Activity Report 2016

Team POLARIS

Performance analysis and optimization of LARge Infrastructures and Systems

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Team POLARIS

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

1.1.1. - Multicore
1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
1.1.4. - High performance computing
1.1.5. - Exascale
1.2. - Networks
1.2.3. - Routing
1.2.5. - Internet of things
1.6. - Green Computing
5.2. - Data visualization
6. - Modeling, simulation and control
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.2.6. - Optimization
6.2.7. - High performance computing
7.3. - Optimization
7.11. - Performance evaluation
7.14. - Game Theory

Other Research Topics and Application Domains:

4.4. - Energy delivery
4.4.1. - Smart grids
4.5.1. - Green computing
6.2. - Network technologies
6.2.1. - Wired technologies
6.2.2. - Radio technology
6.4. - Internet of things
8.3. - Urbanism and urban planning
9.5.7. - Geography
9.6. - Reproducibility
9.7.2. - Open data

1. Members

Research Scientists

Arnaud Legrand [Team leader, CNRS, Researcher, HDR]
Nicolas Gast [Inria, Researcher]
Bruno Gaujal [Inria, Senior Researcher, HDR]
Panayotis Mertikopoulos [CNRS, Researcher]

Faculty Members

Elena-Veronica Belmega [ENSEA, Associate Processor]
Vincent Danjean [Univ. Grenoble Alpes, Associate Processor]
Guillaume Huard [Univ. Grenoble Alpes, Associate Processor]
Florence Perronnin [Univ. Grenoble Alpes, Associate Processor]
Jean-Marc Vincent [Univ. Grenoble Alpes, Associate Processor]
Philippe Waille [Univ. Grenoble Alpes, Associate Processor]

Engineer

Benjamin Briot [Inria, until Jan 2016]

PhD Students

David Beniamine [Inria]
Stephane Durand [Univ. Grenoble Alpes]
Franz Christian Heinrich [Inria]
Baptiste Jonglez [Inria]
Alexandre Marcastel [Univ. Cergy-Pontoise]
Alexis Martin [Inria]
Stephan Plassart [Inria, from Sep 2016]
Benoît Vinot [Schneider Electric, granted by CIFRE]
Rafael Keller Tesser [UFRGS, from August 2015]
Vinicius Garcia Pinto [UFRGS/Univ. Grenoble Alpes, From Sep 2015]

Post-Doctoral Fellows

Josu Doncel [Inria, until Nov 2016]
Ahmed El Rheddane [Inria, until Feb 2016]
Carmen Higuera Chan [Inria, from Sep 2016]
Guillaume Massonnet [Inria, until Aug 2016]
Lucas Mello Schnorr [Inria]

Administrative Assistant

Annie Simon [Inria]

Others

Mathieu Baille [Inria, Intern M1, Jun 2015 until Jul 2016]
Vania Martin [Univ. Grenoble Alpes, Faculty Member]
Sebastien Ochier [Inria, Intern, from Feb 2016 until Jun 2016]
Michael Picard [Inria, Intern, from Jun 2016 until Jul 2016]
Loic Poncet [Inria, Intern, from Jun 2016 until Jul 2016]
Florian Popek [Inria, Intern, until Aug 2016]
Steven Quinito Masnada [CNRS, Intern M2, until Jun 2016]
Angelika Studeny [Inria, Post-Doctoral Fellow, until June 2016]

2. Overall Objectives

2.1. Context

Large distributed infrastructures are rampant in our society. Numerical simulations form the basis of computational sciences and high performance computing infrastructures have become scientific instruments with similar roles as those of test tubes or telescopes. Cloud infrastructures are used by companies in such an intense way that even the shortest outage quickly incurs the loss of several millions of dollars. But every citizen also relies on (and interacts with) such infrastructures via complex wireless mobile embedded devices whose nature is constantly evolving. In this way, the advent of digital miniaturization and interconnection has enabled our homes, power stations, cars and bikes to evolve into smart grids and smart transportation systems that should be optimized to fulfill societal expectations.

Our dependence and intense usage of such gigantic systems obviously leads to very high expectations in terms of performance. Indeed, we strive for low-cost and energy-efficient systems that seamlessly adapt to changing environments that can only be accessed through uncertain measurements. Such digital systems also have to take into account both the users' profile and expectations to efficiently and fairly share resources in an online way. Analyzing, designing and provisioning such systems has thus become a real challenge.

Such systems are characterized by their **ever-growing size**, intrinsic **heterogeneity** and **distributedness**, **user-driven** requirements, and an unpredictable variability that renders them essentially **stochastic**. In such contexts, many of the former design and analysis hypotheses (homogeneity, limited hierarchy, omniscient view, optimization carried out by a single entity, open-loop optimization, user outside of the picture) have become obsolete, which calls for radically new approaches. Properly studying such systems requires a drastic rethinking of fundamental aspects regarding the system's **observation** (measure, trace, methodology, design of experiments), **analysis** (modeling, simulation, trace analysis and visualization), and **optimization** (distributed, online, stochastic).

2.2. Objectives

The goal of the POLARIS project is to **contribute to the understanding of the performance of very large scale distributed systems** by applying ideas from diverse research fields and application domains. We believe that studying all these different aspects at once without restricting to specific systems is the key to push forward our understanding of such challenges and to proposing innovative solutions. This is why we intend to investigate problems arising from application domains as varied as large computing systems, wireless networks, smart grids and transportation systems.

The members of the POLARIS project cover a very wide spectrum of expertise in performance evaluation and models, distributed optimization, and analysis of HPC middleware. Specifically, POLARIS' members have worked extensively on:

Experiment design: Experimental methodology, measuring/monitoring/tracing tools, experiment control, design of experiments, and reproducible research, especially in the context of large computing infrastructures (such as computing grids, HPC, volunteer computing and embedded systems).

Trace Analysis: Parallel application visualization (paje, triva/viva, framesoc/ocelotl, ...), characterization of failures in large distributed systems, visualization and analysis for geographical information systems, spatio-temporal analysis of media events in RSS flows from newspapers, and others.

Modeling and Simulation: Emulation, discrete event simulation, perfect sampling, Markov chains, Monte Carlo methods, and others.

Optimization: Stochastic approximation, mean field limits, game theory, discrete and continuous optimization, learning and information theory.

In the rest of this document, we describe in detail our new results in the above areas.

3. Research Program

3.1. Sound and Reproducible Experimental Methodology

Participants: Vincent Danjean, Nicolas Gast, Guillaume Huard, Arnaud Legrand, Jean-Marc Vincent.

Experiments in large scale distributed systems are costly, difficult to control and therefore difficult to reproduce. Although many of these digital systems have been built by men, they have reached such a complexity level that we are no longer able to study them like artificial systems and have to deal with the same kind of experimental issues as natural sciences. The development of a sound experimental methodology for the evaluation of resource management solutions is among the most important ways to cope with the growing complexity of computing environments. Although computing environments come with their own specific challenges, we believe such general observation problems should be addressed by borrowing good practices and techniques developed in many other domains of science.

This research theme builds on a transverse activity on *Open science and reproducible research* and is organized into the following two directions: (1) *Experimental design* (2) *Smart monitoring and tracing*. As we will explain in more detail hereafter, these transverse activity and research directions span several research areas and our goal within the POLARIS project is foremost to transfer original ideas from other domains of science to the distributed and high performance computing community.

3.2. Multi-Scale Analysis and Visualization

Participants: Vincent Danjean, Guillaume Huard, Arnaud Legrand, Jean-Marc Vincent, Panayotis Mertikopoulos.

As explained in the previous section, the first difficulty encountered when modeling large scale computer systems is to observe these systems and extract information on the behavior of both the architecture, the middleware, the applications, and the users. The second difficulty is to *visualize* and *analyze* such *multi-level traces to understand how the performance* of the application *can be improved*. While a lot of efforts are put into visualizing scientific data, in comparison little effort have gone into to developing techniques specifically tailored for understanding the behavior of distributed systems. Many visualization tools have been developed by renowned HPC groups since decades (e.g., BSC [87], Jülich and TU Dresden [86], [55], UIUC [74], [90], [77] and ANL [104], Inria Bordeaux [61] and Grenoble [106], ...) but most of these tools build on the classical information visualization mantra [95] that consists in always first presenting an overview of the data, possibly by plotting everything if computing power allows, and then to allow users to zoom and filter, providing details on demand. However in our context, the amount of data comprised in such traces is several orders of magnitude larger than the number of pixels on a screen and displaying even a small fraction of the trace leads to harmful visualization artifacts [82]. Such traces are typically made of events that occur at very different time and space scales, which unfortunately hinders classical approaches. Such visualization tools have focused on easing interaction and navigation in the trace (through gantcharts, intuitive filters, pie charts and kiviats) but they are very difficult to maintain and evolve and they require some significant experience to identify performance bottlenecks.

Therefore many groups have more recently proposed in combination to these tools some techniques to help identifying the structure of the application or regions (applicative, spatial or temporal) of interest. For example, researchers from the SDSC [85] propose some segment matching techniques based on clustering (Euclidean or Manhattan distance) of start and end dates of the segments that enables to reduce the amount of information to display. Researchers from the BSC use clustering, linear regression and Kriging techniques [94], [81], [73] to identify and characterize (in term of performance and resource usage) application phases and present aggregated representations of the trace [93]. Researchers from Jülich and TU Darmstadt have proposed techniques to identify specific communication patterns that incur wait states [101], [48]

3.3. Fast and Faithful Performance Prediction of Very Large Systems

Participants: Vincent Danjean, Bruno Gaujal, Arnaud Legrand, Florence Perronnin, Jean-Marc Vincent.

