IN PARTNERSHIP WITH:
CNRS

Université Claude Bernard (Lyon 1)

Ecole normale supérieure de Lyon

Activity Report 2013

Team AVALON

Algorithms and Software Architectures for Distributed and HPC Platforms

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

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Team AVALON

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

Creation of the Team: 2012 February 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]

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]
Nicolas Bard [CNRS, granted by E-Biothon project since February 2013]
Sylvain Bernard [ENS Lyon, granted by ANR CloudPower]
Haiwu He [ENS Lyon, granted by ANR CloudPower]
Julien Carpentier [Inria, granted by XLcloud project, until Apr 2013]
François Rossigneux [Inria, granted by XLcloud project]
Noua Toukourou [Inria, granted by PRACE 2IP]
Amine Bsila [SysFera, provisional assignment since June 2013]
Guillaume Verger [Inria, granted by SEED4C project]
Huaxi Zhang [Inria, granted by SEED4C project]
Laurent Pouilloux [Inria, supported by Hemera Large Scale Initiative]

PhD Students

Mohammed El Mehdi Diouri [ENS Lyon, until September 2013]
Maurice-Djibril Faye [Joint PhD with the University Gaston Berger, St Louis, Sénégal; UGB grant, CMIRA grant, and LIP grant]
Sylvain Gault [Inria]
Vincent Lanore [ENS Lyon]
Arnaud Lefray [ENSI de Bourges]
Georgios Markomanolis [Inria, until July 2013]
Anthony Simonet [Inria]
Ghislain Landry Tsafack Chetsa [Univ. Lyon I]
Violaine Villebonnet [Inria, since October 2013]
Daniel Balouek [NewGeneration SR]

Post-Doctoral Fellows

Zhengxiong Hou [Inria, granted by PRACE 2IP project, until February 2013]
Jonathan Rouzaud-Cornabas [CNRS temporary researcher until September 2013, then Inria, granted by PaaSage project]

Lamiel Toch [Inria, granted by SEED4C project]

Visiting Scientist

Jose Luis Lucas [Univ. Computense of Madrid, from May to August 2013]

Administrative Assistant

Evelyne Blesle [Inria]

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 [59].

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 [60] 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 [64], phase detection for specific HPC applications [69], 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 [63] exposes the number of nodes but hides the complexity of network topology behind a set of collective operations; OpenMP [67] simplifies the management of threads on top of a shared memory machine while OpenACC [66] 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 [61]. 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 [70] 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 [68] 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 [65] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [71] 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 (http://www.ias.u-psud.fr/imEuclid) 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. Software and Platforms

5.1. BitDew

Participants: Gilles Fedak [correspondant], Haiwu He.

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 work focuses on providing reliable storage on top of hybrid distributed computing infrastructures.

5.2. DIET

Participants: Daniel Balouek, Eddy Caron [correspondant], Frédéric Desprez, Maurice-Djibril 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. Pilgrim

Participants: Eddy Caron, Matthieu Imbert [correspondant].

Pilgrim (http://pilgrim.gforge.inria.fr) is an open metrology and prediction performance framework whose goal is to provide easy and powerful tools for instrumenting computer platforms and predicting their behavior. Those tools are aimed at being used not only by humans but also by programs, in particular by resource managers and schedulers. Pilgrim is designed to be a loosely coupled integration of various custom-developed or off-the-shelf tools.

5.4. 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 ENSIB (Bourges).

5.5. SimGrid

Participants: Georgios Markomanolis, 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/.

5.6. HLCMi, L^{2}C, & Gluon++

Participants: Zhengxiong Hou, Vincent Lanore, Christian Perez [correspondant].

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/L^{2}C—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/L^{2}C and HLCM/Charm++ (known as Gluon++).

L^2C (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^2C does not offer network transparency neither language transparency. Instead, L^2C lets the user choose between various kinds of interactions between components, some with ultra low overhead and others that support network transport. L^2C is extensible as additional interaction kinds can be added quite easily. L^2C 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^2C 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++.

5.7. Execo

Participants: Matthieu Imbert [correspondant], Laurent Pouilloux.

Execo (http://execo.gforge.inria.fr) offers a Python API for local or remote, standalone or parallel, processes execution. It is especially well suited for scripting workflows of parallel/distributed operations on local or remote hosts: automating a scientific workflow, conducting computer science experiments, performing automated tests, etc. The core python package is `Execo`. The `execo_g5k` package provides a set of tools and extensions for GRID'5000. The `execo_engine` package and `execo-run` script provide an extendable experiment engine.

Execo currently has more than 10 users in and outside the AVALON team, who rely on it to automate experimental workflows, mainly on GRID'5000 ([26]).

It is distributed under GPLv3 and it is made of 5200 lines of Python.

5.8. Grid'5000

Participants: Frédéric Desprez, Simon Delamare, Laurent Lefèvre, Christian Perez.

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 operational manager of the technical team.

6. New Results

6.1. Energy efficiency of large scale distributed systems

Participants: Ghislain Landry Tsafack Chetsa, Mohammed El Mehdi Diouri, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, François Rossigneux.

6.1.1. Analysis and Evaluation of Different External and Internal Power Monitoring Devices for a Server and a Desktop Machine

Large-scale distributed systems (e.g., datacenters, HPC systems, clouds, large-scale networks, etc.) consume and will consume enormous amounts of energy. Therefore, accurately monitoring the power and energy consumption of these systems is increasingly more unavoidable. The main novelty of this contribution [15] is the analysis and evaluation of different external and internal power monitoring devices tested using two different computing systems, a server and a desktop machine. Furthermore, we also provide experimental results for a variety of benchmarks which exercise intensively the main components (CPU, Memory, HDDs, and NICs) of the target platforms to validate the accuracy of the equipment in terms of power dispersion and energy consumption. We highlight that external wattmeters do not offer the same measures as internal wattmeters. Thanks to the high sampling rate and to the different measured lines, the internal wattmeters allow an improved visualization of some power fluctuations. However, a high sampling rate is not always necessary to understand the evolution of the power consumption during the execution of a benchmark.

6.1.2. Your Cluster is not Power Homogeneous

Future supercomputers will consume enormous amounts of energy. These very large scale systems will gather many homogeneous clusters. We analyze the power consumption of the nodes from different homogeneous clusters during different workloads. As expected, we observe that these nodes exhibit the same level of performance. However, we also show that different nodes from a homogeneous cluster may exhibit heterogeneous idle power energy consumption even if they are made of identical hardware. Hence, we propose an experimental methodology to understand such differences. We show that CPUs are responsible for such heterogeneity which can reach 20% in terms of energy consumption. So energy aware (Green) schedulers must take care of such hidden heterogeneity in order to propose efficient mapping of tasks. To consume less energy, we propose an energy-aware scheduling approach taking into account the heterogeneous idle power consumption of homogeneous nodes [20]. It shows that we are able to save energy up to 17% while exploiting the high power heterogeneity that may exist in some homogeneous clusters.

6.1.3. Energy Consumption Estimations of Fault Tolerance protocols

Energy consumption and fault tolerance are two interrelated issues to address for designing future exascale systems. Fault tolerance protocols used for checkpointing have different energy consumption depending on parameters like application features, number of processes in the execution and platform characteristics. Currently, the only way to select a protocol for a given execution is to run the application and monitor the energy consumption of different fault tolerance protocols. This is needed for any variation of the execution setting. To avoid this time and energy consuming process, we propose an energy estimation framework [16], [17], [7]. It relies on an energy calibration of the considered platform and a user description of the execution setting. We evaluate the accuracy of our estimations with real applications running on a real platform with energy consumption monitoring. Results show that our estimations are highly accurate and allow selecting the best fault tolerant protocol without pre-executing the application.

