: Frédéric Desprez, Parallélisme, 30h, niveau M1, Université Lyon 1, France

Master: Eddy Caron, Gilles Fedak, Laurent Lefèvre, and Christian Perez, Grid and Cloud Computing, 36h, niveau M2, Ecole Normale Supérieure de Lyon, France

PhD.: G. Fedak, MapReduce Programming Model, Formation doctorale, 7h, Université Paris XIII, Villetaneuse, France

Master: G. Fedak, MapReduce and Big Data, Master, 9h, University of Babes-Bolay, Romania

9.2.2. Supervision

HdR : Frédéric Suter, Bridgin a gap between research and production, contributions to scheduling and simulation, École Normale Supérieure de Lyon, December 10, 2014.

PhD in progress: Semen Marchuk, *Contribution pour la conception d'algorithmes d'ordonnancement de taches de calcul dans le domaine de l'animation 3D*, 10/2014, Eddy Caron (dir), Pierre Biecher (CapRézo society, co-dir) and Rodolphe De Carini (CapRézo society, co-dir)

PhD in progress: Radu Carpa, *Efficacité énergétique des échanges de données dans une fédération d'infrastructures distribuées à grande échelle*, 10/2014, Laurent Lefèvre (dir), Olivier Glück (co-dir).

PhD in progress: Sylvain Gault, *Improving MapReduce Performance on Clusters*, 11/2010, Frédéric Desprez (dir), Christian Pérez (co-dir).

PhD in progess: Maurice Djibril Faye, *Déploiement auto-adaptatif d'intergiciel sur plateforme élastique*, Eddy Caron (dir), Ousmane Thiaré (Université Gaston Berger, St Louis, Sénégal, co-dir)

PhD in progress, Vincent Lanore, *Adaptation et dynamicité dans les modèles à composants logiciels pour les applications scientifiques*, C. Pérez.

PhD in progress, Arnaud Lefray, *Mission fonctionnelle et de sécurité dans une informatique en nuage*, Eddy Caron (dir), Christian Toinard (ENSIB, co-dir)

PhD in progress : Anthony Simonet, *Exécution efficace d'applications de traitement intensif de données sur des infrastructures distribuées hybrides*, 10/2011, Gilles Fedak.

PhD in progress: Jonathan Pastor (Nantes), *Conception d'un système de gestion autonomique, coopérative et distribué pour les centrales numériques à large échelle*, 11/2012, Frédéric Desprez, Adrien Lèbre (Ascola, Nantes, co-dir).

PhD in progress: Pedro Silva, *Application model and co-scheduling algorithm for dynamic and evolutive data-intensive application*, 10/2014, Christian Perez (dir), Frédéric Desprez (co-dir).

PhD in progress: Violaine Villebonnet, *Proportionnalité énergétique dans les systèmes distribués à grande échelle*, 9/2013, Laurent Lefevre (dir), Jean-Marc Pierson (IRIT, Toulouse, co-dir)

PhD in progress: Daniel Balouek Thomert, *Ordonnancement et éco-efficacité dans les environnements virtualisés*, Eddy Caron (dir), Laurent Lefevre (co-dir), Gilles Cieza (NewGen society, co-dir)

PhD in progress: Jérôme Richard, *Conception of a software component model with task scheduling for many-core based parallel architecture, application to the Gysela5D code*, 11/2014, Christian Perez (dir), Julien Bigot (CEA, co-dir).

9.2.3. Juries

Eddy Caron has been member of the following PhD Juries:
- Gautier Berthou. "Vers des communications anonymes et efficaces", Grenoble, January 2014, examinateur.
- Mandicou Ba. " Vers une structuration auto-stabilisante des réseaux ad hoc: cas des réseaux de capteurs sans fil.", Université de Reims, May 2014, examinateur.
- Vianney Kengne Tchendji. "Solutions parallèles efficaces sur le modèle CGM d'une classe de problèmes issus de la programmation dynamique", Université de Picardie Jules Verne en cotutelle avec l'Université de Yaoundé, June 2014, exminateur.

