Activity Report 2017

Project-Team MYRIADS

Design and Implementation of Autonomous Distributed Systems

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

Table of contents

Project-Team MYRIADS

Creation of the Team: 2010 January 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.1.6. - Cloud
A1.1.7. - Peer to peer
A1.1.9. - Fault tolerant systems
A1.1.13. - Virtualization
A1.2. - Networks
A1.2.4. - QoS, performance evaluation
A1.2.5. - Internet of things
A1.3. - Distributed Systems
A1.6. - Green Computing
A2.1.7. - Distributed programming
A2.2.3. - Run-time systems
A2.3.2. - Cyber-physical systems
A2.4.2. - Model-checking
A2.6. - Infrastructure software
A2.6.1. - Operating systems
A2.6.2. - Middleware
A2.6.3. - Virtual machines
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A4.7. - Access control
A4.9. - Security supervision
A4.9.1. - Intrusion detection
A4.9.3. - Reaction to attacks
A5.6. - Virtual reality, augmented reality
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.2.6. - Optimization
A6.2.7. - High performance computing
A7.1. - Algorithms
A8.2. - Optimization

Other Research Topics and Application Domains:

B2.3. - Epidemiology
B3.1. - Sustainable development
B3.2. - Climate and meteorology
B4.3. - Renewable energy production
B4.4. - Energy delivery

B4.4.1. - Smart grids
B4.5. - Energy consumption
B4.5.1. - Green computing
B5.1. - Factory of the future
B5.8. - Learning and training
B6.1. - Software industry
B6.1.1. - Software engineering
B6.3. - Network functions
B6.3.3. - Network Management
B6.4. - Internet of things
B6.5. - Information systems
B6.6. - Embedded systems
B8.1. - Smart building/home
B8.2. - Connected city
B8.5. - Smart society
B9.1. - Education
B9.1.1. - E-learning, MOOC
B9.1.2. - Serious games
B9.4.1. - Computer science
B9.6. - Reproducibility
B9.7. - Knowledge dissemination
B9.7.1. - Open access
B9.7.2. - Open data
B9.8. - Privacy

1. Personnel

Research Scientists
Christine Morin [Team leader, Inria, Senior Researcher, HDR]
Deborah Agarwal [Inria, Senior Researcher, Advanced Research Position]
Jean-Pierre Banâtre [Univ de Rennes I, Emeritus, until Oct 2017]
Anne-Cécile Orgerie [CNRS, Researcher]

Faculty Members
Nikos Parlavantzas [INSA Rennes, Associate Professor]
Jean-Louis Pazat [INSA Rennes, Professor, HDR]
Guillaume Pierre [Univ de Rennes I, Professor]
Martin Quinson [Ecole normale supérieure de Rennes, Professor, HDR]
Cédric Tedeschi [Univ de Rennes I, Associate Professor, HDR]

Post-Doctoral Fellows
Ehsan Ahvar [Inria, from Jun 2017]
Benjamin Camus [Inria]
Subarna Chatterjee [Inria, from Mar 2017]

PhD Students
Arif Ahmed [Univ de Rennes I]
Mehdi Belkhiria [Univ de Rennes I, from Dec 2017]

Yasmina Bouizem [Tlemcen University, from Nov 2017]
Clément El Baz [Inria, from Oct 2017]
Ali Fahs [Univ de Rennes I, from Oct 2017]
Anna Giannakou [Inria, until May 2017]
Baptiste Goupille–Lescar [Thales]
Loic Guegan [Ecole normale supérieure de Rennes, from Oct 2017]
Baptiste Goupille-Lescar [Thales, PhD Student, granted by CIFRE]
David Guyon [Univ de Rennes I]
Sabbir Hasan Rochi [INSA Rennes, PhD Student, until May 2017]
Yunbo Li [Inria, until Feb 2017]
The Anh Pham [Inria]
Genc Tato [Inria]
Amir Teshome Wonjiga [Inria]

Technical staff
Subarna Chatterjee [Inria, Feb 2017]
Toufik Boubehziz [Inria, from Mar 2017]
Yue Li [Inria, from Mar 2017]
Manh Linh Pham [Inria]
Sébastien Philippot [Inria]
Subhadeep Sarkar [Inria, from Feb 2017]
Alexandre Van Kempen [Inria, until Feb 2017]
Pascal Morillon [Univ. Rennes I, Engineer]
Matthieu Simonin [Inria, Research engineer]
David Margery [Inria, Research engineer, part-time]

Interns
Gairik Aluni [Inria, from May 2017 until Aug 2017]
Betsegaw Amersho [Univ de Rennes I, until Jun 2017]
Salsabil Amri [Univ de Rennes I, from May 2017 until Jul 2017]
Seyedmojtaba Imani [Inria, from Nov 2017]
Vinothkumar Nagasayanan [Univ de Rennes I, from May 2017 until Aug 2017]
Berenger Nguyen Nhon [Univ de Rennes I, from Jul 2017 until Aug 2017]

Administrative Assistant
Maryse Fouche [Inria]

External Collaborator
Louis Rilling [DGA]

2. Overall Objectives

2.1. General Objectives

MYRIADS is a joint team with INRIA, CNRS, UNIVERSITY RENNES 1, INSA RENNES and ENS RENNES. It is part of IRISA (D1 department on large scale systems) and INRIA RENNES – BRETAGNE ATLANTIQUE.

The objective of MYRIADS is to design and implement systems for autonomous service and resource management in interconnected and distributed clouds. The team tackles the challenges of dependable application execution and efficient resource management in highly distributed clouds.

2.2. Context

The MYRIADS team research activities are conducted in the context of the future of Internet.

Internet of Services. Myriads of applications are provided to more than one billion users [1] all over

the world. Over time, these applications are becoming more and more sophisticated, a given application being a composition of services likely to be executed on various sites located in different geographical locations. The Internet of Services is spreading all domains: home, administration, business, industry and science. Everyone is involved in the Internet of Services: citizens, enterprises, scientists are application, service and resource consumers and/or providers over the Internet.

Outsourcing. Software is provided as a service over the Internet. Myriads of applications are available on-line to billions of users as, for instance, *GoogleApps* (Gmail). After decades in which companies used to host their entire IT infrastructures in-house, a major shift is occurring where these infrastructures are outsourced to external operators such as Data Centers and Computing Clouds. In the Internet of Services, not only software but also infrastructure are delivered as a service. Clouds turned computing and storage into a utility. Just like water or electricity, they are available in virtually infinite amounts and their consumption can be adapted within seconds like opening or closing a water tap. The main transition, however, is the change in business models. Companies or scientists do not need to buy and operate their own data centers anymore. Instead, the compute and storage resources are offered by companies on a "pay-as-you-go" basis. There is no more need for large hardware investments before starting a business. Even more, the new model allows users to adapt their resources within minutes, e.g., scale up to handle peak loads or rent large numbers of computers for a short experiment. The risk of wasting money by either under-utilization or undersized data centers is shifted from the user to the provider.

Sharing and Cooperation. Sharing information and cooperating over the Internet are also important user needs both in the private and the professional spheres. This is exemplified by various services that have been developed in the last decade. Peer-to-peer networks are extensively used by citizens in order to share musics and movies. A service like *Flickr* allowing individuals to share pictures is also very popular. Social networks such as *FaceBook* or *LinkedIn* link millions of users who share various kinds of information within communities. Virtual organizations tightly connected to Grids allow scientists to share computing resources aggregated from different institutions (universities, computing centers...). The EGEE European Grid is an example of production Grid shared by thousands of scientists all over Europe.

2.3. Challenges

The term cloud was coined 10 years ago. Today cloud computing is widely adopted for a wide range of usage: information systems outsourcing, web service hosting, scientific computing, data analytics, back-end of mobile and IoT applications. There is a wide variety of cloud service providers (IaaS, PaaS, SaaS) resulting in difficulties for customers to select the services fitting their needs. Production clouds are powered by huge data centers that customers reach through the Internet. This current model raises a number of issues. Cloud computing generates a lot of traffic resulting in ISP providers needing to increase the network capacity. An increasing amount of always larger data centers consumes a lot of energy. Cloud customers experience poor quality of experience for highly interactive mobile applications as their requests are dealt with in data centers that are several hops away. The centralization of data in clouds also raises (i) security issues as clouds are a target of choice for attackers and (ii) privacy issues with data aggregation. Recently new cloud architectures have been proposed to overcome the scalability, latency, and energy issues of traditional centralized data centers. Various flavors of distributed cloud computing are emerging depending on the resources exploited: resources in the core network (distributed cloud), resources at the edge of the network (edge clouds) and even resources in the people swarms of devices (fog computing) enabling scalable cloud computing. These distributed clouds raise new challenges for resource and application management.

The ultimate goal of Myriads team is making highly distributed clouds sustainable. By sustainability we mean green, efficient and secure clouds. We plan to study highly distributed clouds including edge clouds and fog computing. In this context, we will investigate novel techniques for greening clouds including the

---

[1]According to World Stats, there are 3.67 billion Internet users i.e. more than half of the total world population in June 2016 http://www.internetworldstats.com/stats.htm.

optimization of energy consumption in distributed clouds in the context of smart grids. As more and more critical information system are outsourced in the cloud and personal data captured by sensors embedded in smart objects and smartphones are stored in the cloud, we will investigate security and privacy issues in two directions: cloud security monitoring and personal data protection in cloud-based IoT applications.

System research requires experimental validation based on simulation and/or prototyping. Reproducible experimentation is essential. We will contribute to the design and implementation of simulators well suited to the study of distributed clouds (architecture, energy consumption) and of large scale experimentation platforms for distributed systems enabling reproducible experiments.

3. Research Program

3.1. Introduction

In this section, we present our research challenges along four work directions: resource and application management in distributed cloud architectures for scaling clouds in Section 3.2, energy management strategies for greening clouds in Section 3.3, security and data protection aspects for securing cloud-based information systems and applications in Section 3.4, and methods for experimenting with clouds in Section 3.5.

3.2. Scaling clouds

3.2.1. Resource management in hierarchical clouds

The next generation of utility computing appears to be an evolution from highly centralized clouds towards more decentralized platforms. Today, cloud computing platforms mostly rely on large data centers servicing a multitude of clients from the edge of the Internet. Servicing cloud clients in this manner suggests that locality patterns are ignored: wherever the client issues his/her request from, the request will have to go through the backbone of the Internet provider to the other side of the network where the data center relies. Besides this extra network traffic and this latency overhead that could be avoided, other common centralization drawbacks in this context stand in limitations in terms of security/legal issues and resilience.

At the same time, it appears that network backbones are over-provisioned for most of their usage. This advocates for placing computing resources directly within the backbone network. The general challenge of resource management for such clouds stands in trying to be locality-aware: for the needs of an application, several virtual machines may exchange data. Placing them *close* to each others can significantly improve the performance of the application they compose. More generally, building an overlay network which takes the hierarchical aspects of the platform without being a hierarchical overlay – which comes with load balancing and resilience issues is a challenge by itself.

The results of these works are planned to be integrated into the Discovery initiative [60] which aims at revisiting OpenStack to offer a cloud stack able to manage utility computing platforms where computing resources are located in small computing centers in the backbone's PoPs (Point of Presence) and interconnected through the backbone's internal links.

3.2.2. Resource management in mobile edge clouds

Mobile edge cloud (MEC) infrastructures are composed of compute, storage and networking resources located at the edge of wide-area networks, in immediate proximity to the end users. Instead of treating the mobile operator's network as a high-latency dumb pipe between the end users and the external service providers, MEC platforms aim at deploying cloud functionalities *within* the mobile phone network, inside or close to the mobile access points. Doing so is expected to deliver added value to the content providers and the end users by enabling new types of applications ranging from Internet-of-Things applications to extremely interactive systems (e.g., augmented reality). Simultaneously, it will generate extra revenue streams for the mobile network operators, by allowing them to position themselves as cloud computing operators and to rent their already-deployed infrastructure to content and application providers.

Mobile edge clouds have very different geographical distribution compared to traditional clouds. While traditional clouds are composed of many reliable and powerful machines located in a very small number of data centers and interconnected by very high-speed networks, mobile edge cloud are composed of a very large number of points-of-presence with a couple of weak and potentially unreliable servers, interconnected with each other by commodity long-distance networks. This creates new demands for the organization of a scalable mobile edge computing infrastructure, and opens new directions for research.

The main challenges that we plan to address are:

- How should an edge cloud infrastructure be designed such that it remains scalable, fault-tolerant, controllable, energy-efficient, etc.?

- How should applications making use of edge clouds be organized? One promising direction is to explore the extent to which stream-data processing platforms such as Apache Spark and Apache Flink can be adapted to become one of the main application programming paradigms in such environments.

3.2.3. Self-optimizing applications in multi-cloud environments

As the use of cloud computing becomes pervasive, the ability to deploy an application on a multi-cloud infrastructure becomes increasingly important. Potential benefits include avoiding dependence on a single vendor, taking advantage of lower resource prices or resource proximity, and enhancing application availability. Supporting multi-cloud application management involves two tasks. First, it involves selecting an initial multi-cloud application deployment that best satisfies application objectives and optimizes performance and cost. Second, it involves dynamically adapting the application deployment in order to react to changes in execution conditions, application objectives, cloud provider offerings, or resource prices. Handling price changes in particular is becoming increasingly complex. The reason is the growing trend of providers offering sophisticated, dynamic pricing models that allow buying and selling resources of finer granularities for shorter time durations with varying prices.

Although multi-cloud platforms are starting to emerge, these platforms impose a considerable amount of effort on developers and operations engineers, provide no support for dynamic pricing, and lack the responsiveness and scalability necessary for handling highly-distributed, dynamic applications with strict quality requirements. The goal of this work is to develop techniques and mechanisms for automating application management, enabling applications to cope with and take advantage of the dynamic, diverse, multi-cloud environment in which they operate.

The main challenges arising in this context are:

- selecting effective decision-making approaches for application adaptation,

- supporting scalable monitoring and adaptation across multiple clouds,

- performing adaptation actions in a cost-efficient and safe manner.

3.3. Greening clouds

ICT (Information and Communications Technologies) ecosystem now approaches 5% of world electricity consumption and this ICT energy use will continue grow fast because of the information appetite of Big Data, big networks and big infrastructures as Clouds that unavoidably leads to big power.