6.1.4. Energy Consumption Estimations of Data Broadcasting

Future supercomputers will gather hundreds of millions of communicating cores. The movement of data in such systems will be very energy consuming. We address the issue of energy consumption of data broadcasting in such large scale systems. To this end, in [19], [7], we propose a framework to estimate the energy consumed by different MPI broadcasting algorithms for various execution settings. Validation results show that our estimations are highly accurate and allow to select the least consuming broadcasting algorithm.

6.1.5. A Smart-Grid Based Framework for Consuming Less and Better in Extreme-Scale Infrastructures

As they will gather hundreds of million cores, future exascale supercomputers will consume enormous amounts of energy. Besides being very important, their power consumption will be dynamic and irregular. Thus, in order to consume energy efficiently, powering such systems will require a permanent negotiation between the energy supplier and one of its major customers represented by exascale platforms. We have designed SESAMES [18], [53], a smart and energy-aware service-oriented architecture manager that proposes energy-efficient services for exascale applications and provides an optimized reservation scheduling. The new features of this framework are the design of a smart grid and a multi-criteria green job scheduler. Simulation results show that with the proposed multi-criteria job scheduler, we are able to save up to 2.32 % in terms of energy consumption, 24.22 % in terms of financial cost and reduce up to 7.12 % the emissions of CO_2.

6.1.6. 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 7.3), 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.

6.1.7. 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.

6.1.8. Reservation based Usage for Energy Efficient Clouds: the Climate Architecture

The FSN XLcloud project (cf Section 7.1) 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 will be improved with standby mode features, to shut down automatically the unused hosts. The first release of Climate is planned at the end of January 2014, and we expect an incubation in the next version of OpenStack.

6.2. Simulation of Large Scale Distributed Systems

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

6.2.1. Toward Better Simulation of MPI Applications on Ethernet/TCP Networks

Simulation and modeling for performance prediction and profiling is essential for developing and maintaining HPC code that is expected to scale for next-generation exascale systems, and correctly modeling network behavior is essential for creating realistic simulations. In [11], we proposed an implementation of a flow-based hybrid network model that accounts for factors such as network topology and contention, which are commonly ignored by other approaches. We focused on large-scale, Ethernet-connected systems, as these currently compose 37.8% of the TOP500 index, and this share is expected to increase as higher-speed 10 and 100GbE become more available. The European Mont-Blanc project that studies exascale computing by developing prototype systems with low-power embedded devices will also use Ethernet-based interconnect. Our model is implemented within SMPI, an open-source MPI implementation that connects real applications to the SimGrid simulation framework (cf Section 5.5). SMPI provides implementations of collective communications based on current versions of both OpenMPI and MPICH. SMPI and SimGrid also provide methods for easing the simulation of large-scale systems, including shadow execution, memory folding, and support for both online and offline simulation. We validated our proposed model by comparing traces produced by SMPI with those from real world experiments, as well as with those obtained using other established network models. Our study shows that SMPI has a consistently better predictive power than classical LogP-based models for a wide range of scenarios including both established HPC benchmarks and real applications.

6.2.2. SimGrid: a Sustained Effort for the Versatile Simulation of Large Scale Distributed Systems

SimGrid (cf Section 5.5) is a toolkit for the versatile simulation of large scale distributed systems, whose development effort has been sustained for the last fifteen years. Over this time period SimGrid has evolved from a one-laboratory project in the U.S. into a scientific instrument developed by an international collaboration. The keys to making this evolution possible have been securing of funding, improving the quality of the software, and increasing the user base. We detailed in [55]how we have been able to make advances on all three fronts, on which we plan to intensify our efforts over the upcoming years.

6.2.3. Simulating Multiple Clouds from a Client Point of View: SGCB an AWS Simulator

Validating a new application over a Cloud is not an easy task and it can be costly over public Clouds. Simulation is a good solution if the simulator is accurate enough and if it provides all the features of the target Cloud. In [49], we have proposed an extension of the SimGrid simulation toolkit to simulate the Amazon IaaS Cloud. Based on an extensive study of the Amazon platform and previous evaluations, we have integrated models into the SimGrid Cloud Broker and exposed the same API as Amazon to the users. Our experimental results have shown that our simulator is able to simulate different parts of Amazon for different applications.

6.3. Active Data: A Data-Centric Approach to Data Life-Cycle Management

Participants: Gilles Fedak, Anthony Simonet.

Data-intensive science offers new opportunities for innovation and discoveries, provided that large datasets can be handled efficiently. Data management for data-intensive science applications is challenging; requiring support for complex data life cycles, coordination across multiple sites, fault tolerance, and scalability to support tens of sites and petabytes of data. In [28], we argue that data management for data-intensive science applications requires a fundamentally different management approach than the current ad-hoc task centric approach. We propose Active Data, a fundamentally novel paradigm for data life cycle management. Active Data follows two principles: data-centric and event-driven. We report on the Active Data programming model and its preliminary implementation, and discuss the benefits and limitations of the approach on recognized challenging data-intensive science use-cases.

6.4. HPC Component Model

Participants: Zhengxiong Hou, Vincent Lanore, Christian Perez.

6.4.1. Auto-tuning of Stencil Based Applications

We have finished designing a tuning approach for stencil applications on multi-core clusters [25]. We focused in particular on a 2D Jacobi benchmark application as well as memory bandwidth performance. The tuning approach includes data partitioning within one node, the selection of the number of threads within a multi-core node, a data partitioning for multi nodes, and the number of nodes for a multi-core cluster. This model is based on a set of experiments on machines of GRID'5000 and on the Curie supercomputer.

6.4.2. Static 2D FFT Adaptation through a Component Model based on Charm++

Adaptation algorithms for HPC applications can improve performance but their implementation is often costly in terms of development and maintenance. Component models such as Gluon++, which is built on top of Charm++, propose to separate the business code, encapsulated in components, and the application structure, expressed through a component assembly. Adaptation of component-based HPC applications can be achieved through the optimization of the assembly. We have studied such an approach with the adaptation to network topology and data size of a Gluon++ 2D FFT application. Preliminary experimental results obtained on the GRID'5000 platform show the suitability of the proposed approach.

6.4.3. Towards Scalable Reconfiguration in Component Models

Some HPC applications require reconfiguration of their architecture at runtime; examples include adapting to (cloud) resource elasticity, efficient distributed deployment, Adaptive Mesh Refinement (AMR), and load balancing. This class of applications raises challenges such as handling of concurrent reconfigurations and distributed architecture representation at runtime. To our knowledge, no existing programming model addresses those challenges in the general case with both high programmability and scalability. We have identified a list of specific subproblems and use-cases and we have devised a preliminary component model to address some of them.

6.5. Resource Management and Scheduling

Participants: Eddy Caron, Frédéric Desprez, Gilles Fedak, Jose Luis Lucas, Christian Perez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.5.1. Resource Management Architecture for Fair Scheduling of Optional Computations

Most High-Performance Computing platforms require users to submit a pre-determined number of computation requests (also called jobs). Unfortunately, this is cumbersome when some of the computations are optional, i.e., they are not critical, but their completion would improve results. For example, given a deadline, the number of requests to submit for a Monte Carlo experiment is difficult to choose. The more requests are completed, the better the results are, however, submitting too many might overload the platform. Conversely, submitting too few requests may leave resources unused and misses an opportunity to improve the results.

In cooperation with IRIT (Toulouse), we have proposed a generic client-server architecture and an implementation in DIET, a production GridRPC middleware, which auto-tunes the number of requests [12]. Real-life experiments show significant improvement of several metrics, such as user satisfaction, fairness and the number of completed requests. Moreover, the solution is shown to be scalable.