Evaluating the scalability, robustness, energy consumption and performance of large infrastructures such as exascale platforms and clouds raises severe methodological challenges. The complexity of such platforms mandates empirical evaluation but direct experimentation via an application deployment on a real-world testbed is often limited by the few platforms available at hand and is even sometimes impossible (cost, access, early stages of the infrastructure design, ...). Unlike direct experimentation via an application deployment on a real-world testbed, simulation enables fully repeatable and configurable experiments that can often be conducted quickly for arbitrary hypothetical scenarios. In spite of these promises, current simulation practice is often not conducive to obtaining scientifically sound results. To date, most simulation results in the parallel and distributed computing literature are obtained with simulators that are ad hoc, unavailable, undocumented, and/or no longer maintained. For instance, Naicken et al. [47] point out that out of 125 recent papers they surveyed that study peer-to-peer systems, 52% use simulation and mention a simulator, but 72% of them use a custom simulator. As a result, most published simulation results build on throw-away (short-lived and non validated) simulators that are specifically designed for a particular study, which prevents other researchers from building upon it. There is thus a strong need for recognized simulation frameworks by which simulation results can be reproduced, further analyzed and improved.

The *SimGrid* simulation toolkit [59], whose development is partially supported by POLARIS, is specifically designed for studying large scale distributed computing systems. It has already been successfully used for simulation of grid, volunteer computing, HPC, cloud infrastructures and we have constantly invested on the software quality, the scalability [51] and the validity of the underlying network models [49], [99]. Many simulators of MPI applications have been developed by renowned HPC groups (e.g., at SDSC [97], BSC [45], UIUC [105], Sandia Nat. Lab. [100], ORNL [58] or ETH Zürich [75] for the most prominent ones). Yet, to scale most of them build on restrictive network and application modeling assumptions that make them difficult to extend to more complex architectures and to applications that do not solely build on the MPI API. Furthermore, simplistic modeling assumptions generally prevent to faithfully predict execution times, which limits the use of simulation to indication of gross trends at best. Our goal is to improve the quality of SimGrid to the point where it can be used effectively on a daily basis by practitioners to *reproduce the dynamic of real HPC systems*.

We also develop another simulation software, *PSI* (Perfect SImulator) [63], [56], dedicated to the simulation of very large systems that can be modeled as Markov chains. PSI provides a set of simulation kernels for Markov chains specified by events. It allows one to sample stationary distributions through the Perfect Sampling method (pioneered by Propp and Wilson [88]) or simply to generate trajectories with a forward Monte-Carlo simulation leveraging time parallel simulation (pioneered by Fujimoto [67], Lin and Lazowska [80]). One of the strength of the PSI framework is its expressiveness that allows us to easily study networks with finite and infinite capacity queues [57]. Although PSI already allows to simulate very large and complex systems, our main objective is to push its scalability even further and *improve its capabilities by one or several orders of magnitude*.

3.4. Local Interactions and Transient Analysis in Adaptive Dynamic Systems

Participants: Nicolas Gast, Bruno Gaujal, Florence Perronnin, Jean-Marc Vincent, Panayotis Mertikopoulos.

Many systems can be effectively described by stochastic population models. These systems are composed of a set of n entities interacting together and the resulting stochastic process can be seen as a continuous-time Markov chain with a finite state space. Many numerical techniques exist to study the behavior of Markov chains, to solve stochastic optimal control problems [89] or to perform model-checking [46]. These techniques, however, are limited in their applicability, as they suffer from the *curse of dimensionality*: the state-space grows exponentially with n.

This results in the need for approximation techniques. Mean field analysis offers a viable, and often very accurate, solution for large n. The basic idea of the mean field approximation is to count the number of entities that are in a given state. Hence, the fluctuations due to stochasticity become negligible as the number of entities grows. For large n, the system becomes essentially deterministic. This approximation has been originally developed in statistical mechanics for vary large systems composed of more than 10^{20} particles (called entities here). More recently, it has been claimed that, under some conditions, this approximation can be successfully used for stochastic systems composed of a few tens of entities. The claim is supported by various convergence results [68], [78], [103], and has been successfully applied in various domains: wireless networks [50], computer-based systems [71], [84], [98], epidemic or rumour propagation [60], [76] and bike-sharing systems [64]. It is also used to develop distributed control strategies [102], [83] or to construct approximate solutions of stochastic model checking problems [52], [53], [54].

Within the POLARIS project, we will continue developing both the theory behind these approximation techniques and their applications. Typically, these techniques require a homogeneous population of objects where the dynamics of the entities depend only on their state (the state space of each object must not scale with n the number of objects) but neither on their identity nor on their spatial location. Continuing our work in [68], we would like to be able to handle heterogeneous or uncertain dynamics. Typical applications are caching mechanisms [71] or bike-sharing systems [65]. A second point of interest is the use of mean field or large deviation asymptotics to compute the time between two regimes [92] or to reach an equilibrium state. Last, mean-field methods are mostly descriptive and are used to analyse the performance of a given system. We wish to extend their use to solve optimal control problems. In particular, we would like to implement numerical algorithms that use the framework that we developed in [69] to build distributed control algorithms [62] and optimal pricing mechanisms [70].

3.5. Distributed Learning in Games and Online Optimization

Participants: Nicolas Gast, Bruno Gaujal, Arnaud Legrand, Panayotis Mertikopoulos.

Game theory is a thriving interdisciplinary field that studies the interactions between competing optimizing agents, be they humans, firms, bacteria, or computers. As such, game-theoretic models have met with remarkable success when applied to complex systems consisting of interdependent components with vastly different (and often conflicting) objectives – ranging from latency minimization in packet-switched networks to throughput maximization and power control in mobile wireless networks.

In the context of large-scale, decentralized systems (the core focus of the POLARIS project), it is more relevant to take an inductive, "bottom-up" approach to game theory, because the components of a large system cannot be assumed to perform the numerical calculations required to solve a very-large-scale optimization problem. In view of this, POLARIS' overarching objective in this area is to *develop novel algorithmic frameworks that offer robust performance guarantees when employed by all interacting decision-makers.*

A key challenge here is that most of the literature on learning in games has focused on *static* games with a *finite number of actions* per player [66], [91]. While relatively tractable, such games are ill-suited to practical applications where players pick an action from a continuous space or when their payoff functions evolve over time – this being typically the case in our target applications (e.g., routing in packet-switched networks or energy-efficient throughput maximization in wireless). On the other hand, the framework of online convex optimization typically provides worst-case performance bounds on the learner's *regret* that the agents can attain irrespectively of how their environment varies over time. However, if the agents' environment is determined chiefly by their interactions these bounds are fairly loose, so more sophisticated convergence criteria should be applied.

From an algorithmic standpoint, a further challenge occurs when players can only observe their own payoffs (or a perturbed version thereof). In this bandit-like setting regret-matching or trial-and-error procedures guarantee convergence to an equilibrium in a weak sense in certain classes of games. However, these results apply exclusively to static, finite games: learning in games with continuous action spaces and/or nonlinear payoff functions cannot be studied within this framework. Furthermore, even in the case of finite games, the complexity of the algorithms described above is not known, so it is impossible to decide a priori which algorithmic scheme can be applied to which application.

4. Application Domains

4.1. Large Computing Infrastructures

Supercomputers typically comprise thousands to millions of multi-core CPUs with GPU accelerators interconnected by complex interconnection networks that are typically structured as an intricate hierarchy of network switches. Capacity planning and management of such systems not only raises challenges in term of computing efficiency but also in term of energy consumption. Most legacy (SPMD) applications struggle to benefit from such infrastructure since the slightest failure or load imbalance immediately causes the whole program to stop or at best to waste resources. To scale and handle the stochastic nature of resources, these applications have to rely on dynamic runtimes that schedule computations and communications in an opportunistic way. Such evolution raises challenges not only in terms of programming but also in terms of observation (complexity and dynamicity prevents experiment reproducibility, intrusiveness hinders large scale data collection, ...) and analysis (dynamic and flexible application structures make classical visualization and simulation techniques totally ineffective and require to build on *ad hoc* information on the application structure).

4.2. Next-Generation Wireless Networks

Considerable interest has arisen from the seminal prediction that the use of multiple-input, multiple-output (MIMO) technologies can lead to substantial gains in information throughput in wireless communications, especially when used at a massive level. In particular, by employing multiple inexpensive service antennas, it is possible to exploit spatial multiplexing in the transmission and reception of radio signals, the only physical limit being the number of antennas that can be deployed on a portable device. As a result, the wireless medium can accommodate greater volumes of data traffic without requiring the reallocation (and subsequent re-regulation) of additional frequency bands. In this context, throughput maximization in the presence of interference by neighboring transmitters leads to games with convex action sets (covariance matrices with trace constraints) and individually concave utility functions (each user's Shannon throughput); developing efficient and distributed optimization protocols for such systems is one of the core objectives of Theme 5.