3.3.1. Smart grids and clouds

We propose exploiting Smart Grid technologies to come to the rescue of energy-hungry Clouds. Unlike in traditional electrical distribution networks, where power can only be moved and scheduled in very limited ways, Smart Grids dynamically and effectively adapt supply to demand and limit electricity losses (currently 10% of produced energy is lost during transmission and distribution).

For instance, when a user submits a Cloud request (such as a Google search for instance), it is routed to a data center that processes it, computes the answer and sends it back to the user. Google owns several data centers spread across the world and for performance reasons, the center answering the user's request is more likely to be the one closest to the user. However, this data center may be less energy efficient. This request may have consumed less energy, or a different kind of energy (renewable or not), if it had been sent to this further data center. In this case, the response time would have been increased but maybe not noticeably: a different trade-off between quality of service (QoS) and energy-efficiency could have been adopted.

While Clouds come naturally to the rescue of Smart Grids for dealing with this big data issue, little attention has been paid to the benefits that Smart Grids could bring to distributed Clouds. To our knowledge, no previous work has exploited the Smart Grids potential to obtain and control the energy consumption of entire Cloud infrastructures from underlying facilities such as air conditioning equipment (which accounts for 30% to 50% of a data center's electricity bill) to network resources (which are often operated by several actors) and to computing resources (with their heterogeneity and distribution across multiple data centers). We aim at taking advantage of the opportunity brought by the Smart Grids to exploit renewable energy availability and to optimize energy management in distributed Clouds.

3.3.2. *Energy cost models*

Cloud computing allows users to outsource the computer resources required for their applications instead of using a local installation. It offers on-demand access to the resources through the Internet with a pay-as-you-go pricing model.However, this model hides the electricity cost of running these infrastructures.

The costs of current data centers are mostly driven by their energy consumption (specifically by the air conditioning, computing and networking infrastructure). Yet, current pricing models are usually static and rarely consider the facilities' energy consumption per user. The challenge is to provide a fair and predictable model to attribute the overall energy costs per virtual machine and to increase energy-awareness of users.

Another goal consists in better understanding the energy consumption of computing and networking resources of Clouds in order to provide energy cost models for the entire infrastructure including incentivizing cost models for both Cloud providers and energy suppliers. These models will be based on experimental measurement campaigns on heterogeneous devices. Inferring a cost model from energy measurements is an arduous task since simple models are not convincing, as shown in our previous work. We aim at proposing and validating energy cost models for the heterogeneous Cloud infrastructures in one hand, and the energy distribution grid on the other hand. These models will be integrated into simulation frameworks in order to validate our energy-efficient algorithms at larger scale.

3.3.3. *Energy-aware users*

In a Cloud moderately loaded, some servers may be turned off when not used for energy saving purpose. Cloud providers can apply resource management strategies to favor idle servers. Some of the existing solutions propose mechanisms to optimize VM scheduling in the Cloud. A common solution is to consolidate the mapping of the VMs in the Cloud by grouping them in a fewer number of servers. The unused servers can then be turned off in order to lower the global electricity consumption.

Indeed, current work focuses on possible levers at the virtual machine suppliers and/or services. However, users are not involved in the choice of using these levers while significant energy savings could be achieved with their help. For example, they might agree to delay slightly the calculation of the response to their applications on the Cloud or accept that it is supported by a remote data center, to save energy or wait for the availability of renewable energy. The VMs are black boxes from the Cloud provider point of view. So, the user is the only one to know the applications running on her VMs.

We plan to explore possible collaborations between virtual machine suppliers, service providers and users of Clouds in order to provide users with ways of participating in the reduction of the Clouds energy consumption. This work will follow two directions: 1) to investigate compromises between power and performance/service quality that cloud providers can offer to their users and to propose them a variety of options adapted to their

workload; and 2) to develop mechanisms for each layer of the Cloud software stack to provide users with a quantification of the energy consumed by each of their options as an incentive to become greener.

3.4. Securing clouds

3.4.1. *Security monitoring SLO*

While the trend for companies to outsource their information system in clouds is confirmed, the problem of securing an information system becomes more difficult. Indeed, in the case of infrastructure clouds, physical resources are shared between companies (also called tenants) but each tenant controls only parts of the shared resources, and, thanks to virtualization, the information system can be dynamically and automatically reconfigured with added or removed resources (for example starting or stopping virtual machines), or even moved between physical resources (for example using virtual machine migration). Partial control of shared resources brings new classes of attacks between tenants, and security monitoring mechanisms to detect such attacks are better placed out of the tenant-controlled virtual information systems, that is under control of the cloud provider. Dynamic and automatic reconfigurations of the information system make it unfeasible for a tenant's security administrator to setup the security monitoring components to detect attacks, and thus an automated self-adaptable security monitoring service is required.

Combining the two previous statements, there is a need for a dependable, automatic security monitoring service provided to tenants by the cloud provider. Our goal is to address the following challenges to design such a security monitoring service:

1. to define relevant Service-Level Objectives (SLOs) of a security monitoring service, that can figure in the Service-Level Agreement (SLA) signed between a cloud provider and a tenant;
2. to design heuristics to automatically configure provider-controlled security monitoring software components and devices so that SLOs are reached, even during automatic reconfigurations of tenants' information systems;
3. to design evaluation methods for tenants to check that SLOs are reached.

Moreover in challenges 2 and 3 the following sub-challenges must be addressed:

- although SLAs are bi-lateral contracts between the provider and each tenant, the implementation of the contracts is based on shared resources, and thus we must study methods to combine the SLOs;
- the designed methods should have a minimal impact on performance.

3.4.2. *Data Protection in Cloud-based IoT Services*

The Internet of Things is becoming a reality. Individuals have their own swarm of connected devices (e.g. smartphone, wearables, and home connected objects) continually collecting personal data. A novel generation of services is emerging exploiting data streams produced by the devices' sensors. People are deprived of control of their personal data as they don't know precisely what data are collected by service providers operating on Internet (oISP), for which purpose they could be used, for how long they are stored, and to whom they are disclosed. In response to privacy concerns the European Union has introduced, with the Global Data Protection Regulation (GDPR), new rules aimed at enforcing the people's rights to personal data protection. The GDPR also gives strong incentives to oISPs to comply. However, today, oISPs can't make their systems GDPR-compliant since they don't have the required technologies. We argue that a new generation of system is mandatory for enabling oISPs to conform to the GDPR. We plan to to design an open source distributed operating system for native implementation of new GDPR rules and ease the programming of compliant cloud-based IoT services. Among the new rules, transparency, right of erasure, and accountability are the most challenging ones to be implemented in IoT environments but could fundamentally increase people's confidence in oISPs. Deployed on individuals' swarms of devices and oISPs' cloud-hosted servers, it will enforce detailed data protection agreements and accountability of oISPs' data processing activities. Ultimately we will show to what extend the new GDPR rules can be implemented for cloud-based IoT services.

3.5. Experimenting with Clouds

Cloud platforms are challenging to evaluate and study with a sound scientific methodology. As with any distributed platform, it is very difficult to gather a global and precise view of the system state. Experiments are not reproducible by default since these systems are shared between several stakeholder. This is even worsen by the fact that microscopic differences in the experimental conditions can lead to drastic changes since typical Cloud applications continuously adapt their behavior to the system conditions.

3.5.1. *Experimentation methodologies for clouds*

We propose to combine two complementary experimental approaches: direct execution on testbeds such as Grid'5000, that are eminently believable but rather labor intensive, and simulations (using *e.g.* SimGrid) that are much more light-weighted, but requires are careful assessment. One specificity of the Myriads team is that we are working on these experimental methodologies *per se*, raising the standards of *good experiments* in our community.

We plan to make SimGrid widely usable beyond research laboratories, in order to evaluate industrial systems and to teach the future generations of cloud practitioners. This requires to frame the specific concepts of Cloud systems and platforms in actionable interfaces. The challenge is to make the framework both easy to use for simple studies in educational settings while modular and extensible to suit the specific needs of every advanced industrial-class users.

We aim at leveraging the convergence opportunities between methodologies by further bridging simulation and real testbeds. The predictions obtained from the simulator should be validated against some real-world experiments obtained on the target production platform, or on a similar platform. This (in)validation of the predicted results often improves the understanding of the modeled system. On the other side, it may even happen that the measured discrepancies are due to some mis-configuration of the real platform that would have been undetected without this (in)validation study. In that sense, the simulator constitutes a precious tool for the quality assurance of real testbeds such as Grid'5000.

Scientists need more help to make there Cloud experiments fully reproducible, in the sprit of Open Science exemplified by the HAL Open Archive, actively backed by Inria. Users still need practical solutions to archive, share and compare the whole experimental settings, including the raw data production (particularly in the case of real testbeds) and their statistical analysis. This is a long lasting task to which we plan to collaborate through the research communities gathered around the Grid'5000 and SimGrid scientific instruments.

Finally, since correction and performance can constitute contradictory goals, it is particularly important to study them jointly. To that extend, we want to bridge the performance studies, that constitute our main scientific heritage, to correction studies leveraging formal techniques. SimGrid already includes to exhaustively explore the possible executions. We plan to continue this work to ease the use of the relevant formal methods to the experimenter studying Cloud systems.

3.5.2. *Use cases*

In system research it is important to work on real-world use cases from which we extract requirements inspiring new research directions and with which we can validate the system services and mechanisms we propose. In the framework of our close collaboration with the Data Science Technology department of the LBNL, we will investigate cloud usage for scientific data management. Next-generation scientific discoveries are at the boundaries of datasets, e.g., across multiple science disciplines, institutions and spatial and temporal scales. Today, data integration processes and methods are largely adhoc or manual. A generalized resource infrastructure that integrates knowledge of the data and the processing tasks being performed by the user in the context of the data and resource lifecycle is needed. Clouds provide an important infrastructure platform that can be leveraged by including knowledge for distributed data integration.

4. Application Domains

4.1. Main Application Domains

The Myriads team investigates the design and implementation of system services. Thus its research activities address a broad range of application domains. We validate our research results with selected use cases in the following application domains:

- Smart city services,
- Smart grids,
- Energy and sustainable development,
- Home IoT applications,
- Bio-informatics applications,
- Data science applications,
- Computational science applications,
- Numerical simulations.

5. Highlights of the Year

5.1. Highlights of the Year

- The FogGuru project was accepted and started on September 1st 2017. FogGuru is a European H2020 Maria-Skłodowska-Curie Action (MSCA) European Industrial Doctorate (EID) training project which aims to to train eight talented PhD students with an innovative and intersectoral research program to constitute the next generation of European Cloud and Fog computing experts. It is coordinated by Guillaume Pierre.
- Cédric Tedeschi defended his *habilitation à diriger des recherches* summarizing his research activity of the last seven years.

5.1.1. Awards

- Best paper award for Timothée Haudebourg and Anne-Cécile Orgerie at the International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2017) for the paper entitled "On the Energy Efficiency of Sleeping and Rate Adaptation for Network Devices"
- Christine Morin has been selected to be included in the 2017 list of "N2Women:Stars in Computer Networking and Communications". The "N2Women:Stars in Computer Networking and Communications" is an annual list focusing in amazing women who have had a major impact in networking and/or communications.

BEST PAPER AWARD:

[34]
T. HAUDEBOURG, A.-C. ORGERIE. *On the Energy Efficiency of Sleeping and Rate Adaptation for Network Devices*, in "ICA3PP 2017 - 17th International Conference on Algorithms and Architectures for Parallel Processing", Helsinki, Finland, Springer, August 2017, vol. 10393, pp. 132-146, Best Paper Award [*DOI :* 10.1007/978-3-319-65482-9_9], https://hal.archives-ouvertes.fr/hal-01575513

6. New Software and Platforms

6.1. ConPaaS

KEYWORDS: Cloud computing - PaaS
SCIENTIFIC DESCRIPTION: Contact:

Guillaume Pierre, Guillaume.Pierre@irisa.fr URL:

http://www.conpaas.eu/ Status:

Version 1.4.2 License:

BSD Presentation:

ConPaaS [60] is a runtime environment for hosting applications in the cloud. It aims at offering the full power of the cloud to application developers while shielding them from the associated complexity of the cloud. ConPaaS is designed to host both high-performance scientific applications and online Web applications. It automates the entire life-cycle of an application, including collaborative development, deployment, performance monitoring, and automatic scaling. This allows developers to focus their attention on application-specific concerns rather than on cloud-specific details. Active contributors (from the Myriads team):

Eliya Buyukkaya, Ancuta Iordache, Morteza Neishaboori, Guillaume Pierre, Dzenan Softic, Genc Tato, Teodor Crivat. Impact:

ConPaaS is recognized as one of the major open-source PaaS environments. It is being developed by teams in Rennes, Amsterdan, Berlin and Ljubljana. Technology transfer of ConPaaS technology is ongoing in the context of the MC-DATA EIT ICT Labs project.

FUNCTIONAL DESCRIPTION: ConPaaS is a runtime environment for hosting applications in the cloud. It aims at offering the full power of the cloud to application developers while shielding them from the associated complexity of the cloud. ConPaaS is designed to host both high-performance scientific applications and online Web applications. It automates the entire life-cycle of an application, including collaborative development, deployment, performance monitoring, and automatic scaling. This allows developers to focus their attention on application-specific concerns rather than on cloud-specific details.

- Participants: Ancuta Iordache, Dzenan Softic, Eliya Buyukkaya, Genc Tato, Guillaume Pierre, Morteza Neishaboori and Teodor Crivat
- Contact: Guillaume Pierre
- URL: http://www.conpaas.eu/

6.2. GinFlow

KEYWORDS: Dynamic adaptation - Distributed Applications - Distributed - Distributed computing - Workflow - Framework

FUNCTIONAL DESCRIPTION: GinFlow decentralizes the coordination of the execution of workflow-based applications. GinFlow relyies on an architecture where multiple service agents (SA) coordinate each others through a shared space containing the workflow description and current status. GinFlow allows the user to define several variants of a workflow and to switch from one to the other during run time.