6.5.2. Advanced Promethee-based Scheduler Enriched with User-Oriented Methods

Efficiently scheduling tasks in hybrid Distributed Computing Infrastructures (DCI) is a challenging pursue because the scheduler must deal with a set of parameters that simultaneously characterize the tasks and the hosts originating from different types of infrastructure. In [27], we propose a scheduling method for hybrid DCIs, based on advanced multi-criteria decision methods. The scheduling decisions are made using pairwise comparisons of the tasks for a set of criteria like expected completion time and price charged for computation. The results are obtained with an XtremWeb-like pull-based scheduler simulator using real failure traces for a combination of three types of infrastructure. We also show how such a scheduler should be configured to enhance user satisfaction regardless their profiles, while maintaining good values for makespan and cost. We validate our approach with a statistical analysis on empirical data and show that our proposed scheduling method improves performance by 12-17% compared to other scheduling methods. Experimenting on large time-series and using realistic scheduling scenarios lead us to conclude about time consistency results of the method.

6.5.3. Fair Resource Sharing for Dynamic Scheduling of Workflows on Heterogeneous Systems

Scheduling independent workflows on shared resources in a way that satisfy users Quality of Service is a significant challenge. In [37], we described methodologies for off-line scheduling, where a schedule is generated for a set of known workflows, and on-line scheduling, where users can submit workflows at any moment in time. We consider the on-line scheduling problem in more detail and present performance comparisons of state-of-the-art algorithms for a realistic model of a heterogeneous system.

6.5.4. Image Transfer and Storage Cost Aware Brokering Strategies for Multiple Clouds

Nowadays, Clouds are used for hosting a large range of services. But between different Cloud Service Providers, the pricing model and the price of individual resources are very different. Furthermore hosting a service in one Cloud is the major cause of service outage. To increase resiliency and minimize the monetary cost of running a service, it becomes mandatory to span it between different Clouds. Moreover, due to dynamicity of both the service and Clouds, it could be required to migrate a service at run time. Accordingly, this ability must be integrated into the multi-Cloud resource manager, *i.e.* the Cloud broker. But, when migrating a VM to a new Cloud Service Provider, the VM disk image must be migrated too. Accordingly, data storage and transfer must be taken into account when choosing if and where an application will be migrated.

In [47], we have extended a cost-optimization algorithm to take into account storage costs to approximate the optimal placement of a service. The data storage management consists in taking two decisions: where to upload an image, and keep it on-line during the experiment lifetime or delete it when unused. Based on our experimentations, we have shown that the storage cost of VM disk image must not be neglected as done in previous work. Moreover, we have shown that using the accurate combinations of storage policies can dramatically reduce the storage cost (from 90% to 14% of the total bill).

6.6. Security for Virtualization and Clouds

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

6.6.1. *Improving Users' Isolation in IaaS: Virtual Machine Placement with Security Constraints*

Nowadays virtualization is used as the sole mechanism to isolate different users on Cloud platforms. Due to improper virtualization of micro-architectural components, data leak and modification can occur on public Clouds. Moreover, using the same attack vector (improper virtualization of micro-architectural components), it is possible to induce performance interferences, *i.e.* noisy neighbors. Using this approach, a VM can slow down and steal resources from concurrent VMs. In [43], we have proposed placement heuristics that take into account isolation requirements. We have modified three classical heuristics to take into account these requirements. Furthermore, we have proposed four new heuristics that take into account the hierarchy of the Cloud platforms and the isolation requirements. Finally, we have evaluated these heuristics and compare them with the modified classical ones. We have shown that our heuristics are performing at least as good as classical ones but are scaling better and are faster by a few order of magnitude than the classical ones.

6.6.2. *Security for Cloud Environment through Information Flow Properties Formalization with a First-Order Temporal Logic*

The main slowdown of Cloud activity comes from the lack of reliable security. The on-demand security concept aims at delivering and enforcing the client's security requirements. In [50], we have presented an approach, Information Flow Past Linear Time Logic (IF-PLTL), to specify how a system can support a large range of security properties. We have presented how to control those information flows from lower system events. We have given complete details over IF-PLTL syntax and semantics. Furthermore, that logic enables to formalize a large set of security policies. Our approach is exemplified with the Chinese Wall commercial-related policy. Finally, we have discussed the extension of IF-PLTL with dynamic relabeling to encompass more realistic situations through the dynamic domains isolation policy.

6.6.3. *Security Metrics for the Cloud Computing and Security-aware Virtual Machine Placement*

In a classic Cloud Computing scenario, a client connects to a provider platform/service and submits his computation requirements, sometimes known as Service Level Agree- ments (SLAs). Then, the platform executes the computation taking into account, in its allocation algorithms, criteria like data location, CPU usage or duration of a job. As security in Cloud Computing is a main concern, we propose to consider security as another criteria for jobs scheduling. Thus, two questions need to be answered. The first one is how a client can describe his needs in terms of security level and the second one is how the scheduler could leverage the security to satisfy the client requirements? To provide an answer, a system of security metrics is essential. Indeed, with appropriate metrics, we can quantify and compare the security level of our resources. Moreover, a client can easily describe his security requirements and the scheduler can allocate the fitted resources using these metrics. Unfortunately, such system of metrics is not yet available. Consequently, we developed a system of security metrics specific to the Cloud Computing and scheduling algorithms using these metrics for a Security-Aware Virtual Machine (VM) placement.

6.7. Self-healing of Operational Issues for Grid Computing

Participant: Frédéric Desprez.

Many scientists now formulate their computational problems as scientific workflows. Workflows allow researchers to easily express multi-step computational task. However, their large scale and the number of middleware systems involved in these gateways lead to many errors and faults. Fair quality of service (QoS) can be delivered, yet with important human intervention. Automating such operations is challenging for two reasons. First, the problem is online by nature because no reliable user activity prediction can be assumed, and new workloads may arrive at any time. Therefore, the considered metrics, decisions and actions have to remain simple and to yield results while the application is still executing. Second, it is non-clairvoyant due to the lack of information about applications and resources in production conditions. Computing resources are usually dynamically provisioned from heterogeneous clusters, clouds or desktop grids without any reliable estimate of their availability and characteristics. Models of application execution times are hardly available either, in particular on heterogeneous computing resources.

In collaboration with Rafaël Silva and Tristan Glatard, we proposed a general self-healing process for autonomous detection and handling of operational incidents in scientific workflow executions on grids. Instances are modeled as Fuzzy Finite State Machines (FuSM) where state degrees of membership are determined by an external healing process. Degrees of membership are computed from metrics assuming that incidents have outlier performance, e.g. a site or a particular invocation behaves differently than the others. These metrics make little assumptions on the application or resource characteristics. Based on incident degrees, the healing process identifies incident levels using thresholds determined from the platform history. A specific set of actions is then selected from association rules among incident levels. The healing process is parametrized on real application traces acquired in production on the European Grid Infrastructure (EGI).

To optimize task granularity in distributed scientific workflows, we presented a method that groups tasks when the fineness degree of the application becomes higher than a threshold determined from execution traces. Controlling the granularity of workflow activities executed on grids is required to reduce the impact of task queuing and data transfer time. Our method groups tasks when the fineness degree of the application, which takes into account the ratio of shared data and the queuing/round-trip time ratio, becomes higher than a threshold determined from execution traces. The algorithm also de-groups task groups when new resources arrive. Results showed that under stationary load, our fineness control process significantly reduces the makespan of all applications. Under non-stationary load, task grouping is penalized by its lack of adaptation, but our de-grouping algorithm corrects it in case variations in the number of available resources are not too fast [21].