Another major challenge that occurs here is due to the fact that the efficient physical layer optimization of wireless networks relies on perfect (or close to perfect) channel state information (CSI), on both the uplink and the downlink. Due to the vastly increased computational overhead of this feedback – especially in decentralized, small-cell environments – the ongoing transition to fifth generation (5G) wireless networks is expected to go hand-in-hand with distributed learning and optimization methods that can operate reliably in feedback-starved environments. Accordingly, one of POLARIS' application-driven goals will be to leverage the algorithmic output of Theme 5 into a highly adaptive resource allocation framework for next-gneration wireless systems that can effectively "learn in the dark", without requiring crippling amounts of feedback.

4.3. Energy and Transportation

Participant: Nicolas Gast.

This work is mainly done within the Quanticol European project.

Smart urban transport systems and smart grids are two examples of collective adaptive systems. They consist of a large number of heterogeneous entities with decentralised control and varying degrees of complex autonomous behaviour. Within the QUANTICOL project, we develop an analysis tools to help to reason about such systems. Our work relies on tools from fluid and mean-field approximation to build decentralized algorithms that solve complex optimization problems. We focus on two problems: decentralized control of electric grids and capacity planning in vehicle-sharing systems to improve load balancing.

5. New Software and Platforms

5.1. Framesoc

FUNCTIONAL DESCRIPTION

Framesoc is the core software infrastructure of the SoC-Trace project. It provides a graphical user environment for execution-trace analysis, featuring interactive analysis views as Gantt charts or statistics views. It provides also a software library to store generic trace data, play with them, and build other analysis tools (e.g., Ocelotl).

- Participants: Jean-Marc Vincent and Arnaud Legrand
- Contact: Guillaume Huard
- URL: http://soctrace-inria.github.io/framesoc/

5.2. GameSeer

FUNCTIONAL DESCRIPTION

GameSeer is a tool for students and researchers in game theory that uses Mathematica to generate phase portraits for normal form games under a variety of (user-customizable) evolutionary dynamics. The aim of GameSeer is a) to provide a numerical integration kernel for phase portrait and equilibrium set generation; and b) to provide a graphical user interface that allows the user to employ said capabilities from a simple and intuitive front-end.

- Contact: Panayotis Mertikopoulos
- URL: http://mescal.imag.fr/membres/panayotis.mertikopoulos/publications.html

5.3. Moca

Memory Organization Cartography and Analysis
MOCA is an efficient tool for the collection of complete spatiotemporal memory traces. Its objective is twofold, namely to *a*) avoid missuses of the memory hierarchy (such as false sharing of cache lines or contention); and *b*) to take advantage of the various cache levels and the memory hardware prefetcher. It is based on a Linux kernel module and provides a coarse-grained trace of a superset of all the memory accesses performed by an application over its addressing space during the time of its execution.

KEYWORDS: High-Performance Computing - Performance analysis

- Contact: Guillaume Huard
- URL: https://github.com/dbeniamine/MOCA

5.4. Ocelotl

Multidimensional Overviews for Huge Trace Analysis
FUNCTIONAL DESCRIPTION

Ocelotl is an innovative visualization tool, which provides overviews for execution trace analysis by using a data aggregation technique. This technique enables to find anomalies in huge traces containing up to several billions of events, while keeping a fast computation time and providing a simple representation that does not overload the user.

- Participants: Arnaud Legrand and Jean-Marc Vincent
- Contact: Jean-Marc Vincent
- URL: http://soctrace-inria.github.io/ocelotl/

5.5. PSI

Perfect Simulator
FUNCTIONAL DESCRIPTION

Perfect simulator is a simulation software of markovian models. It is able to simulate discrete and continuous time models to provide a perfect sampling of the stationary distribution or directly a sampling of functional of this distribution by using coupling from the past. The simulation kernel is based on the CFTP algorithm, and the internal simulation of transitions on the Aliasing method.

- Contact: Jean-Marc Vincent
- URL: http://psi.gforge.inria.fr/

5.6. 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.

- Participants: Frederic Suter, Martin Quinson, Arnaud Legrand, Adrien Lebre, Jonathan Pastor, Mario Sudholt, Luka Stanisic, Augustin Degomme, Jean-Marc Vincent, Florence Perronnin and Jonathan Rouzaud-Cornabas
- Partners: CNRS - ENS Rennes
- Contact: Martin Quinson
- URL: http://simgrid.gforge.inria.fr/

5.7. Tabarnac

Tool for Analyzing the Behavior of Applications Running on NUMA ArChitecture
KEYWORDS: High-Performance Computing - Performance analysis - NUMA

- Contact: David Beniamine
- URL: https://dbeniamine.github.io/Tabarnac/

5.8. marmoteCore

Markov Modeling Tools and Environments - the Core
KEYWORDS: Modeling - Stochastic models - Markov model
FUNCTIONAL DESCRIPTION

marmoteCore is a C++ environment for modeling with Markov chains. It consists in a reduced set of high-level abstractions for constructing state spaces, transition structures and Markov chains (discrete-time and continuous-time). It provides the ability of constructing hierarchies of Markov models, from the most general to the particular, and equip each level with specifically optimized solution methods.

This software is developed within the ANR MARMOTE project: ANR-12-MONU-00019.

- Participants: Alain Jean-Marie, Issam Rabhi, Jean-Marc Vincent, Benjamin Briot, Jean-Michel Fourneau and Franck Quessette
- Partner: UVSQ
- Contact: Alain Jean-Marie
- URL: http://marmotecore.gforge.inria.fr/

6. New Results

6.1. Asymptotic Models

The analysis of a set of n stochastic entities interacting with each others can be particularly difficult. The *mean field approximation* is a very effective technique to characterize the transient probability distribution or steady-state regime of such systems when the number of entities n grows very large. The idea of mean-field approximation is to replace a complex stochastic system by a simpler deterministic dynamical system. This dynamical system is constructed by assuming that the objects are asymptotically independent. Each object is viewed as interacting with an average of the other objects (the *mean*-field). When each object has a finite or countable state-space, this dynamical system is usually a non-linear ordinary differential equation (ODE). An introduction to these techniques is provided in the book chapter [29].

- Mean-field games model the rational behavior of an infinite number of indistinguishable players in interaction [79]. An important assumption of mean-field games is that, as the number of player is infinite, the decisions of an individual player do not affect the dynamics of the mass. Each player plays against the mass. A mean-field equilibrium corresponds to the case when the optimal decisions of a player coincide with the decisions of the mass. This leads to a simpler computation of the equilibrium.

  It has been shown in [72], [96] that for some games with a finite number of players, the Nash equilibria converge to mean-field equilibria as the number of players tends to infinity. Hence, many authors argue that mean-field games are a good approximation of symmetric stochastic games with a large number of players. The classical argument is that the impact of one player becomes negligible when the number of players goes to infinity. In [17], [36], we show that, in general, this convergence does not hold. We construct an example for which the mean-field limit only describes a sub-set of the limiting equilibria. Each finite-player game has an equilibrium with a good social cost, this is not the case for the limit game.
- Computer system and network performance can be significantly improved by caching frequently used information. When the cache size is limited, the cache replacement algorithm has an important impact on the effectiveness of caching. In [21], [3], [20] we introduce approximations to determine the cache hit probability of two classes of cache replacement algorithms: the recently introduced h-LRU and LRU(m). These approximations only require the requests to be generated according to a general Markovian arrival process (MAP). This includes phase-type renewal processes and the IRM model as special cases. We provide both numerical and theoretical support for the claim that the proposed TTL approximations are asymptotically exact. We further show, by using synthetic and trace-based workloads, that h-LRU and LRU(m) perform alike, while the latter requires less work when a hit/miss occurs.
- In [16], we consider stochastic models in presence of uncertainty, originating from lack of knowledge of parameters or by unpredictable effects of the environment. We focus on population processes, encompassing a large class of systems, from queueing networks to epidemic spreading. We set up a formal framework for imprecise stochastic processes, where some parameters are allowed to vary in time within a given domain, but with no further constraint. We then consider the limit behaviour of these systems as the population size goes to infinity. We prove that this limit is given by a differential inclusion that can be constructed from the (imprecise) drift. We also we discuss different numerical algorithms to compute bounds of the so-obtained differential inclusions. We are currently working on an implementation of these algorithms in a numerical toolbox.
- In [37], we develop a fluid-limit approach to compute the expected absorbing time T_n of a n-dimensional discrete time Markov chain. We show that the random absorbing time T_n is well approximated by a deterministic time t_n that is the first time when a fluid approximation of the chain approaches the absorbing state at a distance $1/n$. We show the applicability of this approach with three different problems: the coupon collector, the erasure channel lifetime and the coupling times of random walks in high dimensional spaces.

6.2. Simulation

Simgrid is a toolkit providing core functionalities for the simulation of distributed applications in heterogeneous distributed environments. Although it was initially designed to study large distributed computing environments such as grids, we have recently applied it to performance prediction of HPC configurations.