- Participants: Cédric Tedeschi, Hector Fernandez, Javier Rojas Balderrama, Matthieu Simonin and Thierry Priol
- Partner: Université de Rennes 1
- Contact: Cédric Tedeschi
- URL: http://ginflow.inria.fr

6.3. Merkat

FUNCTIONAL DESCRIPTION: Merkat is a platform that allows users of an organization to automatically manage and scale their applications while maximizing the infrastructure's utilization. Merkat is generic and extensible, allowing users to automate the application deployment and management process. Users have the flexibility to control how many resources are allocated to their applications and to define their own resource demand adaptation policies. Merkat applies an unique approach to multiplex the infrastructure capacity between the applications, by implementing a proportional-share market and allowing applications to adapt autonomously to resource price and their given performance objectives. The price of the acquired resources acts as a control mechanism to ensure that resources are distributed to applications according to the user's value for them. Merkat was evaluated on Grid'5000 with several scientific applications.

- Participants: Christine Morin, Nikolaos Parlavantzas and Stefania Costache
- Contact: Nikolaos Parlavantzas
- URL: http://www.irisa.fr/myriads/software/Merkat/

6.4. PaaSage Adapter

KEYWORDS: Cloud computing - Dynamic adaptation - Cloud applications management
FUNCTIONAL DESCRIPTION: The purpose of the Adapter is to transform the current configuration of a cloud application into a target configuration in an efficient and safe way. The Adapter is part of PaaSage, an open-source platform for modeling, deploying and executing applications on different clouds in an optimal manner. The Adapter has the following responsibilities: (1) validating reconfiguration plans, (2) applying the plans to the running system, and (3) maintaining an up-to-date representation of the current system state.

- Contact: Nikolaos Parlavantzas
- URL: https://team.inria.fr/myriads/software-and-platforms/paasage-adapter/

6.5. SAIDS

self-adaptable intrusion detection system
KEYWORDS: Cloud - Security
FUNCTIONAL DESCRIPTION: SAIDS is a self-adaptable intrusion detection system for IaaS clouds. To maintain an effective level of intrusion detection, SAIDS monitors changes in the virtual infrastructure of a Cloud environment and reconfigures its components (security probes) accordingly. SAIDS can also reconfigure probes in the case of a change in the list of running services.

- Authors: Anna Giannakou and Jean-Léon Cusinato
- Contact: Christine Morin

6.6. SimGrid

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

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

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

RELEASE FUNCTIONAL DESCRIPTION:

- Four releases in 2017. Major changes:
  - S4U: many progress, toward SimGrid v4.0. About 80% of the features offered by SimDag and MSG are now integrated, along with examples. Users can now write plugins to extend SimGrid.
  - SMPI: Support MPI 2.2, RMA support, Convert internals to C++.
  - Java: Massive memleaks and performance issues fixed.
  - New models: Multi-core VMs, Energy consumption due to the network
  - All internals are now converted to C++, and most of our internally developped data containers were replaced with std::* constructs.
  - (+ bug fixes, cleanups and documentation improvements)
- Participants: Adrien Lèbre, Arnaud Legrand, Augustin Degomme, Florence Perronnin, Frédéric Suter, Jean-Marc Vincent, Jonathan Pastor, Jonathan Rouzaud-Cornabas, Luka Stanisic, Mario Südholt and Martin Quinson
- Partners: CNRS - ENS Rennes
- Contact: Martin Quinson
- URL: http://simgrid.gforge.inria.fr/

6.7. DiFFuSE

Distributed framework for cloud-based epidemic simulations
KEYWORDS: Simulation - Cloud
FUNCTIONAL DESCRIPTION: The DiFFuSE framework enables simulations of epidemics to take full advantage of cloud environments. The framework provides design support, reusable code, and tools for building and executing epidemic simulations. Notably, the framework automatically handles failures and supports elastic allocation of resources from multiple clouds.

- Contact: Nikolaos Parlavantzas
- URL: https://team.inria.fr/myriads/software-and-platforms/diffuse/

7. New Results

7.1. Scaling Clouds

7.1.1. *Fog Computing*

Participants: Guillaume Pierre, Arif Ahmed, Ali Fahs, Alexandre Van Kempen, Salsabil Amri, Vinothkumar Nagasayanan, Berenger Nguyen Nhon.

Fog computing aims to extend datacenter-based cloud platforms with additional compute, networking and storage resources located in the immediate vicinity of the end users. By bringing computation where the input data was produced and the resulting output data will be consumed, fog computing is expected to support new types of applications which either require very low network latency (e.g., augmented reality applications) or which produce large data volumes which are relevant only locally (e.g., IoT-based data analytics).

Fog computing architectures are fundamentally different from those of traditional cloud platforms: to provide computing resources in physical proximity of any end user, fog computing platforms must necessarily rely on very large numbers of small Points-of-Presence connected to each other with commodity networks whereas clouds are typically organized with a handful of extremely powerful data centers connected by dedicated ultra-high-speed networks. This geographical spread also implies that the machines used in any Point-of-Presence may not be datacenter-grade servers but much weaker commodity machines.

We investigated the challenges of efficiently deploying Docker containers in fog platforms composed of tiny single-board computers such as Raspberry PIs. This operation can be painfully slow, in the order of multiple minutes depending on the container's image size and network condition. We showed that this bad performance is not only due to hardware limitations, but it is largely due to inefficiencies in the way Docker implements the container's image download operation. We proposed a number of optimization techniques which , when combined together, make container deployment up to 4 times faster than the vanilla Docker implementation. A publication on this topic is under submission.

Although fog computing infrastructures are fundamentally distributed, their management part still remains centralized: a single node (or small group of nodes) is in charge of maintaining the list of available server machines, monitoring them, distributing software to them, deciding which server must take care of which task, etc. We therefore aim to reduce the discrepancy between the broadly distributed compute/storage resources and the – currently – extremely centralized control of these resources, by focusing first on the resource scheduling function. This project has just started, and we expect to obtain the first results in 2018.

7.1.2. *Edge Cloud*
Participants: Anne-Cécile Orgerie, Cédric Tedeschi, Matthieu Simonin, Ehsan Ahvar, Genc Tato.

Myriads is involved in the Discovery project, whose goal is to design, develop and experiment a software stack for a distributed cloud platform where resources are directly injected into the backbone of the network [60]. To this end, we designed a novel family of overlay network to operate messaging and routing on top of such a distributed utility computing platform. The big picture of these overlays was described in a workshop [47].

7.1.3. *Community Clouds*
Participant: Jean-Louis Pazat.

In this work we consider an infrastructure based on devices (such as Internet boxes and NAS) owned and operated by end-users. A typical use-case is the sharing of CPU and storage capabilities by a community of users. This sharing is operated by hosting services to local and remote users. The devices of this distributed infrastructure have heterogeneous capabilities and no guaranteed availability. It is therefore challenging to ensure to the guest service a minimal hosting service level, such as availability or QoS.

We consider services build as an application based on micro-services. Such an application is deployed on the infrastructure by instantiating its constituent micro-services on some devices. One micro-services may rely on others micro-services to enable its own service. The performance of the resulting application is therefore highly dependent from the placement for each micro-service instance. Device parameters like CPU capabilities or network bandwidth and latency have a significant impact on the resulting response time of the micro-service, hence the application.

We explore solutions to adapt the placement of the micro-services to the capabilities of the infrastructure. As a first step, we are studying a static system where these capabilities are not variating. The placement decision can be expressed as the solution of an NP-Complete optimization problem. We have shown that a solution for this problem can be found with reasonably good precision using a meta-heuristic called Particle Swarm Optimization. The next step will be to study how this solution can be adapted in a dynamic system by considering the variations of the CPU and Network parameters and the availability of the devices.

This work in done in the context of Bruno Stevant's PhD thesis co-advised by Jean-Louis Pazat (Bruno Stevant is a member of OCIF team).

7.1.4. *Evaluation of Data Stream Processing Frameworks in Clouds*
Participants: Christine Morin, Deborah Agarwal, Subarna Chatterjee.

We address the problem of selecting a correct stream processing framework for a given application to be executed within a specific physical infrastructure. For this purpose, we have performed a thorough comparative analysis of three data stream processing platforms – Apache Flink, Apache Storm, and Twitter Heron (the enhanced version of Apache Storm), that are chosen based on their potential to process both streams and batches in real-time. For the comparative performance analysis of the chosen platforms, we have experimented using 8-node clusters on Grid5000 experimentation testbed and have selected a wide variety of applications ranging from a conventional benchmark (word count application) to sensor-based IoT application (air quality monitoring application) and statistical batch processing application (flight delay analysis application). The work focuses to analyze the performance of the frameworks in terms of the volume and throughput of data streams that each framework can possibly handle. The impact of each framework on the operating system is analyzed by experimenting and studying the resource utilization of the platforms in terms of CPU utilization, memory consumption. The energy consumption of the platforms is also studied to understand the suitability of the platforms towards green computing. Last, but not the least, the fault tolerance of the frameworks is also studied and analyzed. Lessons learnt from this work will precisely enlighten IaaS cloud end-users to wisely choose the correct streaming platform in order to run a particular application within a given set of VMs and will assist the cloud-providers to rationally allocate VMs equipped with a particular stream processing framework to PaaS cloud-users for running a specific streaming application. A paper has been submitted to an international conference in November 2017.

7.1.5. *Stream Processing for Maritime Surveillance*
Participants: Pascal Morillon, Christine Morin, Matthieu Simonin, Cédric Tedeschi.

In the context of maritime surveillance, and of the Sesame Project, we started the design and implementation of a platform dedicated to the batch and real-time processing of AIS messages sent by ships to inform about their identity, position and destination among other pieces of information.

Having use cases in mind such as detecting ships entering a protected areas, or ships having suspect behaviors, we designed a software architecture able to process AIS messages and produce synthetic data so as to answer these questions.

First experiments using a preliminary version of this platform have been conducted over the Grid'5000 platform using an archive of one-month of the AIS messages collected globally during March 2017. In particular, we've been able to index these messages using ElasticSearch [2] and visualize them using Kibana [3].

The architecture has been described in a poster presented at BiDS'17 [56].

7.1.6. *Adaptive deployment for multi-cloud applications*
Participants: Nikos Parlavantzas, Manh Linh Pham.

This work builds on the Adapter system, developed in the context of the PaaSage European project (2012-2016). The Adapter is part of the PaaSage open-source platform, a holistic solution for supporting the automatic deployment and execution of multi-cloud applications. Specifically, the Adapter is responsible for dynamic, cross-cloud application adaptation, taking into account adaptation costs and benefits in making deployment decisions. In 2017, we improved the Adapter and performed a comprehensive evaluation using experiments in a multi-cloud environment. The results demonstrate that Adapter supports automated multi-cloud adaptation while optimizing the performance and cost of the application. The results are described in an article currently under submission.

7.1.7. *Application configuration and reconfiguration in multi-cloud environments*
Participant: Nikos Parlavantzas.

---

[2]https://www.elastic.co/fr/
[3]https://www.elastic.co/products/kibana

Current approaches to cloud application configuration and reconfiguration are typically platform dependent, error prone and provide little support for optimizing application performance and resource utilisation. To address these limitations, we are combining the use of software product lines (SPLs) with performance prediction and automatic adaptation techniques. This work is performed in the context of the thesis of Carlos Ruiz Diaz, a PhD student at the University of Guadalajara, co-advised by Nikos Parlavantzas. The work has produced an SPL-based framework supporting initial configuration and dynamic adaptation in a systematic, platform-independent way.

In 2017, we extended the framework with a proactive adaptation solution that performs vertical VM scaling based on predictions of resource utilisation and performance. The solution targets multi-tier applications deployed on IaaS clouds. Experimental results demonstrate that the solution maintains expected application performance while reducing resource waste [46].

7.1.8. *Adaptive resource management for high-performance, multi-sensor systems*
Participants: Christine Morin, Nikos Parlavantzas, Baptiste Goupille–Lescar.

In the context of our collaboration with Thales Research and Technology and Baptiste Goupille-Lescar's PhD work, we are applying cloud resource management techniques to high-performance, multi-sensor, embedded systems with real-time constraints. The objective is to increase the flexibility and efficiency of resource allocation in such systems, enabling the execution of dynamic sets of applications with strict QoS requirements.

In 2017, we focused on an industrial use case concerning the operation of a multi-function surface active electronically scanned array (AESA) radar. We developed a simulation environment using an industrial high-precision AESA simulator and the Ptolemy II simulation framework, and we are using this environment to explore and evaluate different dynamic application placement solutions [57].

7.2. Greening Clouds

ICT (Information and Communications Technologies) ecosystem now approaches 6% of world electricity consumption and this ICT energy use will continue grow fast because of the information appetite of Big Data, big networks and big infrastructures as Clouds that unavoidably leads to big power.

7.2.1. *Energy Models*
Participants: Ehsan Ahvar, Loic Guegan, Anne-Cécile Orgerie, Martin Quinson.

Cloud computing allows users to outsource the computer resources required for their applications instead of using a local installation. It offers on-demand access to the resources through the Internet with a pay-as-you-go pricing model.However, this model hides the electricity cost of running these infrastructures.

The costs of current data centers are mostly driven by their energy consumption (specifically by the air conditioning, computing and networking infrastructure). Yet, current pricing models are usually static and rarely consider the facilities' energy consumption per user. The challenge is to provide a fair and predictable model to attribute the overall energy costs per virtual machine and to increase energy-awareness of users. We aim at proposing such energy cost models without heavily relying on physical wattmeters that may be costly to install and operate.

Another goal consists in better understanding the energy consumption of computing and networking resources of Clouds in order to provide energy cost models for the entire infrastructure including incentivizing cost models for both Cloud providers and energy suppliers. These models will be based on experimental measurement campaigns on heterogeneous devices. Inferring a cost model from energy measurements is an arduous task since simple models are not convincing, as shown in our previous work. We aim at proposing and validating energy cost models for the heterogeneous Cloud infrastructures in one hand, and the energy distribution grid on the other hand. These models will be integrated into simulation frameworks in order to validate our energy-efficient algorithms at larger scale.

7.2.2. Exploiting Renewable Energy in Clouds

Participants: Benjamin Camus, Yunbo Li, Anne-Cécile Orgerie.

The development of IoT (Internet of Things) equipment, the popularization of mobile devices, and emerging wearable devices bring new opportunities for context-aware applications in cloud computing environments. The disruptive potential impact of IoT relies on its pervasiveness: it should constitute an integrated heterogeneous system connecting an unprecedented number of physical objects to the Internet. Among the many challenges raised by IoT, one is currently getting particular attention: making computing resources easily accessible from the connected objects to process the huge amount of data streaming out of them.