Frédéric Desprez has been member of the following PhD Juries:
- Amit Sangroya, "Towards Dependability and Performance Benchmarking for Cloud Computing Services", Grenoble, April 2014, rapporteur.
- Shubhabrata Roy, "A Complete Framework for Modelling Workload Volatility of VoD System-a Perspective to Probabilistic Management", Lyon, June 2014, examinateur.
- Loic Letondeur, "Planification pour la gestion autonomique de l'élasticité d'applications dans le cloud", Lyon, October 2014, examinateur.
- Rachid Habel, "", Mines Paris, October 2014, président.
- Radu Tudoran, "High Performance Big Data Management Across Cloud Data Centers", Rennes, December 2014, rapporteur.
- Andra Hugo, "Composability of parallel codes on heterogenous architectures", Bordeaux, December14, rapporteur.
- Damien Genet, "Design of Generic Modular solution for PDE solvers for modern architectures", Bordeaux, December 2014, examinateur.
- Lionel Boillot, "Contributions à la modélisation mathématique et à l'algorithmique parallèle pour l'optimisation d'un propagateur d'ondes élastiques en milieu anisotrope", Pau, December 2014, rapporteur.

Laurent Lefevre has been member of the following PhD Juries:
- Truong Khoa Phan: "Design and Management of Networks with Low Power Consumption", University of Nice, France, September 2014, Rapporteur.
- Frédéric Pinel: "Energy - Performance optimization for the Cloud", University of Luxembourg, July 2014, Examinateur
- Thiam Chekhou : "Anti Load Balancing for Energy-Aware Distributed Scheduling of Virtual Machines", University of Toulouse, France, July 2014, Rapporteur
- Felipe Diaz: "Cloud brokering : new value added services and pricing models", Telecom ParisTech, France, February 2014, Rapporteur

Christian Perez has been member of the following PhD Jury:

- Damien Gros: "Protection obligatoire répartie Usage pour le calcul intensif et les postes de travail", University of Orléans, France, June 2014, Rapporteur.
- Hélène Coullon: "Modélisation et implémentation de parallélisme implicite pour les simulations scientifiques basées sur des maillages", University of Orléans, France, September 2014.
- Clément Quinton: "Cloud Environments Selection and Configuration: A Software Product Lines-Based Approach", University of Lille, France, October 2014.

Frédéric Suter has been member of the following PhD Jury:

- Vinicius Pinheiro: "The management of multiple submissions in parallel systems: the fair scheduling approach", University of Sao Paulo, Brasil, February 2014, rapporteur.

9.3. Popularization

Eddy Caron has been interview by New Screens Magazine (N°9) edited by Imaginove. "Pipeline de production pour l'animation: Optimisation des temps de calcul et paralleˊlisme", June 2014.

Gilles Fedak is a member of the Maison des Mathématiques et de l'Informatique steering committee.

Laurent Lefevre has been interviewed by Atlantico Web News Magazine on "Gadget ou véritable promesse ? Comment Amazon compte se chauffer grâce à l'énergie dégagée par les datas centers", September 21, 2014

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] G. MARKOMANOLIS. *Performance Evaluation and Prediction of Parallel Applications*, Ecole normale supérieure de lyon - ENS LYON, January 2014, https://tel.archives-ouvertes.fr/tel-00951125

Articles in International Peer-Reviewed Journals

[2] J. BIGOT, Z. HOU, C. PÉREZ, V. PICHON. *A low level component model easing performance portability of HPC applications*, in "Computing", December 2014, vol. 96, n°12, pp. 1115-1130 [*DOI :* 10.1007/s00607-013-0368-3], https://hal.inria.fr/hal-00911231

[3] H. CASANOVA, F. DESPREZ, G. S. MARKOMANOLIS, F. SUTER. *Simulation of MPI applications with time-independent traces*, in "Concurrency and Computation: Practice and Experience", 2014 [*DOI :* 10.1002/CPE.3278], https://hal.inria.fr/hal-01064561