To address unfairness among workflow executions, we proposed an algorithm to fairly allocate distributed computing resources among workflow executions to multi-user platforms. We consider a non-clairvoyant, online fairness problem where the platform workload, task costs, and resource characteristics are unknown and not stationary. We define a novel metric that quantifies unfairness based on the fraction of pending work in a workflow. It compares workflow activities based on their ratio of queuing tasks, their relative durations, and the performance of resources where tasks are running, as information becomes available during the execution. Our method is implemented and evaluated on 4 different applications executed in production conditions on EGI. Results show that our method can very significantly reduce the standard deviation of the slowdown, and the average value of our metric [22].

7. Partnerships and Cooperations

7.1. National Initiatives

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

7.1.1.1. 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.

7.1.2. French National Research Agency Projects (ANR)

7.1.2.1. 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.

7.1.2.2. 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, 2010-2013

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.

7.1.2.3. ANR COSINUS COOP, Multi Level Cooperative Resource Management, 3.5 years, ANR-09-COSI-001-01, 2009-2013

Participants: Frédéric Desprez, Christian Perez, Noua Toukourou.

The main goals of this project are to set up a cooperation as general as possible between programming models and resource management systems and to develop algorithms for efficient resource selection. In particular, the project targets the SALOME platform and the GRID-TLSE expert-site (http://gridtlse.org/) as an example of programming models, and PadicoTM, DIET and XtreemOS as examples of communication manager, grid middleware and distributed operating systems.

The project is led by Christian Perez.

7.1.2.4. ANR INFRA SONGS, Simulation Of Next Generation Systems, 4 years, ANR-12-INFRA-11, 2012-2015

Participants: Frédéric Desprez, Georgios Markomanolis, 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.

7.1.3. Inria Large Scale Initiative

7.1.3.1. 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.

7.1.3.2. C2S@Exa, Computer and Computational Sciences at Exascale, 4 years, 2013-2017

Participants: Frédéric Desprez, Christian Perez, Laurent Lefèvre.

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.

7.1.4. Inria ADT

7.1.4.1. 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.

7.2. European Initiatives

7.2.1. FP7 Projects

7.2.1.1. PRACE 2IP

Participants: Zhengxiong Hou, Vincent Lanore, Christian Perez.

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.

7.2.1.2. PaaSage

Participants: Amine Bsila, 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.

7.2.2. Collaborations in European Programs, except FP7

7.2.2.1. SEED4C

Program: Celtic-Plus
Project acronym: SEED4C
Project title: Security Embedded Element and Data privacy for the Cloud.
Duration: 2012-2015
Coordinator: Bertrand Marquet (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.

7.2.2.2. COST IC804

Participants: Ghislain Landry Tsafack Chetsa, Mohammed El Mehdi Diouri, Laurent Lefèvre.

Program: COST
Project acronym: IC804
Project title: : Energy efficiency in Large Scale Distributed Systems
Duration: 2009-2013
Coordinator: J.M. Pierson (IRIT Toulouse)
Other partners: 26 research institute and countries
Abstract: The COST Action IC0804 proposes realistic energy-efficient alternate solutions to share IT distributed resources. As large scale distributed systems gather and share more and more computing nodes and Storage resources, their energy consumption is exponentially increasing. While much effort is nowadays put into hardware specific solutions to lower energy consumptions, the need for a complementary approach is necessary at the distributed system level, i.e. middleware, network and applications. This Action characterizes the energy consumption and energy efficiencies of distributed applications. Then based on the current hardware adaptation possibilities and innovative algorithms it proposes adaptive and alternative approaches taking into account the energy saving dimension of the problem. This Action also characterizes the trade-off between energy savings and functional and non-functional parameters, including the economic dimension. Deliverables includes workshop proceedings, books, good practice leaflets fostering consciousness rise at ICT researchers, scientists, managers and users levels. Finally, benefits addresses scientific and societal needs.

7.2.2.3. COST IC0805

Participants: Ghislain Landry Tsafack Chetsa, Mohammed El Mehdi Diouri, Laurent Lefèvre.

Program: COST
Project acronym: IC0805
Project title: Open Network for High-Performance Computing on Complex Environments (ComplexHPC)
Duration: 2009-2013
Coordinator: Emmanuel Jeannot (Inria Bordeaux - Sud Ouest)
Other partners: 26 research institute and countries
Abstract: The main objective of the Action is to develop an integrated approach for tackling the challenges associated with heterogeneous and hierarchical systems for High Performance Computing.

7.2.2.4. 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.

7.3. International Initiatives

7.3.1. Inria International Partners

7.3.1.1. GreenTouch

Participants: Laurent Lefèvre, Jean-Patrick Gelas.

GreenTouch is a consortium of leading Information and Communications Technology (ICT) industry, academic and non-governmental research experts dedicated to fundamentally transforming communications and data networks, including the Internet, and significantly reducing the carbon footprint of ICT devices, platforms and networks.

In this project, we explore the design of virtual home gateway at large scale [62], [72] and participate in the SEASON project.

7.3.2. Inria International Labs

7.3.2.1. Inria-UIUC-NCSA Joint Laboratory for Petascale Computing

Participants: Eddy Caron, Frédéric Desprez, Mohammed El Mehdi Diouri, 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. Dissemination

8.1. Scientific Animation

- Eddy Caron
  - Co-Organization of Inria Booth in SuperComputing 2013, Denver, USA, November 16-22, 2013
  - Local Arrangements Co-Chairs of ICPP 2013 : 42th International Conference on Parallel Processing, Lyon, October 1-4, 2013
  - Program Committee Member: IPDPS 2013, HCW 2013, CLOSER 2013, CloudCom 2013, The First International Workshop Federative and Interoperable Cloud Infrastructures, (FedICI'2013) and CGCIC'13.
  - Reviews: Journal of Grid Computing, JPDC (Journal of Parallel and DIstributed Computing), CPE (Concurrency and Computation: Practice and Experience), TPDS (Transactions on Parallel and Distributed Systems)
- Frédéric Desprez
  - General chair of the VTDC'13 workshop (within HPDC'13), Co-chair of Closer'13, Co-chair of BigDataCloud'2013 (within EuroPAR'13)
  - Program Committee Member: CCGRID'2013, IPDPS'13, CGC2013, NPC2013, BigData'13, BDSE'2013
  - Reviews: Journal of Grid Computing, Parallel Processing Letters, Future Generation Computing Systems
  - Project Evaluation: Appel à Projets Recherche 2013 du Conseil régional d'Aquitaine, Région Languedoc-Roussillon, appel à projets "Chercheur(se)s d'avenir" 2013
  - Member of the selection committee for an associate professor at Université Joseph Fourier (Grenoble)
- Gilles Fedak
  - Member of Editorial Board of Springer Cloud Computing Journal, 2013
  - "Special Issue: Combined Special Issue of MapReduce and its Applications & Advanced topics on wireless sensor networks" Concurrency and Computation: Practice and Experience Issue edited by: Gilles Fedak, Keqiu Li, Kai Lin Volume 25, Issue 1 January 2013
  - Program Committee Member: 9th IEEE International Conference on e-Science, (e-Science 2013), 5th IEEE International Conference on Cloud Computing Technology and Science (CloudCom2013), 25th International Symposium on Computer Architecture and High-Performance Computing (SBAC-PAD), The First International Workshop Federative and Interoperable Cloud Infrastructures, (FedICI'2013), IEEE International Conference on Big Data and Distributed Systems (BDDS 2013), 10th IFIP International Conference on Network and Parallel Computing (NPC 2013), 15th IEEE International Conference on High Performance Computing and Communications (HPCC 2013), 22nd ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC 2013), 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID 2013), International Workshop on High Performance Data Intensive Computing (HPDIC2013), 17th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP 2013).