While computation offloading to edge cloud infrastructures can be beneficial from a Quality of Service (QoS) point of view, from an energy perspective, it is relying on less energy-efficient resources than centralized Cloud data centers. On the other hand, with the increasing number of applications moving on to the cloud, it may become untenable to meet the increasing energy demand which is already reaching worrying levels. Edge nodes could help to alleviate slightly this energy consumption as they could offload data centers from their overwhelming power load and reduce data movement and network traffic. In particular, as edge cloud infrastructures are smaller in size than centralized data center, they can make a better use of renewable energy.

We propose to investigate the end-to-end energy consumption of IoT platforms. Our aim is to evaluate, on concrete use-cases, the benefits of edge computing platforms for IoT regarding energy consumption. We aim at proposing end-to-end energy models for estimating the consumption when offloading computation from the objects to the edge or to the core Cloud, depending on the number of devices and the desired application QoS, in particular trading-off between performance (response time) and reliability (service accuracy).

7.2.3. Smart Grids

Participants: Benjamin Camus, Anne-Cécile Orgerie, Martin Quinson.

We propose exploiting Smart Grid technologies to come to the rescue of energy-hungry Clouds. Unlike in traditional electrical distribution networks, where power can only be moved and scheduled in very limited ways, Smart Grids dynamically and effectively adapt supply to demand and limit electricity losses (currently 10% of produced energy is lost during transmission and distribution).

For instance, when a user submits a Cloud request (such as a Google search for instance), it is routed to a data center that processes it, computes the answer and sends it back to the user. Google owns several data centers spread across the world and for performance reasons, the center answering the user's request is more likely to be the one closest to the user. However, this data center may be less energy efficient. This request may have consumed less energy, or a different kind of energy (renewable or not), if it had been sent to this further data center. In this case, the response time would have been increased but maybe not noticeably: a different trade-off between quality of service (QoS) and energy-efficiency could have been adopted.

While Clouds come naturally to the rescue of Smart Grids for dealing with this big data issue, little attention has been paid to the benefits that Smart Grids could bring to distributed Clouds. To our knowledge, no previous work has exploited the Smart Grids potential to obtain and control the energy consumption of entire Cloud infrastructures from underlying facilities such as air conditioning equipment (which accounts for 30% to 50% of a data center's electricity bill) to network resources (which are often operated by several actors) and to computing resources (with their heterogeneity and distribution across multiple data centers). We aim at taking advantage of the opportunity brought by the Smart Grids to exploit renewable energy availability and to optimize energy management in distributed Clouds.

7.2.4. Involving Users in Energy Saving

Participants: David Guyon, Christine Morin, Anne-Cécile Orgerie.

In a Cloud moderately loaded, some servers may be turned off when not used for energy saving purpose. Cloud providers can apply resource management strategies to favor idle servers. Some of the existing solutions propose mechanisms to optimize VM scheduling in the Cloud. A common solution is to consolidate the mapping of the VMs in the Cloud by grouping them in a fewer number of servers. The unused servers can then be turned off in order to lower the global electricity consumption.

Indeed, current work focuses on possible levers at the virtual machine suppliers and/or services. However, users are not involved in the choice of using these levers while significant energy savings could be achieved with their help. For example, they might agree to delay slightly the calculation of the response to their applications on the Cloud or accept that it is supported by a remote data center, to save energy or wait for the availability of renewable energy. The VMs are black boxes from the Cloud provider point of view. So, the user is the only one to know the applications running on her VMs.

We plan to explore possible collaborations between virtual machine suppliers, service providers and users of Clouds in order to provide users with ways of participating in the reduction of the Clouds energy consumption. This work will follow two directions: 1) to investigate compromises between power and performance/service quality that cloud providers can offer to their users and to propose them a variety of options adapted to their workload; and 2) to develop mechanisms for each layer of the Cloud software stack to provide users with a quantification of the energy consumed by each of their options as an incentive to become greener.

Our results were published in [40], [32], [31].

7.3. Securing Clouds

7.3.1. *Security Monitoring in Clouds*
Participants: Christine Morin, Jean-Louis Pazat, Louis Rilling, Anna Giannakou, Amir Teshome Wonjiga, Clément El Baz.

In the INDIC project we aim at making security monitoring a dependable service for IaaS cloud customers. To this end, we study three topics:

- defining relevant SLA terms for security monitoring,
- enforcing and verifying SLA terms,
- making the SLA terms enforcement mechanisms self-adaptable to cope with the dynamic nature of clouds.

The considered enforcement and verification mechanisms should have a minimal impact on performance.

In 2017, we did a thorough performance evaluation and security correctness analysis of the SAIDS approach, that we proposed in 2015, and that makes a network intrusion detection system (NIDS) deployed in a cloud operator infrastructure self-adaptable. In the performance evaluation we studied the performance impact of SAIDS on the cloud infrastructure operations related to the management of virtual machines (typically creation, migration, and deletion) as well as the scalability of SAIDS with respect to the number of NIDS devices managed. This performance evaluation was done on the Grid'5000 platform. The results showed that SAIDS adds very low overhead and is scalable. The security analysis was done both experimentally and based on a risk analysis. This analysis validated the security correctness of SAIDS. A full paper presenting SAIDS and its evaluation is submitted for publication in 2018. A demo of SAIDS was presented at FIC 2017, Lille, France in January 2017 and at the Inria Industry Days, Paris, France on October 17th, 2017.

Regarding SLA definition and enforcement, in 2017 we evaluated the verification method that we defined in 2016 and that enables a Cloud customer to verify that an NIDS located in the operator infrastructure is configured correctly according to the Service-Level Objectives (SLO) figuring in the SLA. The performance evaluation was done on the Grid'5000 platform and showed that the proposed verification method requires making a trade-off between verification speed and impact on the performance of the production applications deployed in the tenant's virtual machines. The security correctness analysis was based on a risk analysis and showed the constraints on the types of attacks that can be used for verification as well as the limitations due to the tools used in the prototype [55]. A full paper presenting the verification method and its evaluation is submitted for publication in 2018.

After the acquired experience on verifying security monitoring metrics, we started studying how to define relevant SLOs that are verifiable. We plan to get results in 2018 and submit a paper for publication in 2018 or 2019.

Finally, in October 2017 we started studying how security monitoring SLAs could take into account context changes like the evolution of threats and updates to the tenants' software.

Our work done as part of the INDIC project were presented in [59].

7.3.2. *Risk assessment in clouds*
Participant: Christine Morin.

Cloud providers have an incomplete view of their hosted virtual infrastructures managed by a Cloud Management System (CMS) and a Software Defined Network (SDN) controller. For various security reasons (e.g. isolation verification, modeling attack paths in the network), it is necessary to know which virtual machines can interact via network protocols. This requires building a connectivity graph between the virtual machines, that we can extract with the knowledge of the overall topology and the deployed network security policy. Existing methodologies for building such models for physical networks produce incomplete results. Moreover, they are not suitable for cloud infrastructures due to either their intrusiveness or lack of connectivity discovery. We propose a method to compute the connectivity graph, relying on information provided by both the CMS and the SDN controller. Connectivity can first be extracted from knowledge databases, then dynamically updated on the occurrence of cloud-related events. We realized an experimental evaluation of the proposed method to determine its correctness and performance in a realistic context, considering CPU and RAM consumption, the volume of data generated, and execution time for the different portions of the algorithm involved. Experiments were run on the Grid'5000 platform with OpenStack CMS and ONOS SDN controller. Our approach proves on a representative infrastructure to compute exact, complete and up-to- date connectivity graphs in reasonable time [42], [41].

7.3.3. *Personal Data Management in Cloud-based IoT Systems*
Participants: Christine Morin, Jean-Pierre Banâtre, Deborah Agarwal, Subhadeep Sarkar, Louis Rilling.

The Internet of Things (IoT), in today's digital world, encompasses billions of smart connected devices. These devices generate an unprecedented amount of data, which often bears sensitive personal information of individuals. In present service models, the data are processed and managed by service providers, beyond the visibility of the owner of the data. Although the EU General Data Protection Regulation (GDPR) strives to protect citizens and their data by regulation, citizens and service providers need technological advances to gain effective control over their data or to prove compliance with the new regulation. Our primary objective is to enforce, by design, the GDPR at the system level so as to preserve the privacy concerning personal data. We started off with enforcement of the data erasure facility as expressed in the GDPR. Data erasure corresponds to both automatic erasure of data after expiration of their retention period and ad-hoc on request of the data owner. Our first contribution, towards this, is design of a customizable privacy policy, which would allow the end users to express their preferences regarding the purpose of use, location of processing, retention period, sharing and storage policies concerning their personal data. We developed a XML-based policy expression language by defining the required data structures and vocabulary, which will facilitate the end-users to easily express their preferences. Next, we have investigated into the possible way of the implementation of the proposed solution and identified the exploitation of the operation system capabilities as an appropriate means to the cause. For this, we have potentially chosen the Sel4 (or may be some other capability-based microkernel) as our platform of operation. Finally, we have identified the different challenges towards implementation of our solution and did some groundwork towards proposing the solutions to the same. These challenges include efficient identification of replication of data, locating all replicas of a given data segment, and implementing erasure of data in a cross-domain service model.

7.4. Experimenting with Clouds

7.4.1. *Simulation*
Participants: Martin Quinson, Loic Guegan, Toufik Boubehziz, The Anh Pham.

We propose to combine two complementary experimental approaches: direct execution on testbeds such as Grid'5000, that are eminently believable but rather labor intensive, and simulations (using *e.g.* SimGrid) that are much more light-weighted, but requires are careful assessment. One specificity of the Myriads team is that we are working on these experimental methodologies *per se*, raising the standards of *good experiments* in our community. The Grid'5000 operational team is embedded in our research team, ensuring that our work remains aligned with the ground reality.

In 2017, our work was mostly centered on letting SimGrid become a *de facto* standard for the simulation of distributed platforms. We introduced a new programming interface, particularly adapted to the study of abstract algorithms. Beyond the engineering task, this requires to carefully capture the concepts that are important to the practitioners on distributed systems.

SimGrid is not limited to abstract algorithms, and can also be used to simulate real applications. This year, we published a journal article on the many challenges to overcome when designing a simulator of high performance systems. This work was published in the TPDS journal [20].

On the modeling side, our team worked this year toward the improvement of energy models, both for computational facilities and for the network. Despite the scarce availability of real testbeds that allow fine-grained energy measurements, we managed to provide a generic energy consumption model, published in [35], [43].

Finally, we restarted our efforts toward the formal verification of distributed systems. The model-checker that is integrated within SimGrid is already functional ([44]), but more work is necessary to make it efficient. We even found cases for which our reduction algorithm may miss defects in the verified system. This work will certainly motivate much more work in the future years.

7.4.2. Use cases

Participants: Christine Morin, Nikos Parlavantzas, Deborah Agarwal, Manh Linh Pham.

7.4.2.1. Simulation framework for studying between-herd pathogen spread in a region

In the context of the MIHMES project (2012-2017) and in collaboration with INRA researchers, we transformed a legacy application for simulating the spread of bovine viral diarrhea virus (BVDV) to a cloud-enabled application based on the DiFFuSE framework (Distributed framework for cloud-based epidemic simulations). Specifically, the original sequential code was first modified to add single-computer parallelism using OpenMP. We then decomposed the code into separate services that were deployed across multiple clouds and independently scaled. Using this service-based cloud-enabled simulation, we performed a set of experiments that demonstrated that applying DiFFuSE increases performance, allows exploring different cost-performance trade-offs, automatically handles failures, and supports elastic allocation of resources from multiple clouds [45].

7.4.2.2. FluxNet and AmeriFlux Data Analysis

The carbon flux datasets from AmeriFlux (Americas) and FLUXNET (global) are comprised of long-term time series data and other measurements at each tower site. There are over 800 flux towers around the world collecting this data. The non-time series measurements include information critical to performing analysis on the site's data. Examples include: canopy height, species distribution, soil properties, leaf area, instrument heights, etc. These measurements are reported as a variable group where the value plus information such as method of measurement and other information are reported together. Each variable group has a different number and type of parameters that are reported. The current output format is a normalized file. Users have found this file difficult to use.

Our earlier work in the DALHIS Inria associate team focused on building user interfaces to specify the data. This year we jointly worked on developing a Jupyter Notebook that would serve as a tool for users to read in and explore the data in a personalized tutorial type environment. We developed two notebooks and the next step is to start user testing on the notebooks.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Technicolor (2016-2017)

Participant: Guillaume Pierre.

Our collaboration with Technicolor has focused on the design of a scalable and elastic virtual customer premises equipment based on Network Function Virtualization, Software-Defined Networking and Cloud technologies. In 2017 we completed the system design and started implementing the system. The collaboration completed successfully in June 2017. However, the vCPE project within which this collaboration took place was unfortunately interrupted by Technicolor before we could write a paper about this work.

8.2. Bilateral Grants with Industry

8.2.1. Thales Research and Technology (2016-2018)

Participants: Baptiste Goupille-Lescar, Christine Morin, Nikos Parlavantzas.

Our collaboration with Thales Research and Technology focuses on the development of distributed Cyber-Physical Systems, such as those developed by Thales to monitor and react to changing physical environments. These systems need to be highly adaptable in order to cope with the dynamism and diversity of their operating environments. Notably, they require distributed, parallel architectures that support dynamic sets of applications, not known in advance, while providing strong QoS guarantees. The objective of this collaboration is to explore adaptive resource management mechanisms for such systems that can adapt to changes in the requirements and in the availability of resources. This contract funds Baptiste Goupille-Lescar's PhD grant.

8.2.2. Nokia (2015-2018)

Participant: Christine Morin.

Together with CIDRE Inria project-team we are involved in a collaboration with Nokia on security policy adaptation driven by risk evaluation in modern communication infrastructures. To address the need for efficient security supervision mechanisms, approaches such as attack graphs generation, coupled to a risk-based assessment have been used to provide an insight into a system's threat exposure. In comparison to static infrastructures, clouds exhibit a dynamic nature and are exposed to new attack scenarios due to virtualization. The goal of this collaboration is thus to revisit existing methods in the context of clouds. This contract funds Pernelle Mensah's PhD grant. Pernelle is a member of CIDRE project-team.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. CominLabs EPOC project (2013-2017)

Participants: Sabbir Hasan Rochi, Yunbo Li, Anne-Cécile Orgerie, Jean-Louis Pazat.