[4] H. CASANOVA, A. GIERSCH, A. LEGRAND, M. QUINSON, F. SUTER. *Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*, in "Journal of Parallel and Distributed Computing", June 2014, vol. 74, n°10, pp. 2899-2917 [*DOI :* 10.1016/J.JPDC.2014.06.008], https://hal.inria.fr/hal-01017319

[5] L. B. COSTA, H. YANG, E. VAIRAVANATHAN, A. BARROS, K. MAHESHWARI, D. S. KATZ, M. WILDE, M. RIPEANU, S. AL-KISWANY, G. FEDAK. *The Case for Workflow-Aware Storage:An Opportunity Study*, in "Journal of Grid Computing", June 2014, 19 p. [*DOI :* 10.1007/S10723-014-9307-6], https://hal.inria.fr/hal-01108605

[6] M. E. M. DIOURI, M. F. DOLZ, O. GLÜCK, L. LEFÈVRE, P. ALONSO, S. CATALÁN, R. MAYO, E. S. QUINTANA-ORTÍ. *Assessing Power Monitoring Approaches for Energy and Power Analysis of Computers*, in "Sustainable Computing: Informatics and Systems", June 2014, vol. 4, n⁰ 2, pp. 68-82, https://hal.inria.fr/hal-01065998

[7] R. FERREIRA DA SILVA, T. GLATARD, F. DESPREZ. *Controlling fairness and task granularity in distributed, online, non-clairvoyant workflow executions*, in "Journal of Concurrency and Computation: Practice and Experience", 2014, vol. 26, n⁰ 14, pp. 2347-2366 [*DOI :* 10.1002/CPE.3303], https://hal.archives-ouvertes.fr/hal-01016491

[8] A.-C. ORGERIE, M. DIAS DE ASSUNCAO, L. LEFÈVRE. *A Survey on Techniques for Improving the Energy Efficiency of Large Scale Distributed Systems*, in "ACM Computing Surveys", December 2014, vol. 46, n⁰ 4, 31 p. , https://hal.inria.fr/hal-00767582

International Conferences with Proceedings

[9] A. BEN CHEIKH, H. ABBES, G. FEDAK. *Towards Privacy for MapReduce on Hybrid Clouds Using Information Dispersal Algorithm*, in "Data Management in Cloud, Grid and P2P Systems (GLOBE)", Munich, Germany, Lecture Notes in Computer Science, September 2014, vol. 8648 [*DOI :* 10.1007/978-3-319-10067-8_4], https://hal.inria.fr/hal-01108651

[10] L. BOBELIN, A. BOUSQUET, J. BRIFFAUT, E. CARON, J.-F. COUTURIER, A. LEFRAY, J. ROUZAUD-CORNABAS, C. TOINARD. *An Advanced Security-Aware Cloud Architecture*, in "HPCS 2014 - The 2014 International Conference on High Performance Computing & Simulation", Bologne, Italy, IEEE, July 2014, https://hal.inria.fr/hal-01063852

[11] R. CARPA, O. GLÜCK, L. LEFEVRE. *Segment Routing based Traffic Engineering for Energy Efficient Backbone Networks*, in "ANTS2014 : IEEE International Conference on Advanced Networks and Telecommunications Systems", New Delhi, India, December 2014, https://hal.inria.fr/hal-01094359

[12] J. PASTOR, M. BERTIER, F. DESPREZ, A. LÈBRE, F. QUESNEL, C. TEDESCHI. *Locality-aware Cooperation for VM Scheduling in Distributed Clouds*, in "Euro-Par 2014", Porto, Portugal, August 2014, https://hal.inria.fr/hal-00991530

[13] F. ROSSIGNEUX, L. LEFEVRE, J.-P. GELAS, M. DIAS DE ASSUNCAO. *A Generic and Extensible Framework for Monitoring Energy Consumption of OpenStack Clouds*, in "The 4th IEEE International Conference on Sustainable Computing and Communications (Sustaincom 2014)", Sydney, Australia, December 2014, https://hal.inria.fr/hal-01094387