- Jean-Patrick Gelas
  - PC member (Cloud and Distributed Computing), AICCSA 2013 : 10th ACS/IEEE International conference on computer systems and applications, Fes, Morocco, May 27-30, 2013
  - PC member, ExtremeGreen: Extreme Green and Energy Efficiency in Large Scale Distributed Systems , Delft, The Netherlands, May 13-16, 2013
  - Reviews : ExtremGreen2013, Euro-Par2013, EnA-HPC 2013, CGC2013.
- Olivier Glück
  - Organization: local organization committee of ICPP
  - PC member: ExtremeGreen, International Conference on Cloud and Green Computing
  - Reviews: Cluster Computing, Euro-Par
  - University committee: Member of the CEVU (Conseil des Etudes et de la Vie Universitaire) of University Lyon 1, Member of the UFR Faculte des Sciences et Technologies faculty council of University Lyon 1, Member of the computer science department council of University Lyon 1
- Laurent Lefèvre
  - Co-organizer of the Green Days @ Lille event: "HPC & Clouds : are they green ?", Lille, November 28-29, 2013
  - Co-Organization of Inria Booth in SuperComputing 2013, Denver, USA, November 16-22, 2013
  - Co General chair of ICPP 2013 : 42th International Conference on Parallel Processing, Lyon, October 1-4, 2013
  - Co-Program chair of CGC 2013 : Third International Conference on Cloud and Green Computing ,Karlsruhe, Germany, September 30- October 2, 2013
  - Co-Workshop chair of ExtremeGreen 2013 : Extreme Green & Energy Efficiency in Large Scale Distributed Systems , Delft, The Netherlands, May 2013
  - Co-organizer of the Green Days @ Luxembourg event: "Energy efficiency : what else / what next ?", Luxembourg, January 28-29, 2013
  - Associate editor of IEEE Transactions on Cloud Computing (TCC)
  - Member of the selection committee for an associate professor position at Université Antilles Guyane (Pointe à Pitre)
- Christian Perez
  - Co-organization of the The Ninth Workshop of the Inria-Illinois Joint Laboratory on Petascale Computing, Lyon, France, June 12-14, 2013.
  - Programm committee member of the International Conference on Parallel Computing 2013 (ParCo), the 11th International Conference on Service Oriented Computing (ICSOC2013).
  - Reviewing for IEEE Transactions on Parallel and Distributed Systems, Future Generation Computer Systems, Parallel Computing, and Technique et Science Informatique.
  - Member of the selection committee for an associate professor position at University Lyon I/LIP.
- Frédéric Suter
  - Organization: 3rd edition of the SimGrid Users Days.
  - Project Evaluation: Comité d'Evaluation ANR SIMI2, Israel Science Fundation.
  - PC member: CLOSER, VTDC, IPDPS, SYNASC, VECPAR, EuroMPI.
  - Reviews: JPDC, Parallel Computing, TPDS, CCGrid

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

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

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

Licence: Eddy Caron, Architecture, Système et Réseaux, 36h, niveau L3, Ecole Normale Supérieure de Lyon, France

Master: Eddy Caron, Systèmes Distribués, 36h, niveau M1, Ecole Normale Supérieure de Lyon, France

Master: Eddy Caron, Projet Intégré, 36h, niveau M1, Ecole Normale Supérieure de Lyon, France

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

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

Master : Olivier Glück, Administration des Systèmes et des Réseaux, 16h, 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, Nouvelles Technologies, 85h, niveau M2, Université Lyon 1, France

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

Master: Eddy Caron, Gilles Fedak, Christian Perez and Jonathan Rouzaud-Cornabas Grid and Cloud Computing, 36h, niveau M2, Ecole Normale Supérieure de Lyon, France

Master/Doctorat : Gilles Fedak, MapReduce Environments: Design, Performance, Optimizations, II Escola Regional de Alto Desempenho - Região Nordeste, Savaldor de Bahia, Brazil, October 22, 2013

Master : Laurent Lefèvre is responsible of training periods for Research Master in ENS-Lyon

8.2.2. Supervision

HdR: Laurent Lefèvre, *Contributions à la flexibilité et à l'efficacité énergétique des systèmes distribués à grande échelle*, Ecole Normale Supérieure de Lyon, 14 Novembre 2013. [2]

PhD: Mohammed El Mehdi Diouri, *Efficacité énergétique dans le calcul très haute performance : application à la tolérance aux pannes et à la diffusion de données*, ENS LYON, September 27 2013, Isabelle Guerin-Lassous and Laurent Lefèvre and Olivier Glück. [1]

PhD: Ghislain Landry Tsafack Chetsa, *System Profiling and Green Capabilities for Large Scale and Distributed Infrastructures*, Ecole Normale Supérieure de Lyon, December 3rd, 2013. [3]

PhD: Rafaël Silva (Creatis, INSA Lyon), *A science-gateway for workflow executions: online and non-clairvoyant self-healing of workflow executions on grids*, Frédéric Desprez (dir), Tristan Glatard (Creatis, CNRS, co-dir), November 13, 2013.

PhD in progress: George Markomanolis, *Performance evaluation and prediction of parallel applications*, 12/2009, Frédéric Desprez (dir), Frédéric Suter (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: Sylvain Gault, *Ordonnancement de tâches avec le modèle MapReduce*, 11/2010, Frédéric Desprez (dir), Frédéric Suter (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).

8.2.3. Juries

- Eddy Caron
  - Ndye Massata NDIAYE, *Technique de gestion des défaillances dans les grilles informatiques tolérantes aux fautes*, PhD, Université Pierre et Marie Curie, France, September 2013, Reviewer
  - Sébastien Fourestier, *Redistribution dynamique parallèle efficace de la charge pour les problèmes numériques de très grandes tailles*, PhD, Université de Bordeaux I, Juin 2013, Reviewer
- Frédéric Desprez
  - Flavien Quesnel, *Vers une gestion coopérative des infrastructures virtualisées à grande échelle : le cas de l'ordonnancement* , PhD, Université de Nantes, February 2013, Reviewer
  - Thierry Coupaye, *Architecture dynamique des systèmes répartis – Expériences avec les composants logiciels* , HDR, Université Joseph Fourier, March 2013, Reviewer
  - Laurent Lefèvre, *Contributions à la flexibilité et à l'efficacité énergétique des systèmes distribués à grande échelle*, HDR, Ecole normale supérieure de Lyon, November 2013, Jury member
  - Ghislain Tsafack Chetsa, *System Profiling and Green Capabilities for Large Scale and Distributed Infrastructures*, PhD, Ecole normale supérieure de Lyon, December 3rd, 2013, Jury member
- Laurent Lefèvre
  - Yacine Kessaci, *Ordonnancement multi-critère sur Clouds*, PhD, Université Lille1, France, November 2013, Jury member
  - Remigiusz Modrzejewski, *Distribution and Storage in Networks*, PhD, Université de Nice, France, October 2013, Jury member
  - Anton Beloglazov, *Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Data Centers*, PhD, University of Melbourne, Australia, March 2013, Reviewer
- Christian Perez
  - Chen Wang, *Chemistry-Inspired Middleware for Flexible Service Composition and Adaptation*, PhD, Insa Rennes, May 28th, 2013, Reviewer.
  - Stefania Costache, *Market-based Autonomous Resource and Application Management in the Cloud*, PhD, University of Rennes I, July 3rd, 2013, Reviewer.
  - Joseph Emeras, *Workload Traces Analysis and Replay in Large Scale Distributed Systems*, PhD, University of Grenoble, France, October 1st, 2013, Reviewer.
  - Cyril Bordage, *Ordonnancement dynamique, adapté aux architectures hétérogènes, de la méthode multipôle pour les équations de Maxwell en électromagnétique*, PhD, Université de Bordeaux I, France, December 20th, 2013, Reviewer.