In this project, partners aim at focusing on energy-aware task execution from the hardware to application's components in the context of a mono-site data center (all resources are in the same physical location) which is connected to the regular electric Grid and to renewable energy sources (such as windmills or solar cells). In this context, we tackle three major challenges:

- Optimizing the energy consumption of distributed infrastructures and service compositions in the presence of ever more dynamic service applications and ever more stringent availability requirements for services.
- Designing a clever cloud's resource management which takes advantage of renewable energy availability to perform opportunistic tasks, then exploring the trade-off between energy saving and performance aspects in large-scale distributed systems.
- Investigating energy-aware optical ultra high-speed interconnection networks to exchange large volumes of data (VM memory and storage) over very short periods of time.

Sabbir Hasan Rochi has defended his PhD on SLA driven Cloud autoscaling for optimizing energy footprint on May 3rd, 2017. Yunbo Li has defended his PhD on Resource allocation in a Cloud partially powered by renewable sources on June 12th, 2017.

9.1.2. INDIC - Cybersecurity Pole of Excellence (2014-2020)

Participants: Anna Giannakou, Christine Morin, Jean-Louis Pazat, Louis Rilling, Amir Teshome Wonjiga, Clément El Baz.

Our study carried out in the framework of a collaboration with DGA-MI aims at defining and enforcing SLA for security monitoring of virtualized information systems. To this aim we study three topics:

- defining relevant SLA terms for security monitoring,
- enforcing and evaluating SLA terms,
- making the SLA terms enforcement mechanisms self-adaptable to cope with the dynamic nature of clouds.

The considered enforcement and evaluation mechanisms should have a minimal impact on performance. The funding from DGA funds two PhD students: Anna Giannakou and Amir Teshome Wonjiga. Clément El Baz is partially funded by the Brittany Regional Council in the PEC framework.

9.2. National Initiatives

9.2.1. ADEME RennesGrid

Participants: Benjamin Camus, Anne-Cécile Orgerie, Martin Quinson.

The aim of the RennesGrid project is to design and implement a large-scale preindustrial microgrid demonstrator in the territory of Rennes Metropole to organize the shared self-consumption of a group of photovoltaic panels coupled to stationary storage devices. Traditional approaches to power grid management tend to overlook the costs, both energy and economic, of using computers to ensure optimal electricity network management. However, these costs can be significant. It is therefore necessary to take them into account along with the design of IT tools during studies of optimal energy management of smart grids. In addition, telecommunication networks are generally considered to have an ideal functioning, that is to say they can not negatively affect the performance of the electricity network. However, this is not realistic and it is necessary to analyze the impact of phenomena such as congestion, latency, failures related to computer equipment or impact on the batteries of sensors, etc. on strategies for optimal management of the electricity network. In this project, we will closely collaborate with Anne Blavette (CR CNRS in electrical engineering, SATIE, Rennes) and co-supervise a post-doc on evaluating the impact of the IT infrastructure in the management of smart grids.

9.2.2. Inria ADT SaaP (2016-2018)

Participants: Toufik Boubehziz, Martin Quinson.

The SaaP technological development action (SimGrid As A Platform) funded by INRIA targets the refactoring of SimGrid to make it ready to use in production and teaching contexts. Our ultimate goal is to sustain the development of the framework by involving 5 to 10 companies that are using it internally. Our target of the teaching context is thus an intermediate goal, as we think that the best solution to ensure the adoption of our tool by the industrial engineers is that they discover the tool during their studies.

The technical actions envisioned for this ADT are the complete re-factoring of the software (to make it easier to script a new model within the tool kernel) and a reorganization of the interfaces (for a better integration in the Java and python language). This work is lead by Toufik Boubehziz in collaboration with the whole SimGrid community, which provide valuable feedback.

9.2.3. Inria ADT DiFFuSE (2017-2018)

Participants: Nikos Parlavantzas, Christine Morin, Manh Linh Pham.

The DiFFuSE technological development action (Distributed framework for cloud-based epidemic simulations) funded by INRIA focuses on the DiFFuSE framework developed by Myriads in the context of MIHMES (2012-2017). MIHMES was a 5-year collaborative multidisciplinary project funded by ANR under the Investments for the Future Program, and led by BIOEPAR, INRA, ONIRIS. DiFFuSE is a framework that provides design support, reusable code, and tools for building and executing epidemic simulations in the cloud. The main objectives of this ADT are to improve the usability and robustness of DiFFuSE, to provide support to scientists for applying the framework to a new epidemic simulations as well as to provide a thorough evaluation of the framework using multiple case studies.

9.2.4. *Inria IPL Discovery (2015-2019)*

Participants: Ehsan Ahvar, Anne-Cécile Orgerie, Matthieu Simonin, Genc Tato, Cédric Tedeschi.

The Inria IPL Discovery officially started in September 2015. It targets the design, development and deployment of a distributed Cloud infrastructure within the network's backbone. It will be based upon a set of building blocks whose design will take locality as a primary constraint, so as to minimize distant communications and consequently achieve better network traffic, partition management and improved availability.

Its developments are planned to get integrated within the OpenStack framework. Myriads is involved in the design of new overlay networks for such environments so as to support efficient messaging and routing. Myriads is also involved in the energy/cost benefit analysis of distributed edge-cloud architectures.

9.2.5. *Inria IPL CityLab (2015-2018)*

Participants: Subarna Chatterjee, Christine Morin.

The Inria Project Lab (IPL) CityLab@Inria (http://citylab.inria.fr) studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. City-scale experiments of the proposed platforms and services are planned in cities in California and France, thereby learning lessons from diverse setups.

Myriads investigates advanced cloud solutions for the Future Internet, which are critical for the processing of urban data. It leverages its experience in cloud computing and Internet of services while expanding its research activities to the design and implementation of cloud services to support crowd-Xing applications and mobile social applications.

In 2017, Christine Morin was involved in the preparation of a SPOC entitled "Technological challenges of participatory smart cities", which is proposed in the framework of the EIT Digital professional school. She prepared seven sequences on cloud-based urban data management. This SPOC is the English version of the MOOC entitled *"Défis technologiques des villes intelligentes participatives"* run on the FUN platform in Spring and Fall 2017.

In 2017, we also conducted a comparative experimental evaluation of data stream processing environments executed on clusters and clouds. We compared the performance and energy consumption of Heron, Storm and Flink frameworks with three data streaming representative applications.

9.2.6. *Inria IPL Hac Specis (2016-2020)*

Participants: Anne-Cécile Orgerie, Martin Quinson, The Anh Pham.

The goal of the HAC SPECIS (High-performance Application and Computers: Studying PErformance and Correctness In Simulation) project (http://hacspecis.gforge.inria.fr/) is to answer methodological needs of HPC application and runtime developers and to allow to study real HPC systems both from the correctness and performance point of view. To this end, we gather experts from the HPC, formal verification and performance evaluation community.

During his first year of PhD thesis, The Anh Pham conducted an analysis of the formal methods and algorithms used in SimGrid. This work, co-advised by Martin Quinson with Thierry Jéron (team SUMO, formal methods), was important to bridge the gap between the involved communities. The resulting work has been published in a workshop gathering the intersection between the communities of formal methods and HPC [44].

Another PhD thesis will stat in December 2017, co-advised by Laurent Lefèvre (Avalon team, Lyon), Martin Quinson and Anne-Cécile Orgerie. This thesis will focus on simulating the energy consumption of continuum computing between heterogeneous numerical infrastructures for HPC.

9.2.7. COSMIC PRE (2016 - 2018)

Participants: Benjamin Camus, Anne-Cécile Orgerie, Martin Quinson.

The distributed nature of Cloud infrastructures involves that their components are spread across wide areas, interconnected through different networks, and powered by diverse energy sources and providers, making overall energy monitoring and optimization challenging. The COSMIC project aims at taking advantage of the opportunity brought by the Smart Grids to exploit renewable energy availability and to optimize energy management in distributed Clouds. This PRE, led by Anne-Cécile Orgerie also involves Fanny Dufossé from Dolphin team (Inria Lille), Anne Blavette from SATIE laboratory (electrical engineering, Rennes), and Benjamin Camus, who has started a 18 months post-doc in October 2016 in the context of this project. A paper on this project has been presented at SMARTGREENS 2017 and two others are currently under submission.

9.2.8. MIHMES ANR Investissements d'Avenir (2012 - 2017)

Participants: Christine Morin, Manh Linh Pham, Nikos Parlavantzas.

The MIMHES project (http://www.inra.fr/mihmes) led by INRA/BioEpAR aimed at producing scientific knowledge and methods for the management of endemic infectious animal diseases and veterinary public health risks. The role of Myriads was to help MIHMES researchers improve the performance of their simulation applications and take advantage of computing resources provided by clouds. To that end, Myriads developed a framework, named DiFFuSE, that provides design support, reusable code, and tools for building and executing epidemic simulations in the cloud.

In 2017, we further developed DiFFuSE and extended the framework to make use of the PaaSage open-source platform, the main outcome of a European FP7 IP project in which Myriads participated (2012-2016). Thanks to PaaSage, DiFFuSE allows deploying and managing services in multi-cloud environments. We applied DiFFuSE to restructure an application that simulates the spread of the bovine viral diarrhea virus (BVDV) and conducted experiments to evaluate DiFFuSE [45].

9.2.9. SESAME ASTRID project (2016-2019)

Participants: Pascal Morillon, Christine Morin, Matthieu Simonin, Cédric Tedeschi, Mehdi Belkhiria.

The Sesame project (http://www.agence-nationale-recherche.fr/Project-ANR-16-ASTR-0026) led by IMT Atlantique aims at develop efficient infrastructures and tools for the maritime traffic surveillance. The role of Myriads is to define a robust and scalable infrastructure for the real-time and batch processing of vessel tracking information.

9.2.10. PIA ELCI (2015-2018)

Participant: Anne-Cécile Orgerie.

The PIA ELCI project deals with software environment for computation-intensive applications. It is leaded by BULL. In the context of this project, we collaborate with ROMA and Avalon teams from Lyon: we co-supervise a PhD student (Issam Rais) funded by this project with these teams on multi-criteria scheduling for large-scale HPC environments. This collaboration has led to two publications in 2017: a journal article published in IJHPCA and a conference paper presented at EuroPar.

9.2.11. CNRS GDS EcoInfo

Participant: Anne-Cécile Orgerie.

The EcoInfo group deals with reducing environmental and societal impacts of Information and Communications Technologies from hardware to software aspects. This group aims at providing critical studies, lifecycle analyses and best practices in order to improve the energy efficiency of printers, servers, data centers, and any ICT equipment in use in public research organizations. In particular, it has led in December 2016 to the publication of an ADEME report jointly with Deloitte Développement Durable, Futuribles, CREDOC and ADEME on the potential contribution of digital to the reduction of environmental impacts: state of play and challenges for the prospective.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. FogGuru
Participant: Guillaume Pierre.

Title: MSCA ITN EID

Programm: H2020

Duration: September 2017 - August 2021

Coordinator: Guillaume Pierre

Participants:

University of Rennes 1, France (coordinator)

Technisch Universität Berliun, Germany

Elastisys AB, Sweden

U-Hopper srl, Italy

EIT Digital Rennes, France

Las Naves, Spain

FogGuru is a doctoral training project which aims to to train eight talented PhD students with an innovative and inter-sectoral research program to constitute the next generation of European Cloud and Fog computing experts. Besides their scientific and technical education, FogGuru's PhD students will receive extensive training in technological innovation and entrepreneurship as well as soft skills. These combined skills will enable them to fully master the innovation process stemming from fundamental research towards invention and development of innovative products and services, and to real-life deployment, experimentation and engagement with beta-testers.

9.3.1.2. Fed4Fire+ (2017-2022)
Participants: David Margery, Yue Li.

Title: Federation for FIRE Plus

Programm: H2020

Duration: January 2017 - December 2021

Coordinator: Interuniversitair Micro-Electronicacentrum Imec VZW

Partners:

Universidad de Malaga

National Technical University of Athens - NTUA

The Provost, Fellows, Foundation Scholars & the other members of board of the College of the Holy & Undivided Trinity of Queen Elizabeth Near Dublin

Ethniko Kentro Erevnas Kai Technologikis Anaptyxis

GEANT LImited

Institut Jozef Stefan

Mandat International Alias Fondation Pour la Cooperation Internationale

Université Pierre et Marie Curie - Paris 6

Universidad De Cantabria

Fundacio Privada I2CAT, Internet I Innovacio Digital A Catalunya

EURESCOM-European Institute For Research And Strategic Studies in Telecommunications GMBH

Nordunet A/S

Technische Universitaet Berlin

Instytut Chemii Bioorganicznej Polskiej Akademii Nauk

Fraunhofer Gesellschaft zur Foerderung Der Angewandten Forschung E.V.

Universiteit Van Amsterdam

University of Southampton

Martel GMBH

Atos Spain SA

Institut National de Recherche en Informatique et automatique

Inria contact: David Margery

Fed4FIRE+ is a successor project to Fed4FIRE. In Fed4FIRE+, we more directly integrate Grid'5000 into the wider eco-system of experimental platforms in Europe and beyond using results we developed in Fed4FIRE. We have developed a generalized proxy mechanisms to allow users with Fed4FIRE identities to interact with services giving access to different testbeds but not designed to support Fed4FIRE identities. Fed4FIRE+ has started January 1st, 2017.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

9.3.2.1. NESUS

Participant: Anne-Cécile Orgerie.

Program: ICT COST

Project acronym: NESUS

Project title: Network for Sustainable Ultrascale Computing (ICT COST Action IC1305)

Duration: 2014 - 2018

Coordinator: Prof. Jesus Carretero, University Carlos III of Madrid, Spain, http://www.nesus.eu

Other partners: 33 COST countries and 11 non-COST countries

Abstract: Ultrascale systems are envisioned as large-scale complex systems joining parallel and distributed computing systems that will be two to three orders of magnitude larger that today's systems. The EU is already funding large scale computing systems research, but it is not coordinated across researchers, leading to duplications and inefficiencies. The goal of the NESUS Action is to establish an open European research network targeting sustainable solutions for ultrascale computing aiming at cross fertilization among HPC, large scale distributed systems, and big data management. The network will contribute to gluing disparate researchers working across different areas and provide a meeting ground for researchers in these separate areas to exchange ideas, to identify synergies, and to pursue common activities in research topics such as sustainable software solutions (applications and system software stack), data management, energy efficiency, and resilience. Some of the most active research groups of the world in this area are members of this proposal. This Action will increase the value of these groups at the European-level by reducing duplication of efforts and providing a more holistic view to all researchers, it will promote the leadership of Europe, and it will increase their impact on science, economy, and society. Anne-Cécile Orgerie is co-responsible of the focus group on metrics, monitoring, instrumentation and profiling in the Working Group 5 on Energy Efficiency. A joint paper has been accepted in 2017 on this topic at the Elsevier journal on Sustainable Computing.