[14] B. TANG, H. HE, G. FEDAK. *Parallel Data Processing in Dynamic Hybrid Computing Environment Using MapReduce*, in "International Conference on Algorithms and Architectures for Parallel Processing", Dalian, China, August 2014 [*DOI :* 10.1007/978-3-319-11194-0_1], https://hal.inria.fr/hal-01108619

[15] G. L. TSAFACK CHETSA, L. LEFEVRE, J.-M. PIERSON, P. STOLF, G. DA COSTA. *Application-Agnostic Framework for Improving the Energy Efficiency of Multiple HPC Subsystems*, in "PDP2015 : 23rd Euromicro International Conference on Parallel, Distributed and Network-based Processing", Turku, Finland, March 2015, https://hal.inria.fr/hal-01094431

[16] V. VILLEBONNET, G. DA COSTA, L. LEFEVRE, J.-M. PIERSON, P. STOLF. *Towards Generalizing "Big.Little" for Energy Proportional HPC and Cloud Infrastructures*, in "IEEE International Conference on Sustainable Computing and Communications (SustainCom 2014)", Sydney, Australia, December 2014, https://hal.inria.fr/hal-01075819

National Conferences with Proceedings

[17] D. BALOUEK-THOMERT, E. CARON, L. LEFÈVRE. *Gestion adaptative de l'énergie pour les infrastructures de type grappe ou nuage*, in "Conférence d'informatique en Parallélisme, Architecture et Système (ComPAS)", Neuchâtel, Switzerland, April 2014, https://hal.inria.fr/hal-01018238

Conferences without Proceedings

[18] A. SIMONET. *Active Data : Un mode`le pour repre´senter et programmer le cycle de vie des donne´es distribue´es*, in "ComPAS'2014", Neuchâtel, Switzerland, April 2014, https://hal.inria.fr/hal-00984210

Scientific Books (or Scientific Book chapters)

[19] H. ARABNEJAD, J. BARBOSA, F. SUTER. *Fair Resource Sharing for Dynamic Scheduling of Workflows on Heterogeneous Systems*, in "High-Performance Computing on Complex Environments", E. JEANNOT, J. ZILINSKAS (editors), Parallel and Distributed Computing series, Wiley, June 2014, https://hal.inria.fr/hal-00926460

[20] E. CARON, F. DESPREZ, J. ROUZAUD-CORNABAS. *Smart Resources Allocation to improve Cloud Security*, in "Security, Privacy and Trust in Cloud Systems", S. NEPAL, M. PATHAN (editors), Springer Berlin Heidelberg, 2014, vol. Part 1, pp. 103-143 [*DOI :* 10.1007/978-3-642-38586-5_4], https://hal.inria.fr/hal-00931144

[21] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE, J.-C. MIGNOT. *Providing Green Services in HPC Data Centers: A Methodology based on Energy Estimation*, in "Handbook on Data Centers", S. U. KHAN, A. ZOMAYA (editors), Springer, 2015, https://hal.inria.fr/hal-01066002

[22] C. HERZOG, L. LEFEVRE, J.-M. PIERSON. *Actors for Innovation in Green IT*, in "ICT Innovations for Sustainability", Advances in Intelligent Systems and Computing, Springer, August 2014, vol. 310, pp. 49-67, https://hal.inria.fr/hal-01094492

[23] A. LÈBRE, J. PASTOR, M. BERTIER, F. DESPREZ, J. ROUZAUD-CORNABAS, C. TEDESCHI, A.-C. ORGERIE, F. QUESNEL, G. FEDAK. *Beyond The Clouds, How Should Next Generation Utility Computing Infrastructures Be Designed?*, in "Cloud Computing: Challenges, Limitations and R&D Solutions", Z. MAHMOOD (editor), Springer, November 2014, https://hal.inria.fr/hal-01067888

[24] G. L. TSAFACK CHETSA, G. DA COSTA, L. LEFEVRE, J.-M. PIERSON, O. ARIEL, B. ROBERT. *Energy aware approach for HPC systems*, in "High-Performance Computing on Complex Environments", E. JEANNOT, J. ZILINSKAS (editors), John Wiley & Sons, June 2014, https://hal.inria.fr/hal-00925310