- Finite difference methods are, in general, well suited to execution on parallel machines and are thus commonplace in High Performance Computing. Yet, despite their apparent regularity, they often exhibit load imbalance that damages their efficiency. In [38], we characterize the spatial and temporal load imbalance of Ondes3D, a seismic wave propagation simulator used to conduct regional scale risk assessment. Our analysis reveals that this imbalance originates from the structure of the input data and from low-level CPU optimizations. We then show that the CHARM++ runtime can effectively dynamically rebalance the load by migrating data and computation at the granularity of an MPI rank. We propose a methodology that leverages the capabilities of the SimGrid simulation framework and allows to conduct an experimental study at low computational cost.
- The article [35] summarizes our recent work and developments on SMPI, a flexible simulator of MPI applications. In this tool, we took a particular care to ensure our simulator could be used to produce fast and accurate predictions in a wide variety of situations. Although we did build SMPI on SimGrid whose speed and accuracy had already been assessed in other contexts, moving such techniques to a HPC workload required significant additional effort. Obviously, an accurate modeling of communications and network topology was one of the key to such achievements. Another less obvious key was the choice to combine in a single tool the possibility to do both offline and online simulation.

6.3. Trace and Statistical Analysis

- In [19], we present visual analysis techniques to evaluate the performance of HPC task-based applications on hybrid architectures. Our approach is based on composing modern data analysis tools (pjdump, R, ggplot2, plotly), enabling an agile and flexible scripting framework with minor development cost. We validate our proposal by analyzing traces from the full-fledged implementation of the Cholesky decomposition available in the MORSE library running on a hybrid (CPU/GPU) platform. The analysis compares two different workloads and three different task schedulers from the StarPU runtime system. Our analysis based on composite views allows to identify allocation mistakes, priority problems in scheduling decisions, GPU tasks anomalies causing bad performance, and critical path issues.
- Media events are an area of major concern for the science of territory, with a combination of empirical, methodological and theoretical fields of research. The paper [22] presents three variations of increasing complexity around the questions of the application of the concepts of “territory”, “territoriality” and “territorialization” to the description of media events. Each variation is illustrated by recent results from the research project ANR Geomedia on a corpus of international RSS flows produced by newspapers of French, English and Spanish language located in various countries of the world.

6.4. Electricity Markets

The increased penetration of renewable energy sources in existing power systems has led to necessary developments in electricity market mechanisms. Most importantly, renewable energy generation is increasingly made accountable for deviations between scheduled and actual energy generation. However, there is no mechanism to enforce accountability for the additional costs induced by power fluctuations. These costs are socialized and eventually supported by electricity customers. In [1], we propose some metrics for assessing the contribution of all market participants to power regulation needs, as well as an attribution mechanism for fairly redistributing related power regulation costs. We discuss the effect of various metrics used by the attribution mechanisms, and we illustrate, in a game-theoretical framework, their consequences on the strategic behavior of market participants. We also illustrate, by using the case of Western Denmark, how these mechanisms may affect revenues and the various market participants.

6.5. Power control in random wireless networks

Ever since the early development stages of wireless networks, the importance of radiated power has made power control an essential component of network design. In [13], we analyzed the problem of power control in large, random wireless networks that are obtained by "erasing" a finite fraction of nodes from a regular d-dimensional lattice of N transmit-receive pairs. Drawing on tools and ideas from statistical physics, we showed that this problem can be mapped to the Anderson impurity model for diffusion in random media; in this way, by employing the so-called *coherent potential approximation* (CPA) method, we calculated the average power in the system (and its variance) for 1-D and 2-D networks. In this regard, even though infinitely large systems are always unstable beyond a critical value of the users' SINR target, finite systems remain stable with high probability even beyond this critical SINR threshold. We calculated this probability by analyzing the density of low lying eigenvalues of an associated random Schrödinger operator, and we showed that the network can exceed this critical SINR threshold by a factor of at least $O((\log N)^{-2/d})$ before undergoing a phase transition to the unstable regime.

6.6. Energy efficiency in wireless networks

[6] The recent increase in the use of wireless networks for video transmission has led to the increase in the use of rate-adaptive protocols to maximize the resource utilization and increase the efficiency in the transmission. However, a number of these protocols lead to interactions among the users that are subjective in nature and affect the overall performance. In [6], we analyzed the interplay between the wireless network and video transmission dynamics in the light of subjective perceptions of the end users in their interactions – specifically, the trade-off between maximizing the quality of service (QoS) or quality of experience (QoE) and minimizing the transmission cost. By using methods from game theory, we derived an optimized transmission scheme that allows the efficient use of traditional protocols by taking into account the subjective interactions that occur in practical scenarios.

6.7. Cognitive radio and beyond

In cognitive radio networks, secondary (unlicensed) users (SUs) can access the spectrum opportunistically, whenever they sense an opening by the network's primary (licensed) users (PUs). In [7], we analyzed the minimization of overall power consumption over several orthogonal frequency bands under constraints on the minimum quality of service (QoS) and maximum peak and average interference to the network's PUs. To that end, we proposed a projected sub-gradient algorithm which quickly converges to an optimal configuration if the users' channels are fast fading.

Despite such benefits, the conventional cognitive radio network (CCRN) paradgim is not particularly attractive for opportunistic spectrum access because the network's PUs can recapture SU channels at will, thus interrupting the transmission of the latter. To address this crucial limitation, we proposed in [24] a semi-cognitive radio network (SCRN) paradigm where PUs are constrained to first use any free channels before being allowed to capture channels that are in use by SUs. These constraints slightly degrade the performance of the network's PUs, but *a*) they offer remarkable performance improvements to the network's SUs; and *b*) they can be compensated by imposing a monetary (or other) penalty to the network's secondary owners. In [24], we provided a game-theoretic analysis of the performance trade-offs involved for both the PUs and the SUs, and we derived both centralized and distributed learning algorithms that allow the system control process to converge to a stable state.

6.8. Online resource allocation in dynamic wireless networks

The vast majority of works on wireless resource allocation (spectrum, power, etc.) has focused on two limit cases: In the *static regime*, the attributes of the network are assumed effectively static and the system's optimality analysis relies on techniques from (static) optimization. On the other hand, in the so-called *stochastic regime*, the network is assumed to evolve randomly following some fixed probability law, and the allocation of wireless resources is optimized using tools from stochastic optimization and control. In practical wireless networks however, both assumptions fail because of factors that introduce an unpredictable variability to the system (such as user mobility, users going arbitrarily on- and off-line, etc.).

The works [15], [27], [28] treat this problem by providing no-regret learning algorithms for single-user rate maximization and power control in multi-carrier cognitive radio and Internet of Things networks. The extension of these works to multi-antenna systems was carried out in [44], where we derived a matrix exponential learning algorithm for dynamic power allocation and control in time-varying MIMO systems. Building on this, we also showed in [8] that regret minimization techniques can also be applied to the much more challenging problem of energy efficiency maximization in dynamic networks – i.e. the maximization of successfully received bits per Watt of transmitted power in environments that fluctuate unpredictably over time. Finally, as was shown in [39], [23], [9], these unilateral performance gains also extend to large networks comprising hundreds (or even thousands) of users: there, the proposed matrix exponential learning algorithm converges to a stable state within a few iterations, even for very large of antennas and subcarriers.

6.9. Adaptive multi-path routing

Routing plays a crucial part in the efficient operation of packet-switched data networks, especially with regard to latency reduction and energy efficiency. However, in addition to being distributed (so as to cope with the prolific size of today's networks), optimized routing schemes must also be able to adapt to changes in the underlying network (e.g. due to variations in traffic demands, link quality, etc.).

First, to address the issue of latency reduction, we provided in [32] an adaptive multi-flow routing algorithm to select end-to-end paths in packet-switched networks. The algorithm is based only on local information, so it is suitable for distributed implementation; furthermore, it provides guarantees that the network configuration converges to a stable state and exhibits several robustness properties that make it suitable for use in dynamic real-life networks (such as robustness to measurement errors, outdated information and update desynchronization).

Concerning energy efficiency, [41] examines the problem of routing in optical networks with the aim of minimizing traffic-driven power consumption. To tackle this, [41] proposed a pricing scheme which, combined with a distributed learning method based on the Boltzmann distribution of statistical mechanics, exhibits remarkable operation properties even under uncertainty. Specifically, the long-term average of the network's power consumption converges quickly to its minimum value (in practice, within a few iterations of the algorithm), and this convergence remains robust in the face of uncertainty of arbitrarily high magnitude.

6.10. Learning in finite games

One of the most widely used algorithms for learning in finite games is the so-called *best response algorithm* (BRA); nonetheless, even though sevaral worst-case bounds are known for its convergence time, the algorithm's performance in typical game-theoretic scenarios seems to be far better than these worst-case bounds suggest. In [26], [18], [25], [31], we computed the average execution time of the BR algorithm using Markov chain coupling techniques that recast the average execution time of this discrete algorithm as the solution of an ordinary differential equation. In so doing, we showed that the worst-case complexity of the BR algorithm in a potential game with N players and A actions per player is $AN(N-1)$, while its average complexity over random potential games is $O(N)$, independently of A.

In [34], we also studied the convergence rate of the HEDGE algorithm (which, contrary to the BR algorithm, leads to no regret even in adversarial settings). Motivated by applications to data networks where fast convergence is essential, we analyzed the problem of learning in generic N-person games that admit Nash equilibria in pure strategies. Despite the (unbounded) uncertainty in the players' observations, we show that hedging eliminates dominated strategies (a.s.) and, with high probability, it converges locally to pure Nash equilibria at the exponential rate $O(\exp(-c\sum_{j=1}^{t}\gamma_j))$, where γ_j is the algorithm s step size.