9.3.3. Collaborations with Major European Organizations

Partner 1: EPFL, Network architecture lab (Switzerland)

We collaborate with Katerina Argyraki's research group on the integration of networking and cloud computing technologies in order to support placement constraints between cloud resources.

Partner 2: University of Neuchâtel, dept. of Computer Science (Switzerland)

We collaborate with Pascal Felber's research group on energy efficiency in Clouds and in particular on the design of energy cost models for virtual machines. A joint journal paper has been accepted in 2017 for publication in Sustainable Computing: Informatics and Systems, Elsevier.

Partner 3: Catholic University of Louvain (Belgium)

We collaborate with Etienne Riviere's research group on the efficient service placement and discovery in a SaaS context.

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. DALHIS

Participants: Christine Morin, Deborah Agarwal, Anna Giannakou, Amir Teshome Wonjiga, Subarna Chatterjee.

Title: Data Analysis on Large Heterogeneous Infrastructures for Science

International Partner (Institution - Laboratory - Researcher):

Lawrence Berkeley National Laboratory (United States) - Data Science and Technology department - Deb Agarwal

Start year: 2016

See also: https://project.inria.fr/dalhis/

Data produced by scientific instruments (large facilities like telescopes or field data), large-scale experiments, and high-fidelity simulations are increasing in magnitude and complexity. Existing data analysis methods, tools and infrastructure are often difficult to use and unable to provide the complete data management, collaboration, and curation environment needed to manage these complex, dynamic, and large-scale data analysis environments. The goal of the Inria-LBL DALHIS associate team involving the Myriads (PI) and Avalon Inria project-teams and the Data Science and Technology (DST) department at Lawrence Berkeley National Laboratory (LBL) is to create a collaborative distributed software ecosystem to manage data lifecycle and enable data analytics on distributed data sets and resources. Specifically, our goal is to build a dynamic software stack that is user-friendly, scalable, energy-efficient and fault tolerant. Our research will determine appropriate execution environments that allow users to seamlessly execute their end-to-end dynamic data analysis workflows in various resource environments and scales while meeting energy-efficiency, performance and fault tolerance goals. We will engage in deep partnerships with scientific teams (Fluxnet in environmental science and SNFactory and LSST experiences in cosmology) and use a mix of user research with system software R&D to address specific challenges that these communities face. In 2017, we worked on evaluating data streaming environments (see Section 7.1.4) and on producing tools to help users (scientist in the climate and environment community) to explore the carbon flux datasets from AmeriFlux (Americas) and FLUXNET (global) (see Section 7.4.2.2). We also worked on two facets of security in the context of HPC distributed computing infrastructures: (i) building a workflow for data analysis for anomaly detection and (ii) using the block-chain technology to leverage data integrity at the network and some portion of computation levels.

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

9.4.2.1. SUSTAM associated team

Participants: Anne-Cécile Orgerie, Yunbo Li.

Anne-Cécile Orgerie participates in the associated team named SUSTAM (Sustainable Ultra Scale compuTing, dAta and energy Management) leaded by Laurent Lefévre (Avalon team, Lyon) with Prof. Manish Parashar (RDI2, Rutgers University, NJ, USA). The SUSTAM associate team will focus on the joint design of a multi-criteria orchestration framework dealing with resources, data and energy management in an sustainable way.

9.4.3. Inria International Partners

9.4.3.1. Informal International Partners

Partner: Rutgers University, dept. of Computer Science (New Jersey, United States)

We collaborate with Manish Parashar's research group on energy efficiency in edge Clouds and in particular on the design of energy cost models for such environments involving renewable energy. A joint paper has been presented at IEEE/ACM CCGrid 2017.

Partner: Northeastern University, dept. of Computer Science (Massachusetts, United States)

We collaborate with Gene Cooperman 's research group on the study of large-scale distributed systems. More specifically, we actively collaborate on virtualization technologies and system snapshoting (we obtained a postdoc funding on that topic from the Brittany Regional Council, but all applicants declined in the last minute). We plan to reinforce and extend our collaboration to formal methods for distributed systems in the next year.

Partner: University of Guadalajara (Mexico)

We collaborate with the team of Prof. Hector Duran-Limon on application and resource management in the cloud. In 2017, we produced a joint publication [46]. Nikos Parlavantzas is co-advising a PhD student enrolled in the University of Guadalajara (Carlos Ruiz Diaz).

Partner: Tlemcen University (Algeria)

We collaborate with Djawida Dib on energy-efficient fault-tolerant resource and application management in containerized clouds. Christine Morin and Nikos Parlavantzas have been co-advising Yasmina Bouizem, a PhD student enrolled in the University of Tlemcen from December 2016.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Deb Agarwal, senior scientist at Lawrence Berkeley National Laboratory, who has been awarded an Inria International Chair for the 2015-2019 period, visited Myriads team during two months from August 1st to October 15th, 2017.

Professor Gene Cooperman, Northeastern University, Boston, USA, visited the Myriads team for one week in July to reinforce our collaboration on the virtualization of large-scale distributed systems.

9.5.1.1. Internships

Betsegaw Lemma Amersho

Date: Feb-June 2017

Institution: University of Rennes 1 & Aalto University (Finland)

Supervisors: Anne-Cécile Orgerie and Martin Quinson

Vinothkumar Nagasayanan

Date: May-August 2017

Institution: University of Rennes 1 & TU Berlin (Germany)

Supervisor: Guillaume Pierre

Salsabil Amri

Date: May-August 2017

Institution: University of Rennes 1

Supervisor: Guillaume Pierre

Bérenger Nguyen Nhon

Date: July-August 2017

Institution: University of Rennes 1

Supervisor: Guillaume Pierre

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

- Genc Tato started a 6-month research visit at the Catholic University of Louvain in November 2017 to work with Etienne Riviere on service placement and discovery in a SaaS context.

- Amir Teshome Wonjiga did a 3-month research internship in the Data Science and Technology department of the Lawrence Berkeley National Laboratory from April to June 2017. He worked with Sean Peisert, staff scientist, on ensuring data integrity in the workflow of high performance applications.

- Anne-Cécile Orgerie visited for 1 week the team of Prof. Manish Parashar in the RDI2 laboratory at Rutgers University in October 2017.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Martin Quinson was the main organizer of the research school SUD gathering the SimGrid user community, held on 20-24 November, in Le Bono, Brittany.

- Guillaume Pierre was a co-organizer of the NetV IRISA/Technicolor Workshop on Network Virtualization (Rennes, February 1st 2017)

10.1.1.2. Member of the Organizing Committees

Guillaume Pierre is general chair for the ACM/IFIP/Usenix Middleware 2018 conference. Christine Morin and Benjamin Camus are respectively sponsor and publicity & web chairs for this conference.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- Anne-Cécile Orgerie was Program vice co-chair for GreenCom 2017: IEEE International Conference on Green Computing and Communications.

- Anne-Cécile Orgerie was co-chair of the Track on Performance Modeling and Evaluation for ICA3PP 2017: International Conference on Algorithms and Architectures for Parallel Processing.

10.1.2.2. Member of the Conference Program Committees

- Christine Morin served as program committee member for the following conferences: ACM International Conference on High Performance Distributed Systems (**HPDC**) 2017, IEEE International Conference on Distributed Computing Systems (**ICDCS**) 2017, ACM/IEEE **SC** 2017, IEEE **Big Data** 2017, IEEE **CloudCom** 2017, IEEE/ACM *International Symposium on Cluster, Cloud and Grid Computing* (**CC-GRID**) 2017, *IEEE Conference on Network unction Virtualization and Software Defined Networks* (**NFV-SDN**) 2017, **ComPas** 2017 and workshop: **Resilience** 2017.

- Anne-Cécile Orgerie served as a program committee member for the following conferences: PDP 2017, ICC 2017, I-SPAN 2017, Compas 2017, CBDCom 2017, SC'17, SD3C 2017, and ISPA 2017.

- Guillaume Pierre served as a program committee member of ACM SAC 2017, ISFEC 2017, ISPDC 2017, CloudCom 2017, and HotWeb 2017.
- Nikos Parlavantzas served as a program committee member of ISPDC 2017, VHPC 17, and CLOSER 2017.
- Cédric Tedeschi served as a program committee member for ICCS 2017, Closer 2017 and Compas 2017.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Christine Morin is associate editor in the IEEE's Transactions on Parallel and Distributed Systems' Editorial Board.

10.1.3.2. Reviewer - Reviewing Activities

- Anne-Cécile Orgerie served as a reviewer for SIMPAT Simulation Modelling Practice and Theory in 2017.
- Cédric Tedeschi served as a reviewer for TPDS in 2017.

10.1.4. Invited Talks

- Anne-Cécile Orgerie gave a talk for the seminar organized at CEA *maison de la simulation* in December 2017 on modeling and simulating the energy consumption of HPC infrastructures.
- Anne-Cécile Orgerie gave a talk for the Inria Scientific Days in June 2017 on measuring the energy consumption and improving the energy efficiency of computing infrastructures.
- Anne-Cécile Orgerie gave a talk for a seminar organized by ENS Rennes for high school students in March 2017 on greening ICT.
- Anne-Cécile Orgerie gave a talk for a seminar organized by the Seine-et-Marne DSDEN during a training day for teachers of primary, secondary and high schools on ICT and sustainability in March 2017 on the energy consumption of ICT.
- Anne-Cécile Orgerie gave a talk for a public conference organized by ArmorScience, an association based in Lannion, in February 2017 on reducing the energy impacts of ICT.
- Anne-Cécile Orgerie gave a talk for a seminar at ENSSAT in Lannion in the Informatics department in February 2017 on saving energy in large-scale distributed systems.
- Anne-Cécile Orgerie gave a talk for the IRISA-Technicolor Workshop on Network Virtualization on February 2017 on greening the networks.
- Guillaume Pierre gave invited talks at the RESCOM summer school on virtualization and containerization (Le Croisic, France, June 2017), the 11th Cloud Control Workshop (Sweden, June 2017), in a seminar of the DIMA team at TU Berlin (November 2017), and a guest lecture at INSA Rennes (France, November 2017).

10.1.5. Leadership within the Scientific Community

- Anne-Cécile Orgerie is co-responsible for the Green axis of the CNRS GDR RSD (Network and Distributed Systems working group).
- Anne-Cécile Orgerie is secretary of the ASF: the French chapter of ACM SIGOPS.
- Cédric Tedeschi is a member of the steering committee of the Compas conference.

10.1.6. Scientific Expertise

Christine Morin was a member of the junior researcher selection committee at Inria Rennes Bretagne Atlantique.

Christine Morin was a member of the selection committee for an assistant-professor position at the University of Lyon.

Guillaume Pierre and Martin Quinson were members of the selection committee for a professor position at the University of Rennes 1.

Anne-Cécile Orgerie was the coordinator of the scientific challenge on "Digitizing Energy" for the Inria's scientific strategic plan 2018-2022.

Anne-Cécile Orgerie was a member of the selection committee for two assistant professor positions at IMT-Atlantique in Nantes.

Anne-Cécile Orgerie was a member of the admission jury for the *second concours* to recruit normalien students at Ecole Normale Supérieure of Rennes.

Jean-Louis Pazat is the coordinator of experts in Information Technology for the evaluation of international bilateral collaborations at the ministry of research and education.

Nikos Parlavantzas acted as an expert reviewer for ANR and SNSF projects.

10.1.7. Research Administration

- Christine Morin is a member of the board of the Project-Team Committee of Inria Rennes Bretagne Atlantique.
- Christine Morin is a member of the University of Rennes 1 board of directors and of the International Affairs Commission and its board.
- Anne-Cécile Orgerie is officer (chargée de mission) for the IRISA cross-cutting axis on Green IT.
- Martin Quinson is the leader of the "Large Scale Systems" department of IRISA.
- Louis Rilling is a member of the scientific board of the research component ("club recherche") of "Pôle d'Excellence Cyber".

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Christine Morin is responsible for the "Advanced Cloud Infrastructures" (ACI) teaching unit of the Master in Computer Science (SIF) of the University of Rennes 1 and of the EIT Digital Master School at the University of Rennes 1.

Christine Morin:

- Master 1: FreeRoom project (spanning the two semesters) co-supervised with Cédric Tedeschi, Adrien Capaine, and Paul Couderc, University of Rennes 1, France.
- Master 2: Advanced Computing Infrastructures" (ACI), 15 hours ETD, EIT Digital Master School, University of Rennes 1, France.