8.3. Popularization

- Laurent Lefèvre
  - Interview extract in Blog "Web developpement durable" : Internet & Data Center : bilan énergétique positif ou négatif , July 17, 2013
  - Interview for La Vie Journal, "Using Internet with less energy", June 13, 2013
  - Interview for Liberation News Paper, "Datacenters : la donnée écolo", April 15, 2013
  - Participant in the Panel on Green Computing, GDR GPL, Nancy, April 2013
  - Co-Editor for the Future Internet Assembly (FIA) Book 2013 : Validated Results and New Horizons, Alex Galis editor, LNCS, Volume 7858, Spinger, May 2013
- Jean-Patrick Gelas :
  - Invited speaker to Ecocity 2013, "Greening Clouds and High Performance Computing: Challenges and Initiatives", Nantes, France, September 23, 2013
  - Invited speaker to JDN2013 (Journe´e sur le de´cisionnel dans le Nuage), "Energy monitoring framework for high-performance cloud infrastructure", Bron, France, May 16, 2013
  - Invited speaker to JSF2013 (Journée Scientifique) about Energy, "Propositions to improve Data Centers and Clouds Energy Efficiency", Lyon, France, June 26, 2013

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. E. M. DIOURI. , *Efficacité énergétique dans le calcul très haute performance : application à la tolérance aux pannes et à la diffusion de données*, Ecole normale supérieure de lyon - ENS LYON, September 2013, http://hal.inria.fr/tel-00881094

[2] L. LEFEVRE. , *Contributions à la flexibilité et à l'efficacité énergétique des systèmes distribués à grande échelle*, Ecole normale supérieure de lyon - ENS LYON, November 2013, Habilitation à Diriger des Recherches, http://hal.inria.fr/tel-00925154

[3] G. L. TSAFACK CHETSA. , *Profilage système et leviers verts pour les infrastructures distribuées à grande échelle*, Ecole normale supérieure de lyon - ENS LYON, December 2013, Thèse rédigée en Aout 2013, http://hal.inria.fr/tel-00925320

Articles in International Peer-Reviewed Journals

[4] G. ANTONIU, J. BIGOT, C. BLANCHET, L. BOUGÉ, F. BRIANT, F. CAPPELLO, A. COSTAN, F. DESPREZ, G. FEDAK, S. GAULT, K. KEAHEY, B. NICOLAE, C. PÉREZ, A. SIMONET, F. SUTER, B. TANG, R. TERREUX. *Towards Scalable Data Management for Map-Reduce-based Data-Intensive Applications on Cloud and Hybrid Infrastructures*, in "International Journal of Cloud Computing (IJCC)", 2013, vol. 2, n^{o} 2/3 [*DOI :* 10.1504/IJCC.2013.055265], http://hal.inria.fr/hal-00767029

[5] J. BIGOT, Z. HOU, C. PÉREZ, V. PICHON. *A low level component model easing performance portability of HPC applications*, in "Computing", November 2013 [*DOI :* 10.1007/s00607-013-0368-3], http://hal.inria.fr/hal-00911231

[6] M. E. M. DIOURI, G. L. TSAFACK CHETSA, O. GLÜCK, L. LEFEVRE, J.-M. PIERSON, P. STOLF, G. DA COSTA. *Energy efficiency in HPC with and without knowledge of applications and services*, in "International Journal of High Performance Computing Applications", August 2013, vol. 27, n^o^ 3, pp. 232-243 [*DOI :* 10.1177/1094342013495304], http://hal.inria.fr/hal-00925313

[7] M. E. M. DIOURI, G. L. TSAFACK CHETSA, O. GLÜCK, L. LEFÈVRE, J.-M. PIERSON, P. STOLF, G. DA COSTA. *Energy efficiency in high-performance computing with and without knowledge of applications and services*, in "International Journal of High Performance Computing Applications", August 2013, vol. 27, n^o^ 3, pp. 232-243, http://hal.inria.fr/hal-00870615

[8] A.-C. ORGERIE, M. DIAS DE ASUNCAO, L. LEFEVRE. *A Survey on Techniques for Improving the Energy Efficiency of Large Scale Distributed Systems*, in "ACM Computing Surveys", December 2014, vol. 46, n^o^ 4, To appear, http://hal.inria.fr/hal-00767582

[9] G. L. TSAFACK CHETSA, L. LEFÈVRE, J.-M. PIERSON, P. STOLF, G. DA COSTA. *Exploiting Performance Counters to Predict and Improve Energy Performance of HPC Systems*, in "Future Generation Computer Systems", August 2013 [*DOI :* 10.1016/J.FUTURE.2013.07.010], http://hal.inria.fr/hal-00925306

Articles in Non Peer-Reviewed Journals

[10] F. BERTHOUD, B. BOUTHERIN, R. DAVID, R. FERRET, L. LEFÈVRE. *Réduire la consommation électrique des centres de données*, in "La Revue Durable - Dossier "Les Technologies de l'Information et de la Comumunication et l'Impératif de la Sobriété"", August 2013, vol. 49, http://hal.inria.fr/hal-00925171

International Conferences with Proceedings

[11] P. BEDARIDE, A. DEGOMME, S. GENAUD, A. LEGRAND, G. MARKOMANOLIS, M. QUINSON, M. STILLWELL, F. SUTER, B. VIDEAU. *Toward Better Simulation of MPI Applications on Ethernet/TCP Networks*, in "PMBS13 - 4th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems", Denver, United States, November 2013, http://hal.inria.fr/hal-00919507

[12] F. CAMILLO, E. CARON, R. GUIVARCH, A. HURAULT, C. KLEIN, C. PÉREZ. *Resource Management Architecture for Fair Scheduling of Optional Computations*, in "8th international conference on P2P, parallel, grid, cloud and internet computing (3PGCIC)", Compiegne, France, October 2013, pp. 113-120 [*DOI :* 10.1109/3PGCIC.2013.23], http://hal.inria.fr/hal-00865668

[13] E. CARON, L. ANH DUNG, A. LEFRAY, C. TOINARD. *Definition of security metrics for the Cloud Computing and security-aware virtual machine placement algorithms*, in "International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC 2013)", Beijing, China, IEEE Computer Society, 2013, pp. 125-131, http://hal.inria.fr/hal-00932401

[14] E. CARON, M.-D. FAYE, J. ROUZAUD-CORNABAS, O. THIARE. *Grid Middleware Modelling for Self-Adaptive Deployment*, in "The International Symposium on Autonomous Decentralized System (ISADS 2013)", Mexico City, Mexico, IEEE Computer Society, 2013, pp. 51-56, IEEE Catalog Number: CFP13099-PRT, http://hal.inria.fr/hal-00932400

[15] M. E. M. DIOURI, M. F. DOLZ, O. GLÜCK, L. LEFÈVRE, P. ALONSO, S. CATALÁN, R. MAYO, E. S. QUINTANA-ORTÍ. *Solving some Mysteries in Power Monitoring of Servers: Take Care of your Wattmeters!*, in "Energy Efficiency in Large Scale Distributed Systems conference (EE-LSDS)", Vienne, Austria, April 2013, pp. 3-18 [*DOI :* 10.1007/978-3-642-40517-4_1], http://hal.inria.fr/hal-00806504

[16] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE, F. CAPPELLO. *ECOFIT: A Framework to Estimate Energy Consumption of Fault Tolerance protocols during HPC executions*, in "13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)", Delft, Netherlands, May 2013, http://hal.inria.fr/hal-00806500

[17] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE, F. CAPPELLO. *Towards an Energy Estimator for Fault Tolerance Protocols*, in "18th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP)", Shenzhen, China, February 2013, pp. 313–314 [*DOI :* 10.1145/2442516.2442561], http://hal.inria.fr/hal-00806499