These results are strongly related to the long-term rationality properties (elimination of dominated strategies, convergence to pure Nash equilibria and evolutionarily stable states, etc.) of an underlying class of game dynamics based on regularization and Riemannian geometry. Specifically, in [42], we introduced a class of evolutionary game dynamics whose defining element is a state-dependent geometric structure on the set of population states. When this geometric structure satisfies a certain integrability condition, the resulting dynamics preserve many further properties of the replicator and projection dynamics and are equivalent to a class of reinforcement learning dynamics studied in [10]. Finally, as we showed in [2], these properties also hold even in the presence of noise, i.e. when the players only have noisy observations of their payoff vectors.

6.11. Learning in games with continuous action spaces

A key limitation of existing game-theoretic learning algorithms is that they invariably revolve around games with a finite number of actions per players. However, this assumption is often unrealistic (especially in network-based applications of game theory), a factor which severely limits the applicability of learning techniques in real-life problems.

To address this issue, we studied in [14] a class of control problems that can be formulated as potential games with continuous action sets, and we proposed an actor-critic reinforcement learning algorithm that provably converges to equilibrium in said class. The method employed is to analyse the learning process under study through a mean-field dynamical system that evolves in an infinite-dimensional function space (the space of probability distributions over the players' continuous controls). To do so, we extend the theory of finite-dimensional two-timescale stochastic approximation to an infinite-dimensional, Banach space setting, and we proved that the continuous dynamics of the process converge to equilibrium in the case of potential games. These results combine to give a provably-convergent learning algorithm in which players do not need to keep track of the controls selected by the other agents.

Finally, to address cases where mixing over a continuum of actions is unrealistic, we examined in [40] the convergence properties of a class of learning schemes for N-person games with continuous action spaces based on a continuous optimization technique known as "dual averaging". To study this multi-agent, pure-strategy learning process, we introduced the notion of *variational stability* (VS), and we showed that stable equilibria are locally attracting with high probability whereas globally stable states are globally attracting with probability 1. Finally, we examined the scheme's convergence speed and we showed that if the game admits a strict equilibrium and the players' mirror maps are surjective, then, with high probability, the process converges to equilibrium in a finite number of steps, no matter the level of uncertainty in the players' observations (or payoffs).

6.12. Stochastic optimization

A key feature of modern data networks is their distributed nature and the stochasticity surrounding users and their possible actions. To account for these issues in a general optimization context, we proposed in [4] a distributed, asynchronous algorithm for stochastic semidefinite programming which is a stochastic approximation of the continous-time matrix exponential scheme derived in [9]. This algorithm converges almost surely to an ϵ-approximation of an optimal solution requiring only an unbiased estimate of the gradient of the problem's stochastic objective. When applied to throughput maximization in wireless multiple-input and multiple-output (MIMO) systems, the proposed algorithm retains its convergence properties under a wide array of mobility impediments such as user update asynchronicities, random delays and/or ergodically changing channels.

More generally, in view of solving convex optimization problems with noisy gradient input, we also analyzed in [43] the asymptotic behavior of gradient-like flows that are subject to stochastic disturbances. For concreteness, we focused on the widely studied class of mirror descent methods for constrained convex programming and we examined the dynamics' convergence and concentration properties in the presence of noise. In the small noise limit, we showed that the dynamics converge to the solution set of the underlying problem with probability 1. Otherwise, in the case of persistent noise, we estimated the measure of the dynamics' long-run concentration around interior solutions and their convergence to boundary solutions that are sufficiently "robust". Finally, we showed that a rectified variant of the method with a decreasing sensitivity parameter converges irrespective of the magnitude of the noise or the structure of the underlying convex program, and we derived an explicit estimate for its rate of convergence.

6.13. Benchmarking

In modern High Performance Computing architectures, the memory subsystem is a common performance bottleneck. When optimizing an application, the developer has to study its memory access patterns and adapt accordingly the algorithms and data structures it uses. The objective is twofold: on one hand, it is necessary to avoid missuses of the memory hierarchy such as false sharing of cache lines or contention in a NUMA interconnect. On the other hand, it is essential to take advantage of the various cache levels and the memory hardware prefetcher. Still, most profiling tools focus on CPU metrics. The few of them able to provide an overview of the memory patterns involved by the execution rely on hardware instrumentation mechanisms and have two drawbacks. The first one is that they are based on sampling which precision is limited by hardware capabilities. The second one is that they trace a subset of all the memory accesses, usually the most frequent, without information ab out the other ones. In [30] we present Moca, an efficient tool for the collection of complete spatio-temporal memory traces. Moca is based on a Linux kernel module and provides a coarse grained trace of a superset of all the memory accesses performed by an application over its addressing space during the time of its execution. The overhead of Moca is reasonable when taking into account the fact that it is able to collect complete traces which are also more precise than the ones collected by comparable tools.

Benchmarking has proven to be crucial for the investigation of the behavior and performances of a system. However, the choice of relevant benchmarks still remains a challenge. To help the process of comparing and choosing among benchmarks, in [33] we propose a solution for automatic benchmark profiling. It computes unified benchmark profiles reflecting benchmarks' duration, function repartition, stability, CPU efficiency, parallelization and memory usage. It identifies the needed system information for profile computation, collects it from execution traces and produces profiles through efficient and reproducible trace analysis treatments. The paper presents the design, implementation and the evaluation of the approach.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry: Alcatel Lucent-Bell

A common laboratory between Inria and the Alcatel Lucent-Bell Labs was created in early 2008 and consists on three research groups (ADR). POLARIS leads the ADR on self-optimizing networks (SELFNET). The researchers involved in this project are Bruno Gaujal and Panayotis Mertikopoulos.

- Contract with Schneider Electric (2015–2018). Distributed optimization in electrical distribution networks. Associated to a CIFRE PhD grant (Benoît Vinot, started in 4/2015). Partners: Inria (Polaris), Schneider Electric, G2ELab.

7.2. National Initiatives

7.2.1. ANR

- *GAGA (2014–2017)*
  GAGA is an ANR starting grant (JCJC) whose aim is to explore the Geometric Aspects of GAmes. The GAGA team is spread over three different locations in France (Paris, Toulouse and Grenoble), and is coordinated by Vianney Perchet (ENS Cachan). Its aim is to perform a systematic study of the geometric aspects of game theory and, in so doing, to establish new links between application areas that so far appeared unrelated (such as the use of Hessian Riemannian optimization techniques in wireless communication networks).

- *MARMOTE (2013–2016)*
  Partners: Inria Sophia (MAESTRO), Inria Rocquencourt (DIOGEN), Université Versailles-Saint-Quentin (PRiSM lab), Telecom SudParis (SAMOVAR), Université Paris-Est Créteil (*Spécification et vérification de systèmes*), Université Pierre-et-Marie-Curie/LIP6.
  The project aims at realizing a software prototype dedicated to Markov chain modeling. It gathers seven teams that will develop advanced resolution algorithms and apply them to various domains (reliability, distributed systems, biology, physics, economy).
- *NETLEARN (2013–2017)*
  Partners: Université Versailles – Saint-Quentin (PRiSM lab), Université Paris Dauphine, Inria Grenoble (POLARIS), Institut Mines–Telecom (Telecom ParisTech), Alcatel–Lucent Bell Labs (ALBF), and Orange Labs.
  The main objective of the project is to propose a novel approach of distributed, scalable, dynamic and energy efficient algorithms for mobile network resource management. This new approach relies on the design of an orchestration mechanism of a portfolio of algorithms. The ultimate goal of the proposed mechanism is to enhance the user experience, while at the same time ensuring the more efficient utilization of the operator's resources.
- *ORACLESS (2016–2021)*
  ORACLESS is an ANR starting grant (JCJC) coordinated by Panayotis Mertikopoulos. The goal of the project is to develop highly adaptive resource allocation methods for wireless communication networks that are provably capable of adapting to unpredictable changes in the network. In particular, the project will focus on the application of online optimization and online learning methodologies to multi-antenna systems and cognitive radio networks.
- *ANR SONGS, 2012–2016*. Partners: Inria Nancy (Algorille), Inria Sophia (MASCOTTE), Inria Bordeaux (CEPAGE, HiePACS, RunTime), Inria Lyon (AVALON), University of Strasbourg, University of Nantes.

  The last decade has brought tremendous changes to the characteristics of large scale distributed computing platforms. Large grids processing terabytes of information a day and the peer-to-peer technology have become common even though understanding how to efficiently exploit such platforms still raises many challenges. As demonstrated by the USS SimGrid project funded by the ANR in 2008, simulation has proved to be a very effective approach for studying such platforms. Although even more challenging, we think the issues raised by petaflop/exaflop computers and emerging cloud infrastructures can be addressed using similar simulation methodology.

  The goal of the SONGS project (Simulation of Next Generation Systems) is to extend the applicability of the SimGrid simulation framework from grids and peer-to-peer systems to clouds and high performance computation systems. Each type of large-scale computing system will be addressed through a set of use cases and led by researchers recognized as experts in this area. Any sound study of such systems through simulations relies on the following pillars of simulation methodology: Efficient simulation kernel; Sound and validated models; Simulation analysis tools; Campaign simulation management. Such aspects are also addressed in the SONGS project.

7.2.2. National Organizations

- Jean-Marc Vincent is member of the scientific committees of the CIST (Centre International des Sciences du Territoire).
- *REAL.NET (2016)*
  REAL.NET is a CNRS PEPS starting grant (JCJC) coordinated by Panayotis Mertikopoulos. Its objective is to provide dynamic control methodologies for nonstationary stochastic optimization problems that arise in wireless communication networks.