Anne-Cécile Orgerie:

- Master 1: Cloud & Big Data - 12 hours of lecture and 12 hours of practical sessions at ENS Rennes
- Master 2: Green ICT - 3 hours of invited lecture at Telecom SudParis
- Master 2: Green ICT - 15 hours of lecture at IMT-Atlantique Nantes

Guillaume Pierre (at the University of Rennes 1):

- License 3 MIAGE: Systèmes (25 hours ETD)
- License 3 informatique: Organisation et utilisation des systèmes d'exploitation 2 (67 hours ETD)
- Master 1: Service Technologies (24 hours ETD)
- Master 1: Approche algorithmique des applications et systèmes répartis (32 hours ETD)
- Master 2: Techniques de développement logiciel dans le Cloud (39 hours ETD)
- Master 2: Advanced Cloud Infrastructures (10 hours ETD)

Martin Quinson (at ENS Rennes):

- Licence 3: Programming and Software Engineering (30 hours ETD); ARCSYS - architecture et systèmes (60 hours ETD); Pedagogy (15 hours ETD).
- Agrégation Science Industrielle: Programming and Software Engineering (20 hours ETD); Operating Systems and C programming (20 hours ETD); Networking (20 hours ETD).
- Master 2: Pedagogy and Scientific Mediation for Computer Science (30 hours EDT)

Jean-Louis Pazat (at INSA Rennes):

- Licence 3: Parallel and Multicore Programming (36 hours ETD)
- Master 1: Parallel and Distributed Programming (36 hours ETD)
- Master 1: Scientific Parallel Programming (36 hours ETD)
- Licence 3: Networks and SOA (20 hours ETD)
- Master1 / Master 2: Introduction to IoT (24 hours ETD)
- Master2: Student Project in Cloud Computing (12 hours ETD)

Nikos Parlavantzas (at INSA Rennes):

- Master 2: Project in Large-Scale Systems (26 hours ETD)
- Master 1: Clouds (20 hours ETD)
- Master 1: Performance Evaluation (32 hours ETD)
- Master1/2: Introduction to IoT (8 hours ETD)
- Master 1: Operating Systems (36 hours ETD)
- Master 1: Big Data and Applications (15 hours ETD)
- Master 1: Parallel and Distributed Programming (12 hours ETD)
- Licence 3: Multicore Programming (12 hours ETD)
- Licence 3: Networks and SOA (14 hours ETD)
- Master 1: Software Development Project (30 hours ETD)

Cédric Tedeschi (220 hours ETD at Univ. Rennes 1):

- Master 1: Cooperation and concurrency in Systems and Networks
- Master 1: Parallel programming
- Master 2: Internet of Services and Infrastructures

  E-learning

    MOOC : Christine Morin, *Défis technologiques des villes intelligentes participatives*, 1 week in March 2017 and 1 week in November 2017, FUN (http://www.fun-mooc.fr/courses/inria/41009S02/session02/about), Inria MOOC-Lab, leaders, engineers, students, continuing education, respectively 3857 and 3022 registered participants for the March and November sessions.

    SPOC : Christine Morin, Technological challenges of participatory smart cities, 1 week, eitdigitalx, EIT Digital professional school, leaders and engineers, continuing education, unknown number of registered participants

    Pedagogical resources : Christine Morin, cloud-based urban data management , video and quizz, Master, https://www.eitdigital.eu/news-events/news/article/eit-digital-professional-school-launches-a-smart-cities-course-for-companies-and-cities/

10.2.2. Supervision

HdR: Cédric Tedeschi, Distributed Chemically-Inspired Runtimes for Large Scale Adaptive Computing Platforms, defended on April 11th, 2017.

PhD: Anna Giannakou, Self-adaptable security monitoring for IaaS Cloud environments, defended on July 6th, 2017, Louis Rilling, Christine Morin, Jean-Louis Pazat.

PhD: Yunbo Li, Resource allocation in a Cloud partially powered by renewable energy sources, defended on June 12th, 2017, Anne-Cécile Orgerie, Jean-Marc Menaud (Ascola).

PhD: Sabbir Hasan, SLA Driven Cloud autoscaling for optimizing energy footprint, defended on May 3rd, 2017, Thomas Ledoux (Ascola), Jean-Louis Pazat.

PhD: Ismael Cuadrado Cordero, Energy-efficient and network-aware resource allocation in Cloud infrastructures, defended on February 9th, 2017, Christine Morin, Anne-Cécile Orgerie.

PhD in progress: Mehdi Belkhiria, Dynamic Stream Processing for Maritime Traffic Surveillance, started in December 2017, advised by Cédric Tedeschi and Christine Morin

PhD in progress: Ali Jawad Fahs, "Decentralized Fog Computing Infrastructure Control", started in October 2017, Guillaume Pierre.

PhD in progress: Clément El Baz, Reactive security monitoring in clouds, started in October 2017, Louis Rilling, Christine Morin.

PhD in progress: Loic Guegan, Simulating Internet of Things, started in October 2017, advised by Martin Quinson and Anne-Cécile Orgerie

PhD in progress (*co-tutelle*): Yasmina Bouizem, Energy-efficient, fault-tolerance mechanisms for containerized cloud applications, started in November 2016, Didi Fedoua (Tlemcen University, Algeria), Djawida Dib (Tlemcen University, Algeria), Christine Morin, Nikos Parlavantzas.

PhD in progress: The Anh Pham, Dynamic Formal Verification of High Performance Runtimes and Applications, started in November 2016, Martin Quinson, Thierry Jéron.

PhD in progress: Arif Ahmed, "Scalable Decentralized Edge Cloud Infrastructures", started in October 2016, Guillaume Pierre.

PhD in progress: Baptiste Goupille-Lescar, Designing agile, distributed cyber-physical systems with advanced collaboration capabilities, started in January 2016, Eric Lenormand (Thales), Christine Morin, Nikos Parlavantzas.

PhD in progress: Pernelle Mensah, Security policy adaptation driven by risk evaluation in modern communication infrastructures, started in December 2015, Samuel Dubus (Alcatel-Lucent), Christine Morin, Guillaume Piolle (Cidre), Eric Totel (Cidre).

PhD in progress: Genc Tato, Locality-aware Lazy Overlay Networks for WANS, started in December 2015, Marin Bertier, Cédric Tedeschi.

PhD in progress: Amir Teshome, Definition and enforcement of Service-Level Agreements for Cloud security monitoring, started in October 2015, Louis Rilling, Christine Morin.

PhD in progress: Issam Rais, Multi criteria scheduling for large scale High Performance Computing environments, started in October 2015, Anne-Cécile Orgerie, Anne Benoit (ROMA), Laurent Lefèvre (Avalon).

PhD in progress: David Guyon, Supporting energy-awareness for cloud users, started in September 2015, Anne-Cécile Orgerie, Christine Morin.

PhD in progress: Bruno Stevant, Resource allocation strategies for service distribution at the Internet edge to optimize end-to-end latency, started in December 2014 (part-time), Jean-Louis Pazat.

10.2.3. Juries

- Christine Morin was a member in the HdR committee of Cédric Tedeschi, University of Rennes 1, April 11th, 2017.

- Christine Morin chaired the PhD committee of Sabbir Hasan, Insa Rennes, May 3rd, 2017.
- Christine Morin was an external reviewer in the PhD committee of Jiajun Cao, Northeastern University, September 2017.
- Christine Morin was the external reviewer in the PhD committee of Faiza Samreen, Lancaster University, November 23rd, 2017.
- Christine Morin chaired the PhD committee of Pedro Silva, ENS Lyon, December 11th, 2017.
- Anne-Cécile Orgerie was a member in the PhD committee of Millian Poquet, Université de Grenoble, December 19th, 2017.
- Anne-Cécile Orgerie was a member in the PhD committee of Inès de Courchelle, Université de Toulouse 3, November 20th, 2017.
- Jean-Louis Pazat chaired the PhD committee of Jean-Baptiste Barreau, INSA de Rennes, July 10, 2017.
- Jean-Louis Pazat was an external reviewer in the PhD committee of Riadh Karchoud, December 14, 2017.
- Guillaume Pierre was a member in the PhD committee of George Ioannidis, Ecole Polytechnique Fédérale de Lausanne, December 4th, 2017.

10.3. Popularization

Martin Quinson is a member of the ISO national working group (Info Sans Ordi – CS without computer), that aims at inventing, improving and spread unplug activities. These outreach activities are meant to introduce fundamental concepts of Computer Science in schools without the need of any electrical devices. In 2017, the group produced a special issue to the Tangente magazine (read by maths teachers in France). Martin Quinson authored two articles in this special issue.

Martin Quinson is on the scientific board of the Blaise Pascal foundation, boosting the scientific outreach in the domains of Maths and Computer Science. He participated this year to a working group aiming at the development of an online game that teaches maths to pupils.

Martin Quinson was in the scientific organizing committees for several outreach events toward the popularization of Maths and Computer Science this year: «Maths Vivantes» had a very general audience: a booth on a public space of Rennes for one afternoon (March). «Maths C2+» invited about 20 pupils invited at ENS Rennes for 3 days (March). «Metier en tout Genre» invited several dozen of high-school students for one day on the Rennes university campus (November).

11. Bibliography

Major publications by the team in recent years

[1] S. COSTACHE, S. KORTAS, C. MORIN, N. PARLAVANTZAS. *Market-based Autonomous Resource and Application Management in Private Clouds*, in "Journal of Parallel and Distributed Computing", February 2017, vol. 100, pp. 85-102 [*DOI :* 10.1016/J.JPDC.2016.10.003], https://hal.archives-ouvertes.fr/hal-01378536

[2] A. DEGOMME, A. LEGRAND, G. MARKOMANOLIS, M. QUINSON, M. L. STILLWELL, F. SUTER. *Simulating MPI applications: the SMPI approach*, in "IEEE Transactions on Parallel and Distributed Systems", February 2017, 14 p. [*DOI :* 10.1109/TPDS.2017.2669305], https://hal.inria.fr/hal-01415484

[3] E. FELLER, L. RAMAKRISHNAN, C. MORIN. *Performance and energy efficiency of big data applications in cloud environments: A Hadoop case study*, in "Journal of Parallel and Distributed Computing", 2015, vol. 79-80, pp. 80-89 [*DOI :* 10.1016/J.JPDC.2015.01.001], https://hal.archives-ouvertes.fr/hal-01271141

[4] E. FELLER, L. RILLING, C. MORIN. *Energy-Aware Ant Colony Based Workload Placement in Clouds*, in "The 12th IEEE/ACM International Conference on Grid Computing (GRID-2011)", Lyon, France, September 2011, https://hal.inria.fr/inria-00626042

[5] A. IORDACHE, C. MORIN, N. PARLAVANTZAS, P. RITEAU. *Resilin: Elastic MapReduce over Multiple Clouds*, Inria, October 2012, n⁰ RR-8081, https://hal.inria.fr/hal-00737030

[6] A. IORDACHE, G. PIERRE, P. SANDERS, J. G. D. F. COUTINHO, M. STILLWELL. *High Performance in the Cloud with FPGA Groups*, in "9th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2016)", Shanghai, China, December 2016, https://hal.inria.fr/hal-01356998

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

[8] J. ROJAS BALDERRAMA, M. SIMONIN, C. TEDESCHI. *GinFlow: A Decentralised Adaptive Workflow Execution Manager*, in "30th IEEE International Parallel & Distributed Processing Symposium", Chicago, United States, Proceedings of the 30th IEEE International Parallel & Distributed Processing Symposium, May 2016, https://hal.archives-ouvertes.fr/hal-01292135

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] I. CUADRADO-CORDERO. *Microclouds : an approach for a network-aware energy-efficient decentralised cloud*, Université Rennes 1, February 2017, https://tel.archives-ouvertes.fr/tel-01498190

[10] A. GIANNAKOU. *Self-adaptable Security Monitoring for IaaS Cloud Environments*, Insa Rennes, July 2017, https://hal.inria.fr/tel-01653831

[11] M. S. HASAN. *Smart management of renewable energy in clouds : from infrastructure to application*, INSA de Rennes, May 2017, https://tel.archives-ouvertes.fr/tel-01617963

[12] Y. LI. *Resource allocation in a Cloud partially powered by renewable energy sources*, Ecole nationale supérieure Mines-Télécom Atlantique, June 2017, https://tel.archives-ouvertes.fr/tel-01595953

[13] C. TEDESCHI. *Distributed Chemically-Inspired Runtimes for Large Scale Adaptive Computing Platforms*, Université de Rennes 1, April 2017, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01665776

Articles in International Peer-Reviewed Journals

[14] F. ALMEIDA, M. DIAS DE ASSUNCAO, J. BARBOSA, V. BLANCO, I. BRANDIC, G. DA COSTA, M. F. DOLZ, A. C. ELSTER, M. JARUS, H. KARATZA, L. LEFÈVRE, I. MAVRIDIS, A. OLEKSIAK, A.-C. ORGERIE, J.-M. PIERSON. *Energy Monitoring as an Essential Building Block Towards Sustainable Ultrascale Systems*, in "Sustainable Computing : Informatics and Systems", 2017, forthcoming, https://hal.inria.fr/hal-01627757

[15] N. BELDICEANU, B. DUMAS FERIS, P. GRAVEY, M. S. HASAN, C. JARD, T. LEDOUX, Y. LI, D. LIME, G. MADI-WAMBA, J.-M. MENAUD, P. MOREL, M. MORVAN, M.-L. MOULINARD, A.-C. ORGERIE, J.-L.

PAZAT, O. H. ROUX, A. SHARAIHA. *Towards energy-proportional Clouds partially powered by renewable energy*, in "Computing", January 2017, vol. 99, n° 1, 20 p. [*DOI :* 10.1007/S00607-016-0503-Z], https://hal.inria.fr/hal-01340318

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

[17] M. CALLAU-ZORI, L. SAMOILA, A.-C. ORGERIE, G. PIERRE. *An experiment-driven energy consumption model for virtual machine management systems*, in "Sustainable Computing : Informatics and Systems", 2017, forthcoming, https://hal.archives-ouvertes.fr/hal-01632962

[18] S. COSTACHE, D. DIB, N. PARLAVANTZAS, C. MORIN. *Resource Management in Cloud Platform as a Service Systems: Analysis and Opportunities*, in "Journal of Systems and Software", May 2017 [*DOI :* 10.1016/J.JSS.2017.05.035], https://hal.archives-ouvertes.fr/hal-01534992

[19] S. COSTACHE, S. KORTAS, C. MORIN, N. PARLAVANTZAS. *Market-based Autonomous Resource and Application Management in Private Clouds*, in "Journal of Parallel and Distributed Computing", February 2017, vol. 100, pp. 85-102 [*DOI :* 10.1016/J.JPDC.2016.10.003], https://hal.archives-ouvertes.fr/hal-01378536

[20] A. DEGOMME, A. LEGRAND, G. MARKOMANOLIS, M. QUINSON, M. L. STILLWELL, F. SUTER. *Simulating MPI applications: the SMPI approach*, in "IEEE Transactions on Parallel and Distributed Systems", February 2017, 14 p. [*DOI :* 10.1109/TPDS.2017.2669305], https://hal.inria.fr/hal-01415484

[21] D. DIB, N. PARLAVANTZAS, C. MORIN. *SLA-based PaaS profit optimization*, in "Concurrency and Computation: Practice and Experience", August 2017 [*DOI :* 10.1002/CPE.4251], https://hal.archives-ouvertes.fr/hal-01585288

[22] M. S. HASAN, F. ALVARES, T. LEDOUX, J.-L. PAZAT. *Investigating Energy Consumption and Performance Trade-off for Interactive Cloud Application*, in "IEEE Transactions on Sustainable Computing", June 2017 [*DOI :* 10.1109/TSUSC.2017.2714959], https://hal.archives-ouvertes.fr/hal-01540159