[18] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE. *SESAMES: a Smart-Grid Based Framework for Consuming Less and Better in Extreme-Scale Infrastructures*, in "GreenCom - 4th IEEE International Conference on Green Computing and Communications", Pékin, China, August 2013, pp. 187-194 [*DOI :* 10.1109/GREENCOM-ITHINGS-CPSCOM.2013.52], http://hal.inria.fr/hal-00870619

[19] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE, J.-C. MIGNOT. *Energy Estimation for MPI Broadcasting Algorithms in Large Scale HPC Systems*, in "EuroMPI - 20th European MPI Users' Group Meeting, Recent Advances in Message Passing Interface", Madrid, Spain, September 2013, pp. 111–116 [*DOI :* 10.1145/2488551.2488565], http://hal.inria.fr/hal-00870640

[20] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE, J.-C. MIGNOT. *Your Cluster is not Power Homogeneous: Take Care when Designing Green Schedulers!*, in "IGCC - 4th IEEE International Green Computing Conference", Arlington, United States, June 2013, pp. 1-10 [*DOI :* 10.1109/IGCC.2013.6604506], http://hal.inria.fr/hal-00870637

[21] R. FERREIRA DA SILVA, T. GLATARD, F. DESPREZ. *On-line, non-clairvoyant optimization of workflow activity granularity on grids*, in "19th International Conference Euro-Par 2013", Aachen, Germany, 2013, vol. 8097, pp. 255-266, http://hal.inria.fr/hal-00849988

[22] R. FERREIRA DA SILVA, T. GLATARD, F. DESPREZ. *Workflow fairness control on online and non-clairvoyant distributed computing platforms*, in "19th International Conference Euro-Par 2013", Aachen, Germany, 2013, vol. 8097, pp. 102-113, http://hal.inria.fr/hal-00849993

[23] R. FERREIRA DA SILVA, G. JUVE, E. DEELMAN, T. GLATARD, F. DESPREZ, D. THAIN, B. TOVAR, M. LIVNY. *Toward Fine-grained Online Task Characteristics Estimation in Scientific Workflows*, in "8th Workshop on Workflows in Support of Large-Scale Science (WORKS'13)", Denver, Colorado, United States, ACM, 2013, pp. 58–67 [*DOI :* 10.1145/2534248.2534254], http://hal.inria.fr/hal-00945369

[24] T. HIROFUCHI, A. LÈBRE, L. POUILLOUX. *Adding a Live Migration Model Into SimGrid, One More Step Toward the Simulation of Infrastructure-as-a-Service Concerns*, in "5th IEEE International Conference on Cloud Computing Technology and Science (IEEE CloudCom 2013)", Bristol, United Kingdom, December 2013, http://hal.inria.fr/hal-00861882

[25] Z. HOU, C. PÉREZ. *Performance Evaluation and Tuning of 2D Jacobi Iteration on Many-core Machines*, in "15th IEEE International Conference on High Performance Computing and Communications (HPCC 2013)", Zhangjiajie, China, November 2013, http://hal.inria.fr/hal-00862216

[26] M. IMBERT, L. POUILLOUX, J. ROUZAUD-CORNABAS, A. LÈBRE, T. HIROFUCHI. *Using the EXECO toolbox to perform automatic and reproducible cloud experiments*, in "1st International Workshop on UsiNg

and building ClOud Testbeds (UNICO, collocated with IEEE CloudCom 2013", Bristol, United Kingdom, September 2013, http://hal.inria.fr/hal-00861886

[27] M. MOCA, C. LITAN, G. COSMIN, G. FEDAK. *Advanced Promethee-based Scheduler Enriched with User-Oriented Methods*, in "In Proceedigs of the 10th IEEE Conference on Economics of Grids, Clouds, Systems, and Services (GECON 2013)", Zaragoza, Spain, Lecture Notes in Computer Science, Springer International Publishing, 2013, vol. 8193, pp. 161-172, http://hal.inria.fr/hal-00931985

[28] A. SIMONET, G. FEDAK, M. RIPEANU, S. AL-KISWANY. *Active Data: A Data-Centric Approach to Data Life-Cycle Management*, in "PDSW '13 - 8th Parallel Data Storage Workshop", Denver, United States, K. SCHWAN, D. HILDEBRAND (editors), ACM, November 2013, pp. 39-44 [*DOI :* 10.1145/2538542.2538566], http://hal.inria.fr/hal-00921080

[29] G. L. TSAFACK CHETSA, L. LEFÈVRE, J. PIERSON, P. STOLF, G. DA COSTA. *A User Friendly Phase Detection Methodology for HPC Systems' Analysis*, in "The 2013 IEEE International Conference on Green Computing and Communications (GreenCom 2013)", Beijing, China, August 2013, pp. 118 - 125 [*DOI :* 10.1109/GREENCOM-ITHINGS-CPSCOM.2013.43], http://hal.inria.fr/hal-00925299

[30] G. L. TSAFACK CHETSA, L. LEFÈVRE, P. STOLF. *A Three Step Blind Approach for Improving HPC Systems' Energy Performance*, in "EE-LSDS 2013, Energy Efficiency in Large Scale Distributed Systems conference", Vienne, Austria, Springer, April 2013, pp. 168-181 [*DOI :* 10.1007/978-3-642-40517-4_15], http://hal.inria.fr/hal-00925302

[31] P. VETTER, T. AYHAN, K. KANONAKIS, B. LANNOO, K. L. LEE, L. LEFÈVRE, C. MONNEY, F. SALIOU, X. YIN. *Towards Energy Efficient Wireline Networks, An Update From GreenTouch*, in "OptoElectronics and Communications Conference (OECC) 2013", Kyoto, Japan, 2013, http://hal.inria.fr/hal-00925191

National Conferences with Proceedings

[32] S. GAULT. *Ordonnancement dynamique des transferts dans MapReduce sous contrainte de bande passante*, in "ComPAS'13 / RenPar'21 - 21eme Rencontres francophones du Parallélisme", Grenoble, France, Inria Grenoble - Rhône-Alpes, January 2013, http://hal.inria.fr/hal-00820361

[33] V. LANORE, C. KLEIN. *De l'ordonnancement des applications multi-niveaux*, in "ComPAS'2013", Grenoble, France, January 2013, 8 p. , Track parallélisme, paper 9, http://hal.inria.fr/hal-00764007

Conferences without Proceedings

[34] D. BALOUEK, A. LÈBRE, F. QUESNEL. *Flauncher and DVMS – Deploying and Scheduling Thousands of Virtual Machines on Hundreds of Nodes Distributed Geographically*, in "IEEE International Scalable Computing Challenge (SCALE 2013), held in conjunction with CCGrid'2013", Delft, Netherlands, 2013, http://hal.inria.fr/hal-00920094

[35] N. BARD, S. BOIN, F. BOTHOREL, P. CHAUMEIL, P. COLLINET, M. DAYDÉ, B. DEPARDON, F. DESPREZ, M. FLÉ, A. FRANC, J.-M. FRIGERIO, O. GASCUEL, S. GUINDON, J.-F. GIBRAT, D. GIROU, P.-F. LAVALLÉE, V. LEFORT, G. LESAGE, M. RUGERI, E. RUINET, C. SEGUIN, S. THÉROND. *E-Biothon : Une plate-forme pour accélérer les recherches en biologie, santé et environnement*, in "Journées SUCCES 2013", Paris, France, November 2013, http://hal.inria.fr/hal-00927495

Scientific Books (or Scientific Book chapters)

[36] Z. AFOULKI, A. BOUSQUET, J. BRIFFAUT, L. CLÉVY, J. ROUZAUD-CORNABAS, C. TOINARD, B. VENELLE. *PIGA-Cloud : une protection obligatoire des environnements d'informatique en nuage*, in "RNTI SM2. Ingénierie des protocoles et Nouvelles technologies de la répartition", F. KRIEF, J.-M. PIERSON, P. ROOSE (editors), RNTI, 2013, http://hal.inria.fr/hal-00840731