8. Partnerships and Cooperations

8.1. European Initiatives

8.1.1. FP7 & H2020 Projects

8.1.1.1. Mont-Blanc 2

Program: FP7 Programme

Project acronym: Mont-Blanc 2

Project title: Mont-Blanc: European scalable and power efficient HPC platform based on low-power embedded technology

Duration: October 2013 - September 2016

Coordinator: BSC (Barcelone)

Other partners: BULL - Bull SAS (France), STMicroelectronics - (GNB SAS) (France), ARM - (United Kingdom), JUELICH - (Germany), BADW-LRZ - (Germany), USTUTT - (Germany), CINECA - (Italy), CNRS - (France), Inria - (France), CEA - (France), UNIVERSITY OF BRISTOL - (United Kingdom), ALLINEA SW LIM - (United Kingdom)

Abstract: Energy efficiency is already a primary concern for the design of any computer system and it is unanimously recognized that future Exascale systems will be strongly constrained by their power consumption. This is why the Mont-Blanc project has set itself the following objective: to design a new type of computer architecture capable of setting future global High Performance Computing (HPC) standards that will deliver Exascale performance while using 15 to 30 times less energy. Mont-Blanc 2 contributes to the development of extreme scale energy-efficient platforms, with potential for Exascale computing, addressing the challenges of massive parallelism, heterogeneous computing, and resiliency. Mont-Blanc 2 has great potential to create new market opportunities for successful EU technology, by placing embedded architectures in servers and HPC.

The Mont-Blanc 2 proposal has 4 objectives:
1. To complement the effort on the Mont-Blanc system software stack, with emphasis on programmer tools (debugger, performance analysis), system resiliency (from applications to architecture support), and ARM 64-bit support.
2. To produce a first definition of the Mont-Blanc Exascale architecture, exploring different alternatives for the compute node (from low-power mobile sockets to special-purpose high-end ARM chips), and its implications on the rest of the system.
3. To track the evolution of ARM-based systems, deploying small cluster systems to test new processors that were not available for the original Mont-Blanc prototype (both mobile processors and ARM server chips).
4. To provide continued support for the Mont-Blanc consortium, namely operations of the Mont-Blanc prototype, and hands-on support for our application developers

8.1.1.2. QUANTICOL

Program: The project is a member of Fundamentals of Collective Adaptive Systems (FOCAS), a FET-Proactive Initiative funded by the European Commission under FP7.

Project acronym: QUANTICOL

Project title: A Quantitative Approach to Management and Design of Collective and Adaptive Behaviours

Duration: 04 2013 – 03 2017

Coordinator: Jane Hillston (University of Edinburgh, Scotland)

Other partners: University of Edinburgh (Scotland); Istituto di Scienza e Tecnologie della Informazione (Italy); IMT Lucca (Italy) and University of Southampton (England).

Abstract: The main objective of the QUANTICOL project is the development of an innovative formal design framework that provides a specification language for collective adaptive systems (CAS) and a large variety of tool-supported, scalable analysis and verification techniques. These techniques will be based on the original combination of recent breakthroughs in stochastic process algebras and associated verification techniques, and mean field/continuous approximation and control theory. Such a design framework will provide scalable extensive support for the verification of developed models, and also enable and facilitate experimentation and discovery of new design patterns for emergent behaviour and control over spatially distributed CAS.

8.1.1.3. HPC4E

Title: HPC for Energy

Program: H2020

Duration: 01 2016 – 01 2018

Coordinator: Barcelona Supercomputing Center

Inria contact: Stephane Lanteri

Other partners:

- Europe: Lancaster University (ULANC), Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Repsol S.A. (REPSOL), Iberdrola Renovables Energía S.A. (IBR), Total S.A. (TOTAL).
- Brazil: Fundação Coordenação de Projetos, Pesquisas e Estudos Tecnoclógicos (COPPE), National Laboratory for Scientific Computation (LNCC), Instituto Tecnológico de Aeronáutica (ITA), Petroleo Brasileiro S. A. (PETROBRAS), Universidade Federal do Rio Grande do Sul (INF-UFRGS), Universidade Federal de Pernambuco (CER-UFPE)

Abstract: The main objective of the HPC4E project is to develop beyond-the-state-of-the-art high performance simulation tools that can help the energy industry to respond future energy demands and also to carbon-related environmental issues using the state-of-the-art HPC systems. The other objective is to improve the cooperation between energy industries from EU and Brazil and the cooperation between the leading research centres in EU and Brazil in HPC applied to energy industry. The project includes relevant energy industrial partners from Brazil and EU, which will benefit from the project's results. They guarantee that TRL of the project technologies will be very high. This includes sharing supercomputing infrastructures between Brazil and EU. The cross-fertilization between energy-related problems and other scientific fields will be beneficial at both sides of the Atlantic.

8.1.2. Collaborations with Major European Organizations

EPFL: Laboratoire pour les communications informatiques et leurs applications 2, Institut de systèmes de communication ISC, Ecole polytechnique fédérale de Lausanne (Switzerland). We collaborate with Jean-Yves Leboudec (EPFL) and Pierre Pinson (DTU) on electricity markets.

TU Wien: Research Group Parallel Computing, Technische Universität Wien (Austria). We collaborate with Sascha Hunold on experimental methodology and reproducibility of experiments in HPC. In particular we co-organize the REPPAR workshop on "Reproducibility in Parallel Computing".

BSC (Barcelona): Barcelona Supercomputer Center (Spain). We collaborate with the performance evaluation group through the HPC4E project, the Mont-blanc 2 project, and the JLESC.

University of Edinburgh and Istituto di Scienza e Tecnologie della Informazione: we strongly collaborate through the Quanticol European project.

8.2. International Initiatives

8.2.1. Inria International Labs

8.2.1.1. North America

- JLESC (former JLPC) (Joint Laboratory for Extreme-Scale Computing) with University of University of Illinois Urbana Champaign, Argonne Nat. Lab and BSC. Several members of POLARIS are partners of this laboratory, and have done several visits to Urbana-Champaign or NCSA.

8.2.2. Inria Associate Teams not involved in an Inria International Labs

8.2.2.1. EXASE

Title: Exascale Computing Scheduling and Energy

International Partner (Institution - Laboratory - Researcher):

Universidade Federal do Rio Grande do Sul (Brazil) - INF (INF) - Nicolas MAILLARD

Start year: 2014

See also: https://team.inria.fr/exase/

The main scientific goal of this collaboration for the three years is the development of state-of-the-art energy-aware scheduling algorithms for exascale systems. Three complementary research directions have been identified : (1) Fundamentals for the scaling of schedulers: develop new scheduling algorithms for extreme exascale machines and use existing workloads to validate the proposed scheduling algorithms (2) Design of schedulers for large-scale infrastructures : propose energy-aware schedulers in large-scale infrastructures and develop adaptive scheduling algorithms for exascale machines (3) Tools for the analysis of large scale schedulers : develop aggregation methodologies for scheduler analysis to propose synthetic visualizations for large traces analysis and then analyze schedulers and energy traces for correlation analysis

8.2.3. Inria International Partners

8.2.3.1. Declared Inria International Partners

- POLARIS has strong connections with both UFRGS (Porto Alegre, Brazil) and USP (Sao Paulo, Brazil). The creation of the LICIA common laboratory (see next section) has made this collaboration even tighter.
- POLARIS has strong bounds with the University of Illinois Urbana Champaign and Barcelona Supercompter Center, within the (Joint Laboratory on Petascale Computing, see previous section).

8.2.4. Participation in Other International Programs

8.2.4.1. South America

- *LICIA:* The CNRS, Inria, the Universities of Grenoble, Grenoble INP, and Universidade Federal do Rio Grande do Sul have created the LICIA (*Laboratoire International de Calcul intensif et d'Informatique Ambiante*). LICIA's main research themes are high performance computing, language processing, information representation, interfaces and visualization as well as distributed systems. Jean-Marc Vincent is the director of the laboratory on the French side and visited Porto Alegre for two weeks in November 2016.

  More information can be found at http://www.inf.ufrgs.br/licia/.
- *ECOS-Sud:* POLARIS is a member of the Franco-Chilean collaboration network LEARN with CONICYT (the Chilean national research agency), formed under the ECOS-Sud framework. The main research themes of this network is the application of continuous optimization and game-theoretic learning methods to traffic routing and congestion control in data networks. Panayotis Mertikopoulos was an invited researcher at the University of Chile in October 2016.

  More information can be found at http://www.conicyt.cl/pci/2016/02/11/programa-ecos-conicyt-adjudica-proyectos-para-el-ano-2016.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

- Matthieu Jonckeere (Buenos Aires University) visited for 3 weeks.
- Mario Bravo (University of Santiago, Chile) visited POLARIS for 1 week in Feb. 2016.
- Mathias Staudigl (Maastricht University) visited POLARIS for 2 weeks in July 2016.