Research Reports

[25] A. BSILA, N. FERRY, G. HORN, T. KIRKHAM, M. MALAWSKI, N. PARLAVANTZAS, C. PÉREZ, J. ROUZAUD-CORNABAS, D. ROMERO, A. ROSSINI, A. SOLBERG, H. SONG. *D3.1.1 - Upperware Prototype*, March 2014, https://hal.inria.fr/hal-00991340

[26] Y. CANIOU, E. CARON, G. LE MAHEC, H. NAKADA. *Data Management API within the GridRPC*, June 2014, OGF Reference: GFD 186, https://hal.archives-ouvertes.fr/hal-01009375

[27] E. CARON, J. ROUZAUD-CORNABAS. *Improving users' isolation in IaaS: Virtual Machine Placement with Security Constraints*, January 2014, n⁰ RR-8444, https://hal.inria.fr/hal-00924296

[28] F. DESPREZ, J.-L. LUCAS-SIMARRO, R. MORENO VOZMEDIANO, J. ROUZAUD-CORNABAS. *Image Transfer and Storage Cost Aware Brokering Strat. for Multiple Clouds*, January 2014, n⁰ RR-8445, Cloud Brokering, Resource Allocation, Storage, Data Transfer, SimGrid Cloud Broker, https://hal.inria.fr/hal-00924351

[29] L. POUILLOUX, J. ROUZAUD-CORNABAS. *Performance analysis and models for collocated VMs running on multi-core physical machines*, February 2014, n⁰ RR-8473, 30 p. , https://hal.inria.fr/hal-00945881

[30] J. C. SANTOS DOS ANJOS, G. FEDAK, C. GEYER. *BIGhybrid - A Toolkit for Simulating MapReduce on Hybrid Infrastructures*, September 2014, https://hal.archives-ouvertes.fr/hal-01062709

Other Publications

[31] J. RICHARD. *Implementation and Evaluation of 3D FFT Parallel Algorithms Based on Software Component Model*, LIP - ENS Lyon ; Orleans-Tours, September 2014, 50 p. , https://hal.inria.fr/hal-01082575

References in notes

[32] K. BERGMAN, S. BORKAR, D. CAMPBELL, W. CARLSON, W. DALLY, M. DENNEAU, P. FRANZON, W. HARROD, J. HILLER, S. KARP, S. KECKLER, D. KLEIN, R. LUCAS, M. RICHARDS, A. SCARPELLI, S. SCOTT, A. SNAVELY, T. STERLING, R. S. WILLIAMS, K. YELICK, K. BERGMAN, S. BORKAR, D. CAMPBELL, W. CARLSON, W. DALLY, M. DENNEAU, P. FRANZON, W. HARROD, J. HILLER, S. KECKLER, D. KLEIN, P. KOGGE, R. S. WILLIAMS, K. YELICK. *ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems*, in "DARPA Information Processing Techniques Office", Washington, DC, September 28 2008, 278 p.

[33] H. COULLON. *Implicit parallelism modeling and implementation for mesh-based scientific simulations*, Université d'Orléans, September 2014, https://hal.archives-ouvertes.fr/tel-01094327

[34] 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, n⁰ 5, pp. 658–671 [*DOI : 10.1109/TPDS.2009.76*], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4906989

[35] A. GEIST, S. DOSANJH. *IESP Exascale Challenge: Co-Design of Architectures and Algorithms*, in "Int. J. High Perform. Comput. Appl.", November 2009, vol. 23, n⁰ 4, pp. 401–402, http://dx.doi.org/10.1177/1094342009347766

[36] 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

[37] 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

[38] G. MADEC. *NEMO ocean engine*, Institut Pierre-Simon Laplace (IPSL)France, 2008, n⁰ 27, ISSN No 1288-1619

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

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

[41] 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

[42] 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

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

[44] 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