[37] 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, http://hal.inria.fr/hal-00926460

[38] D. BALOUEK, A. CARPEN AMARIE, G. CHARRIER, F. DESPREZ, E. JEANNOT, E. JEANVOINE, A. LÈBRE, D. MARGERY, N. NICLAUSSE, L. NUSSBAUM, O. RICHARD, C. PÉREZ, F. QUESNEL, C. ROHR, L. SARZYNIEC. *Adding Virtualization Capabilities to the Grid'5000 Testbed*, in "Cloud Computing and Services Science", I. IVANOV, M. SINDEREN, F. LEYMANN, T. SHAN (editors), Communications in Computer and Information Science, Springer International Publishing, 2013, vol. 367, pp. 3-20 [*DOI :* 10.1007/978-3-319-04519-1_1], http://hal.inria.fr/hal-00946971

[39] S. BETGE-BREZETZ, A. BOUSQUET, J. BRIFFAUT, E. CARON, L. CLÉVY, M.-P. DUPONT, G.-B. KAMGA, J.-M. LAMBERT, A. LEFRAY, B. MARQUET, J. ROUZAUD-CORNABAS, L. TOCH, C. TOINARD, B. VENELLE. *Seeding the Cloud: An Innovative Approach to Grow Trust in Cloud Based Infrastructures*, in "The Future Internet - Future Internet Assembly 2013: Validated Results and New Horizons", A. GALIS, A. GAVRAS (editors), Lecture Notes in Computer Science, May 2013, vol. 7858, http://hal.inria.fr/hal-00840730

[40] 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], http://hal.inria.fr/hal-00931144

[41] 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, http://hal.inria.fr/hal-00925310

Research Reports

[42] P. BEDARIDE, S. GENAUD, A. DEGOMME, A. LEGRAND, G. MARKOMANOLIS, M. QUINSON, M. STILLWELL, F. SUTER, B. VIDEAU. , *Improving Simulations of MPI Applications Using A Hybrid Network Model with Topology and Contention Support*, Inria, May 2013, n^{o} RR-8300, 22 p. , http://hal.inria.fr/hal-00821446

[43] E. CARON, J. ROUZAUD-CORNABAS. , *Improving users' isolation in IaaS: Virtual Machine Placement with Security Constraints*, Inria, January 2014, n^{o} RR-8444, http://hal.inria.fr/hal-00924296

[44] E. CARON, L. TOCH, J. ROUZAUD-CORNABAS. , *Comparison on OpenStack and OpenNebula performance to improve multi-Cloud architecture on cosmological simulation use case*, Inria, December 2013, n^{o} RR-8421, 23 p. , http://hal.inria.fr/hal-00916908

[45] S. DELAMARE, G. FEDAK, D. KONDO, O. LODYGENSKY, P. KACSUK, J. KOVACS, F. ARAUJO. , *Advanced QoS Prototype for the EDGI Infrastructure*, Inria, May 2013, n^o RR-8295, http://hal.inria.fr/hal-00819907

[46] S. DELAMARE, G. FEDAK, D. KONDO, O. LODYGENSKY, P. KACSUK, J. KOVACS, F. ARAUJO. , *Intermediate QoS Prototype for the EDGI Infrastructure*, Inria, May 2013, n^o RR-8294, http://hal.inria.fr/hal-00819903

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

[48] F. DESPREZ, G. MARKOMANOLIS, F. SUTER. , *Evaluation of Profiling Tools for the Acquisition of Time Independent Traces*, Inria, July 2013, n^o RT-0437, 43 p. , http://hal.inria.fr/hal-00842396

[49] F. DESPREZ, J. ROUZAUD-CORNABAS. , *SimGrid Cloud Broker: Simulating the Amazon AWS Cloud*, Inria, November 2013, n^o RR-8380, 30 p. , http://hal.inria.fr/hal-00909120

[50] A. LEFRAY, J. ROUZAUD-CORNABAS, J. BRIFFAUT, C. TOINARD. , *Security for Cloud Environment through Information Flow Properties Formalization with a First-Order Temporal Logic*, Inria, December 2013, n^o RR-8420, 30 p. , http://hal.inria.fr/hal-00916882

[51] A. LÈBRE, J. PASTOR, M. BERTIER, F. DESPREZ, J. ROUZAUD-CORNABAS, C. TEDESCHI, P. ANEDDA, G. ZANETTI, R. NOU, T. CORTES, E. RIVIÈRE, T. ROPARS. , *Beyond The Cloud, How Should Next Generation Utility Computing Infrastructures Be Designed?*, Inria, July 2013, n^o RR-8348, http://hal.inria.fr/hal-00854204

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

Scientific Popularization

[53] M. E. M. DIOURI, O. GLÜCK, L. LEFÈVRE. *Smart Energy Management for Greener Supercomputing*, in "European Research Consortium for Informatics and Mathematics News", January 2013, vol. 2013, n^o 92, http://hal.inria.fr/hal-00943503

Other Publications

[54] F. BERTHOUD, L. LEFÈVRE, M. PARRY. *Que faites vous de vos e-déchets ?*, in "JRES 2013 : 10 èmes journées Réseaux", Montpellier, France, December 2013, JRES 2013 : 10 èmes journées Réseaux, http://hal.inria.fr/hal-00925576

[55] H. CASANOVA, A. GIERSCH, A. LEGRAND, M. QUINSON, F. SUTER. , *SimGrid: a Sustained Effort for the Versatile Simulation of Large Scale Distributed Systems*, 2013, 4 pages, submission to WSSSPE'13, http://hal.inria.fr/hal-00926437

[56] C. HERZOG, L. LEFÈVRE, J.-M. PIERSON. *Linking Academia and Industry for a Green (IT) Society*, in "Best Poster Award, ICT4S Conference : First International Conference on ICT for Sustainability", Zurich, Switzerland, February 2013, http://hal.inria.fr/hal-00925645

[57] Z. HOU, C. PÉREZ. , *Auto-tuning 2D Stencil Applications on Multi-core Parallel Machines*, March 2013, PRACE 2IP White Paper, http://hal.inria.fr/hal-00862270

[58] A. LEFRAY, E. CARON, J. ROUZAUD-CORNABAS, Z. HUAXI YULIN, A. BOUSQUET, J. BRIFFAUT, C. TOINARD. , *Security-Aware Models for Clouds*, June 2013, ACM Symposium on High-Performance Parallel and Distributed Computing, http://hal.inria.fr/hal-00840734

References in notes

[59] 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.

[60] R. GE, X. FENG, S. SONG, H.-C. CHANG, D. LI, K. W. CAMERON. *PowerPack: Energy Profiling and Analysis of High-Performance Systems and Applications*, in "IEEE Trans. Parallel Distrib. Syst.", May 2010, vol. 21, n° 5, pp. 658–671, http://dx.doi.org/10.1109/TPDS.2009.76

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

[62] J.-P. GELAS, L. LEFEVRE, T. ASSEFA, M. LIBSIE. *Virtualizing Home Gateways for Large Scale Energy Reduction in Wireline Networks*, in "Electronic Goes Green 2012 (EGG)", Berlin, Allemagne, September 2012, http://hal.inria.fr/hal-00766056

[63] 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

[64] 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

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

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

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

[68] 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

[69] 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

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

[71] 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

[72] P. VETTER, L. LEFEVRE, L. GASCA, K. KANONAKIS, L. KAZOVSKY, A. LEE, C. MONNEY, X. QIU, F. SALIOU, A. WONFOR. *Research Roadmap for Green Wireline Access*, in "Workshop on Green Communications and Networking during IEEE ICC’12", Ottawa, Canada, Canada, IEEE, 2012, http://hal.inria.fr/hal-00766811