8.3.2. Visits to International Teams

8.3.2.1. Sabbatical programme

- Florence Perronnin spent one year in sabbatical leave (rachat de service) at the Université Versailles-Saint-Quentin (DAVID lab)

8.3.2.2. Research Stays Abroad

- Panayotis Mertikopoulos was an invited professor at the University of Athens, Athens, Greece, for four months (March–June 2016).
- Panayotis Mertikopoulos was an invited professor at LUISS Guido Carli University, Rome, Italy, for one month (Sept. 2016).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organization

9.1.1.1. General Chair, Scientific Chair

- Panayotis Mertikopoulos was the general co-chair of the GEL 2016 international workshop on Geometry, Evolution and Learning in Games

9.1.1.2. Member of the Organizing Committees

- Nicolas Gast served as Publicity Chair for ACM Sigmetrics 2016
- Jean-Marc Vincent served in the Organization Committee of the "Decision-making and Optimization under Uncertainty" summer school for young researchers, and the joint LICIA-EXASE workshop

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

Arnaud Legrand was chair of the performance evaluation track of Europar 2016.

9.1.2.2. Member of the Conference Program Committees

The members of the team regularly review numerous papers for international conferences.

- Bruno Gaujal and Nicolas Gast were members of the Technical Program Committee of ACM Sigmetrics 2016
- Nicolas Gast was a member of the Technical Program Committee of ACM E-Energy 2016.
- Panayotis Mertikopoulos was a member of the Technical Program Committee of VALUETOOLS 2016.
- E. Veronica Belmega was a member of the Technical Program Committee of IEEE Globecom 2016, IEEE ICC 2016, IEEE BlackSeaComm 2016, and IEEE WCNC 2016
- Jean-Marc Vincent was a member of the Technical Program Committees of ASMTA 2016, SimulTech 2016 and EPEW 2016
- Arnaud Legrand was a member of the Technical Program Committees of HiPC 2016, ICPP 2016, and COMPAS 2016.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- E. Veronica Belmega is an executive editor for the Transactions on Emerging Telecommunications Technologies.

9.1.3.2. Reviewer - Reviewing Activities

The members of the POLARIS team regularly review articles for JPDC, DAM, IEEE Transactions on Networking/Automatic Control/Cloud Computing/Parallel and Distributed Computing/Information Theory/Signal Processing/Wireless Communications, SIAM Journal on Optimization/Control and Optimization, and others.

9.1.4. Invited Talks

- Bruno Gaujal was a plenary speaker at ECQT 2016 (European Conference on Queuing Theory)
- Bruno Gaujal was a plenary speaker at "Journées SDA2 2016" (Groupe de Travail "Systèmes Dynamiques, Automates et Algorithmes" of the CNRS thematic group GDR Informatique Mathématique.
- Panayotis Mertikopoulos was a plenary speaker at the ADGO 2016 workshop on Algorithms and Dynamics for Games and Optimization
- Panayotis Mertikopoulos gave an invited tutorial on "*Online Optimization for Wireless Communications*" at the Orange workshop on Learning and Networks.
- Panayotis Mertikopoulos gave a two-part tutorial on "*Game Theory, Learning and Cognitive Radio*" at CROWNCOM 2016.

9.1.5. Leadership within the Scientific Community

Arnaud Legrand has organized a series of webinars on reproducible research and whose aim is to introduce the audience to one particular aspect of reproducible research and to illustrate how this aspect can be addressed with state-of-the-art tools. To this end, experts of a given topic are invited and their seminar is screencast so that researchers from other universities can easily follow it. So far, the following topics have been covered:

1. Introduction (reproducible research, challenges, ethic, ...) to reproducible research. Producing replicable articles and managing a laboratory notebook.
2. Controlling your experimental environment
3. Numerical reproducibility
4. Logging and backing up your work: a not so short introduction to git for research
5. Preserving software: ensuring availability and tracking provenance
6. Reproducible science in bioinformatics: current status, solutions and research opportunities

All the corresponding videos and materials are available at the following address: https://github.com/alegrand/RR_webinars/

9.1.6. Scientific Expertise

- E. Veronica Belmega was a member of the jury for the GRETSI–GdR ISIS thematic research group "best thesis" award
- Panayotis Mertikopoulos is a member of the steering committee (*comité de liaison*) of the optimization and decision theory group of the French Society for Industrial and Applied Mathematics (SMAI).
- Jean-Marc Vincent is a member of the scientific committees of the CIST (Centre International des Sciences du Territoire)

9.1.7. Research Administration

- Bruno Gaujal is member of the "bureau du LIG" (Laboratoire d'informatique de Grenoble)

- Bruno Gaujal is member of the “bureau du CP” of Inria Grenoble.
- Bruno Gaujal was member of the CR2 admissibility jury of Inria-Grenoble.
- Panayotis Mertikopoulos serves as the graduate students liaison (*chargé de mission doctorants*) for the Laboratoire d’Informatique de Grenoble
- Arnaud Legrand was coordinator of the Inria evaluation of the “Distributed and High Performance Computing” theme.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Bruno Gaujal and Nicolas Gast, *“Advanced Performance Evaluation”*, 18h (M2), ENSIMAG

Master: Guillaume Huard, *“Conception des Systèmes d’Exploitation”* (M1), Université Grenoble-Alpes

Master: Arnaud Legrand and Jean-Marc Vincent, *“Scientific Methodology and Performance Evaluation”*, 15h (22.5h) M2, M2R MOSIG

Master: Arnaud Legrand, *“Parallel Systems”*, 21h (31.5h) M2R, M2R Mosig.

Master: Arnaud Legrand, *“Scientific Methodology and Performance Evaluation”*, 24h (36h), ENS Lyon.

Master: Panayotis Mertikopoulos, *“Selected Topics in the Theory of Stochastic Processes”*, 16h M2, University of Athens, Athens, Greece

Master: Florence Perronnin, *“Simulation”*, M1, Université Versailles – Saint-Quentin

Master: Florence Perronnin, *“Probabilités–Simulation”*, RICM4 Polytech Grenoble

Master: Arnaud Legrand and Jean-Marc Vincent, Probability and simulation, performance evaluation 72 h, (M1), RICM, Polytech Grenoble.

Master: Jean-Marc Vincent, Mathematics for computer science, 18 h , (M1) Mosig.

DU: Jean-Marc Vincent, *“Informatique et sciences du numérique”*, 20 h, (Professeurs de lycée).

9.2.2. Supervision

Post-Doc: Angelika Studeny, Université Grenoble-Alpes (Jean-Marc Vincent)

PhD: David Beniamine, *“Analyse du comportement mémoire d’application paralleles de calcul scientifique”*, Université Grenoble-Alpes, 05/12/2016 (Guillaume Huard)

PhD: Joaquim Assunção, *“Fitting techniques to knowledge discovery through stochastic models”*, 08/12/2016 (Jean-Marc Vincent)

PhD in progress: Alexandre Marcastel, *“Online resource allocation in dynamic wireless networks”*, 10/2016 (Panayotis Mertikopoulos, E. V. Belmega)

PhD in progress: Stéphane Durand, *“Game theory and control in distributed systems”* (Bruno Gaujal)

PhD in progress: Stephan Plassart, *“Optimization of critical embedded systems”* (Bruno Gaujal)

PhD in progress: Baptiste Jonglez, *“Diversity exploitation in communication networks”* (Bruno Gaujal)

PhD in progress: Umar Ozeer, OrangeLabs, 12/2016 (Jean-Marc Vincent)

PhD in progress: Christian Heinrich, *“Performance Evaluation of HPC Systems Through Simulation”*, 12/2015 (Arnaud Legrand)

PhD in progress: Rafael Keller Tesser, *“Performance Evaluation of Dynamic Load Balancing for Legacy Iterative Parallel Applications*”, 12/2015 (Arnaud Legrand, Cotutelle with Philippe Navaux from UFRGS)

PhD in progress: Vinicius Garcia Pinto, *"Visual Performance Analysis of HPC applications running over Dynamic Task-based Runtimes"*, 12/2015 (Arnaud Legrand, Cotutelle with Nicolas Maillard and Lucas Schnorr from UFRGS)

9.2.2.1. Internships

- Jean-Marc Vincent supervised the eng. internship of Benjamin Briot (02–03/2016)
- Jean-Marc Vincent supervised the master internship Mathieu Baille (Magistère + TER M1)
- Guillaume Huard supervised the Licence internship of Michael Picard
- Guillaume Huard supervised the Licence internship of Loic Poncet
- Arnaud Legrand and Vincent Danjean supervised the eng. internship of Florian Popek (06–08/2016)
- Arnaud Legrand supervised the master internship of Steven Quinito Masnada (02–08/2016) in collaboration with the CORSE team (Brice Videau and Frédéric Desprez)

9.2.3. Juries

- E. Veronica Belmega was a member of the jury for the PhD defense of Kenza Hamidouche (12/2016, Supélec)
- Bruno Gaujal was a reviewer for the HDR defense of Bruno Scherrer
- Bruno Gaujal was a member of the jury for the HDR defense of Patrick Loiseau
- Arnaud Legrand was reviewer for the PhD defense of Rafife Nheili

9.3. Popularization

The POLARIS team is actively involved in various scientific popularization activities. In addition to participating in the Fête de la Science (Guillaume Huard, Florence Perronnin, Jean-Marc Vincent), the POLARIS team also participates in the organization of the "Conférence Inria". Jean-Marc Vincent has also organized (and/or participated in) several training courses for computer science teachers (at a high school level) and has supervised several MathC2+ trainees.

Jean-Marc Vincent also organized the 2-day-long atelier "Computer Science without Computers" for elementary school students; the atelier was nominated for the Shannon trophy (awarded by the Institut Henri Poincaré).