[23] M. S. HASAN, Y. KOUKI, T. LEDOUX, J.-L. PAZAT. *Exploiting Renewable sources: when Green SLA becomes a possible reality in Cloud computing*, in "IEEE Transactions on Cloud Computing", June 2017, vol. 5, n° 2, pp. 249-262 [*DOI :* 10.1109/TCC.2015.2459710], https://hal.archives-ouvertes.fr/hal-01187907

[24] M. KURPICZ, A.-C. ORGERIE, A. SOBE, P. FELBER. *Energy-proportional Profiling and Accounting in Heterogeneous Virtualized Environments*, in "Sustainable Computing : Informatics and Systems", 2017, pp. 1-36, forthcoming, https://hal.archives-ouvertes.fr/hal-01633435

[25] Y. LI, A.-C. ORGERIE, I. RODERO, B. LEMMA AMERSHO, M. PARASHAR, J.-M. MENAUD. *End-to-end Energy Models for Edge Cloud-based IoT Platforms: Application to Data Stream Analysis in IoT*, in "Future Generation Computer Systems", 2017, pp. 1-12, forthcoming [*DOI :* 10.1016/J.FUTURE.2017.12.048], https://hal.archives-ouvertes.fr/hal-01673501

International Conferences with Proceedings

[26] A. BENOIT, L. LEFÈVRE, A.-C. ORGERIE, I. RAÏS. *Shutdown Policies with Power Capping for Large Scale Computing Systems*, in "Euro-Par: International European Conference on Parallel and Distributed Computing", Santiago de Compostela, Spain, F. F. RIVERA, T. F. PENA, J. C. CA-BALEIRO (editors), Lecture Notes in Computer Science, springer, August 2017, vol. 10417, pp. 134 - 146 [*DOI :* 10.1109/COMST.2016.2545109], https://hal.archives-ouvertes.fr/hal-01589555

[27] B. CAMUS, F. DUFOSSÉ, A.-C. ORGERIE. *A stochastic approach for optimizing green energy consumption in distributed clouds*, in "SMARTGREENS 2017 - International Conference on Smart Cities and Green ICT Systems", Porto, Portugal, April 2017, https://hal.inria.fr/hal-01475431

[28] R.-A. CHERRUEAU, D. PERTIN, A. SIMONET, A. LEBRE, M. SIMONIN. *Toward a Holistic Framework for Conducting Scientific Evaluations of OpenStack*, in "The 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)", Madrid, Spain, May 2017, pp. 544-548 [*DOI :* 10.1109/CCGRID.2017.87], https://hal.archives-ouvertes.fr/hal-01664515

[29] I. CUADRADO-CORDERO, A.-C. ORGERIE, J.-M. MENAUD. *Comparative Experimental Analysis of the Quality-of-Service and Energy-Efficiency of VMs and Containers' Consolidation for Cloud Applications*, in "SoftCOM: International Conference on Software, Telecommunications and Computer Networks", Split, Croatia, September 2017, https://hal.archives-ouvertes.fr/hal-01578325

[30] L. GUEGAN, N. MURAD, J. M. LEBRETON, S. BONHOMMEAU. *Integrating Over Sea Radio Channel for Sea Turtles Localization in the Indian Ocean*, in "IEEE RADIO 2017", Cape Town, South Africa, September 2017, https://arxiv.org/abs/1712.03146 , https://hal.archives-ouvertes.fr/hal-01632295

[31] D. GUYON, A.-C. ORGERIE, C. MORIN, D. AGARWAL. *How Much Energy can Green HPC Cloud Users Save?*, in "PDP 2017 - 25th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing", Saint Petersbourg, Russia, March 2017, https://hal.inria.fr/hal-01439874

[32] D. GUYON, A.-C. ORGERIE, C. MORIN. *GLENDA: Green Label towards Energy proportioNality for IaaS DAta centers*, in "E2DC: International Workshop on Energy Efficient Data Centres (e-Energy Workshop)", Hong Kong, Hong Kong SAR China, May 2017 [*DOI :* 10.1145/3077839.3084028], https://hal.archives-ouvertes.fr/hal-01514948

[33] M. S. HASAN, F. ALVARES DE OLIVEIRA JR., T. LEDOUX. *GPaaScaler: Green Energy aware Platform Scaler for Interactive Cloud Application*, in "IEEE/ACM International Conference on Utility and Cloud Computing", Austin, United States, December 2017, https://hal.archives-ouvertes.fr/hal-01650136

[34] *Best Paper*
T. HAUDEBOURG, A.-C. ORGERIE. *On the Energy Efficiency of Sleeping and Rate Adaptation for Network Devices*, in "ICA3PP 2017 - 17th International Conference on Algorithms and Architectures for Parallel Processing", Helsinki, Finland, Springer, August 2017, vol. 10393, pp. 132-146, Best Paper Award [*DOI :* 10.1007/978-3-319-65482-9_9], https://hal.archives-ouvertes.fr/hal-01575513.

[35] F. C. HEINRICH, T. CORNEBIZE, A. DEGOMME, A. LEGRAND, A. CARPEN-AMARIE, S. HUNOLD, A.-C. ORGERIE, M. QUINSON. *Predicting the Energy Consumption of MPI Applications at Scale Using a Single Node*, in "Cluster 2017", Hawaii, United States, IEEE, September 2017, https://hal.inria.fr/hal-01523608

[36] Y. LI, A.-C. ORGERIE, J.-M. MENAUD. *Balancing the use of batteries and opportunistic scheduling policies for maximizing renewable energy consumption in a Cloud data center*, in "PDP 2017 - 25th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing", St Petersburg, Russia, March 2017, https://hal.inria.fr/hal-01432752

[37] Y. LI, A.-C. ORGERIE, I. RODERO, M. PARASHAR, J.-M. MENAUD. *Leveraging Renewable Energy in Edge Clouds for Data Stream Analysis in IoT*, in "CCGrid 2017 - IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Madrid, Spain, May 2017, https://hal.inria.fr/hal-01472358

[38] G. MADI WAMBA, Y. LI, A.-C. ORGERIE, N. BELDICEANU, J.-M. MENAUD. *Cloud workload prediction and generation models*, in "SBAC-PAD: International Symposium on Computer Architecture and High Performance Computing", Campinas, Brazil, October 2017, https://hal.archives-ouvertes.fr/hal-01578354

[39] G. MADI WAMBA, Y. LI, A.-C. ORGERIE, N. BELDICEANU, J.-M. MENAUD. *Green energy aware scheduling problem in virtualized datacenters*, in "ICPADS 2017", Shenzen, China, December 2017, https://hal.archives-ouvertes.fr/hal-01582936

[40] D. MARGERY, D. GUYON, A.-C. ORGERIE, C. MORIN, G. FRANCIS, C. PALANSURIYA, K. KAVOUS-SANAKIS. *A CO2 emissions accounting framework with market-based incentives for Cloud infrastructures*, in "SMARTGREENS: International Conference on Smart Cities and Green ICT Systems", Porto, Portugal, April 2017, https://hal.inria.fr/hal-01486185

[41] P. MENSAH, S. DUBUS, W. KANOUN, C. MORIN, G. PIOLLE, E. TOTEL. *Connectivity Extraction in Cloud Infrastructures*, in "CNSM 2017 - 13th International Conference on Network and Service Management", Tokyo, Japan, November 2017, pp. 1-5, https://hal.inria.fr/hal-01593346

[42] P. MENSAH, S. DUBUS, W. KANOUN, C. MORIN, G. PIOLLE, E. TOTEL. *Connectivity Graph Reconstruction for Networking Cloud Infrastructures*, in "NCA 2017 - 16th IEEE International Symposium on Network Computing and Applications", Cambridge, MA, United States, October 2017, https://hal.inria.fr/hal-01612988

[43] A.-C. ORGERIE, B. LEMMA AMERSHO, T. HAUDEBOURG, M. QUINSON, M. RIFAI, D. LOPEZ PACHECO, L. LEFÈVRE. *Simulation Toolbox for Studying Energy Consumption in Wired Networks*, in "CNSM: International Conference on Network and Service Management", Tokyo, Japan, November 2017, pp. 1-5, https://hal.archives-ouvertes.fr/hal-01630226

[44] T. A. PHAM, T. JÉRON, M. QUINSON. *Verifying MPI Applications with SimGridMC*, in "Correctness 2017 - First International Workshop on Software Correctness for HPC Applications", Denver, United States, November 2017 [*DOI :* 10.1145/3145344.3145345], https://hal.inria.fr/hal-01632421

[45] M. L. PHAM, N. PARLAVANTZAS, C. MORIN, S. ARNOUX, L. QI, P. GONTIER, P. EZANNO. *DiFFuSE, a Distributed Framework for Cloud-based Epidemic Simulations: a Case Study in Modelling the Spread of Bovine Viral Diarrhea Virus*, in "CloudCom 2017 - 9th IEEE International Conference on Cloud Computing Technology and Science", Hong-Kong, China, December 2017, https://hal.inria.fr/hal-01612979

[46] C. RUIZ, H. A. DURAN-LIMON, N. PARLAVANTZAS. *An RLS Memory-based Mechanism for the Automatic Adaptation of VMs on Cloud Environments*, in "Workshop on Adaptive Resource Management and Scheduling for Cloud Computing", Washington, DC, United States, July 2017 [*DOI :* 10.1145/3110355.3110358], https://hal.archives-ouvertes.fr/hal-01637794

[47] G. TATO, M. BERTIER, C. TEDESCHI. *Designing Overlay Networks for Decentralized Clouds*, in "Future-Cloud 2017 - International Workshop on the Future of Cloud Computing and Cloud Services", Hong-Kong, China, December 2017, https://hal.inria.fr/hal-01634409

[48] A. VAN KEMPEN, T. CRIVAT, B. TRUBERT, D. ROY, G. PIERRE. *MEC-ConPaaS: An experimental single-board based mobile edge cloud*, in "IEEE Mobile Cloud Conference", San Francisco, United States, April 2017, https://hal.inria.fr/hal-01446483

[49] A. T. WONJIGA, L. RILLING, C. MORIN. *Verification for Security Monitoring SLAs in IaaS Clouds: the Example of a Network IDS*, in "NOMS 2018 - Network Operations and Management Symposium", Taipei, Taiwan, April 2018, https://hal.inria.fr/hal-01663941

National Conferences with Proceedings

[50] S. DELAMARE, P. MORILLON, L. NUSSBAUM. *Réalisation d'expériences avec Grid'5000*, in "JRES2017 - Journées Réseaux de l'enseignement et de la recherche", Nantes, France, November 2017, https://hal.inria.fr/hal-01639524

Conferences without Proceedings

[51] N. MURAD, L. GUEGAN, S. BONHOMMEAU. *Why satellite localization beacons are not adapted for marine turtles' study: a sea wireless sensors network solution*, in "GIIS'17 - Global Information Infrastructure and Networking Symposium", Saint Pierre, Réunion, IEEE, October 2017, https://hal.archives-ouvertes.fr/hal-01631305

Scientific Books (or Scientific Book chapters)

[52] J. G. F. COUTINHO, M. STILLWELL, K. ARGYRAKI, G. IOANNIDIS, A. IORDACHE, C. KLEINEWEBER, A. KOLIOUSIS, J. MCGLONE, G. PIERRE, C. RAGUSA, P. SANDERS, T. SCHÜTT, T. YU, A. WOLF. *The HARNESS Platform: A Hardware-and Network-Enhanced Software System for Cloud Computing*, in "Software Architecture for Big Data and the Cloud", I. MISTRIK, R. BAHSOON, N. ALI, M. HEISEL, B. MAXIM (editors), Morgan Kaufmann, June 2017, https://hal.inria.fr/hal-01507344

Research Reports

[53] R.-A. CHERRUEAU, M. SIMONIN, A. VAN KEMPEN. *EnosStack: A LAMP-like stack for the experimenter*, Inria Rennes - Bretagne Atlantique, January 2018, n^o RR-9146, pp. 1-14, https://hal.inria.fr/hal-01689726

[54] L. M. PHAM, N. PARLAVANTZAS, C. MORIN, S. ARNOUX, L. QI, P. GONTIER, P. EZANNO. *DiFFuSE, a Distributed Framework for Cloud-based Epidemic Simulations: a Case Study in Modelling the Spread of Bovine Viral Diarrhea Virus*, Inria Rennes - Bretagne Atlantique, September 2017, n^o RR-9094, https://hal.archives-ouvertes.fr/hal-01586945

[55] A. TESHOME, L. RILLING, C. MORIN. *Verification for Security Monitoring SLAs in IaaS Clouds: the Example of a Network IDS*, Inria Rennes Bretagne Atlantique, August 2017, n^o RR-9091, https://hal.inria.fr/hal-01577814

Other Publications

[56] R. FABLET, N. BELLEC, L. CHAPEL, C. FRIGUET, R. GARELLO, P. GLOAGUEN, G. HAJDUCH, S. LEFÈVRE, F. MERCIOL, P. MORILLON, C. MORIN, M. SIMONIN, R. TAVENARD, C. TEDESCHI, R.

VADAINE. *Next Step for Big Data Infrastructure and Analytics for the Surveillance of the Maritime Traffic from AIS & Sentinel Satellite Data Streams*, November 2017, pp. 1-4, BiDS' 2017 - Conference on Big Data from Space, Poster, https://hal.inria.fr/hal-01671323

[57] B. GOUPILLE-LESCAR, E. LENORMAND, C. MORIN, N. PARLAVANTZAS. *Introducing Dynamic Adaptation in High Performance Real-Time Computing Platforms for Sensors*, September 2017, ANDARE 2017 - 1st Workshop on AutotuniNg and aDaptivity AppRoaches for Energy efficient HPC Systems, Poster, https://hal.inria.fr/hal-01624262

[58] F. C. HEINRICH, A. CARPEN-AMARIE, A. DEGOMME, S. HUNOLD, A. LEGRAND, A.-C. ORGERIE, M. QUINSON. *Predicting the Performance and the Power Consumption of MPI Applications With SimGrid*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01446134

References in notes

[59] *Automatiser la reconfiguration des sondes de détection d'intrusion pour le cloud*, July 2017, n° 47, http://emergences.inria.fr/2017/newsletter-n47/newsletter_show_all

[60] 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, https://hal.inria.fr/hal-00854204