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journals

[1] F. BONA, N. GAST, J.-Y. LE BOUDEC, P. PINSON, D.-C. TOMOZEI. *Attribution mechanisms for ancillary service costs induced by variability in power delivery*, in "IEEE Transactions on Power Systems", 2016, 10 p. [*DOI :* 10.1109/TPWRS.2016.2598760], https://hal.inria.fr/hal-01403913

[2] M. BRAVO, P. MERTIKOPOULOS. *On the robustness of learning in games with stochastically perturbed payoff observations*, in "Games and Economic Behavior", June 2016, https://hal.inria.fr/hal-01098494

[3] N. GAST, B. VAN HOUDT. *Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms*, in "Queueing Systems", June 2016, This paper is an extended version of the ACM SIGMETRICS 2015 paper that is accessible at https://hal.inria.fr/hal-01143838 [*DOI :* 10.1007/s11134-016-9487-9], https://hal.inria.fr/hal-01334354

[4] B. GAUJAL, P. MERTIKOPOULOS. *A stochastic approximation algorithm for stochastic semidefinite programming*, in "Probability in the Engineering and Informational Sciences", 2016, vol. 30, n^{o} 3, pp. 431-454, https://hal.archives-ouvertes.fr/hal-01382288

[5] J.-P. GAYON, G. MASSONNET, C. RAPINE, G. STAUFFER. *Constant approximation algorithms for the one warehouse multiple retailers problem with backlog or lost-sales*, in "European Journal of Operational Research", April 2016, vol. 250, n^{o} 1, pp. 155 - 163 [*DOI :* 10.1016/J.EJOR.2015.10.054], https://hal.archives-ouvertes.fr/hal-01284347

[6] S. GUPTA, E. V. BELMEGA, M. Á. VÁZQUEZ-CASTRO. *Game theoretical analysis of rate adaptation protocols conciliating QoS and QoE*, in "EURASIP Journal on Wireless Communications and Networking (EURASIP JWCN)", 2016, vol. 2016, 75 p. [*DOI :* 10.1186/S13638-016-0569-5], https://hal.archives-ouvertes.fr/hal-01308581

[7] R. MASMOUDI, E. VERONICA BELMEGA, I. FIJALKOW. *Efficient spectrum scheduling and power management for opportunistic users*, in "EURASIP Journal on Wireless Communications and Networking", April 2016 [*DOI :* 10.1186/S13638-016-0594-4], https://hal.archives-ouvertes.fr/hal-01301779

[8] P. MERTIKOPOULOS, E. V. BELMEGA. *Learning to be green: Robust energy efficiency maximization in dynamic MIMO-OFDM systems*, in "IEEE Journal on Selected Areas in Communications", April 2016, vol. 34, n^{o} 4, pp. 743 - 757, https://hal.archives-ouvertes.fr/hal-01382276

[9] P. MERTIKOPOULOS, A. L. MOUSTAKAS. *Learning in an uncertain world: MIMO covariance matrix optimization with imperfect feedback*, in "IEEE Transactions on Signal Processing", January 2016, vol. 64, n^{o} 1, pp. 5-18, https://hal.archives-ouvertes.fr/hal-01382278

[10] P. MERTIKOPOULOS, W. H. SANDHOLM. *Learning in games via reinforcement and regularization*, in "Mathematics of Operations Research", November 2016, vol. 41, n^{o} 4, pp. 1297-1324, https://hal.archives-ouvertes.fr/hal-01382286

[11] P. MERTIKOPOULOS, W. H. SANDHOLM. *Learning in games via reinforcement learning and regularization*, in "Mathematics of Operations Research", November 2016, 34 pages, 6 figures [*DOI :* 10.1287/MOOR.2016.0778], https://hal.inria.fr/hal-01073491

[12] P. MERTIKOPOULOS, Y. VIOSSAT. *Imitation dynamics with payoff shocks*, in "International Journal of Game Theory", March 2016, vol. 45, n^{o} 1-2, pp. 291-320, https://hal.archives-ouvertes.fr/hal-01382283

[13] A. L. MOUSTAKAS, P. MERTIKOPOULOS, N. BAMBOS. *Power Optimization in Random Wireless Networks*, in "IEEE Transactions on Information Theory", September 2016, vol. 62, n^{o} 9, pp. 5030-5058, https://hal.archives-ouvertes.fr/hal-01382277

[14] S. PERKINS, P. MERTIKOPOULOS, D. S. LESLIE. *Mixed-strategy learning with continuous action sets*, in "IEEE Transactions on Automatic Control", 2016, https://hal.archives-ouvertes.fr/hal-01382280

Invited Conferences

[15] A. MARCASTEL, E. V. BELMEGA, P. MERTIKOPOULOS, I. FIJALKOW. *Interference Mitigation via Pricing in Time-Varying Cognitive Radio Systems*, in "International conference on NETwork Games, COntrol and OPtimization", Avignon, France, November 2016, https://hal.archives-ouvertes.fr/hal-01387049

International Conferences with Proceedings

[16] L. BORTOLUSSI, N. GAST. *Mean Field Approximation of Uncertain Stochastic Models*, in "46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2016)", Toulouse, France, June 2016, https://hal.inria.fr/hal-01302416

[17] J. DONCEL, N. GAST, B. GAUJAL. *Are mean-field games the limits of finite stochastic games?*, in "The 18th Workshop on MAthematical performance Modeling and Analysis", Nice, France, June 2016, https://hal.inria.fr/hal-01321020

[18] S. DURAND, B. GAUJAL. *Complexity and Optimality of the Best Response Algorithm in Random Potential Games*, in "Symposium on Algorithmic Game Theory (SAGT) 2016", Liverpool, United Kingdom, September 2016, pp. 40-51 [*DOI :* 10.1007/978-3-662-53354-3_4], https://hal.archives-ouvertes.fr/hal-01404643

[19] V. GARCIA PINTO, L. STANISIC, A. LEGRAND, L. MELLO SCHNORR, S. THIBAULT, V. DANJEAN. *Analyzing Dynamic Task-Based Applications on Hybrid Platforms: An Agile Scripting Approach*, in "3rd Workshop on Visual Performance Analysis (VPA)", Salt Lake City, United States, November 2016, Held in conjunction with SC16, https://hal.inria.fr/hal-01353962

[20] N. GAST. *Construction of Lyapunov functions via relative entropy with application to caching*, in "The 18th Workshop on MAthematical performance Modeling and Analysis", Nice, France, June 2016, https://hal.inria.fr/hal-01321017

[21] N. GAST, B. VAN HOUDT. *Asymptotically Exact TTL-Approximations of the Cache Replacement Algorithms LRU(m) and h-LRU*, in "28th International Teletraffic Congress (ITC 28)", Würzburg, Germany, September 2016, https://hal.inria.fr/hal-01292269

[22] C. GRASLAND, R. LAMARCHE-PERRIN, M. LE TEXIER, H. PECOUT, S. DE RUFFRAY, A. STUDENY, J.-M. VINCENT. *Territoire, territorialité et territorialisation des événements médiatiques*, in "CIST2016 - En quête de territoire(s) ?", Grenoble, France, Collège international des sciences du territoire (CIST), March 2016, pp. 207-213, https://hal.archives-ouvertes.fr/hal-01353612

[23] P. MERTIKOPOULOS, E. V. BELMEGA, L. SANGUINETTI. *Distributed learning for resource allocation under uncertainty*, in "GLOBALSIP '16: Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing", 2016, https://hal.archives-ouvertes.fr/hal-01382284

[24] A. S. SHAFIGH, P. MERTIKOPOULOS, S. GLISIC. *A novel dynamic network architecture model based on stochastic geometry and game theory*, in "ICC '16: Proceedings of the 2016 IEEE International Conference on Communications", 2016, https://hal.archives-ouvertes.fr/hal-01382274

Conferences without Proceedings

[25] S. DURAND, B. GAUJAL. *Average complexity of the Best Response Algorithm in Potential Games* , in "Atelier Evalution de Performance 2016", Toulouse, France, March 2016, https://hal.archives-ouvertes.fr/hal-01396906

[26] S. DURAND, B. GAUJAL. *Average complexity of the Best Response Algorithm in Potential Games*, in "17ème conférence dela Société française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2016)", Compiegne, France, February 2016, https://hal.archives-ouvertes.fr/hal-01396902

[27] A. MARCASTEL, E. VERONICA BELMEGA, P. MERTIKOPOULOS, I. FIJALKOW. *Online Interference Mitigation via Learning in Dynamic IoT Environments*, in "IEEE WORKSHOP GLOBECOM 2016", Washington, DC, United States, December 2016, https://hal.archives-ouvertes.fr/hal-01387046

[28] A. MARCASTEL, E. VERONICA BELMEGA, P. MERTIKOPOULOS, I. FIJALKOW. *Online Power Allocation for Opportunistic Radio Access in Dynamic OFDM Networks*, in "2016 IEEE 84th Vehicular Technology Conference (VTC2016-Fall)", Montreal, Canada, September 2016, https://hal.archives-ouvertes.fr/hal-01387044

Scientific Books (or Scientific Book chapters)

[29] L. BORTOLUSSI, N. GAST. *Mean-Field Limits Beyond Ordinary Differential Equations*, in "Formal Methods for the Quantitative Evaluation of Collective Adaptive Systems", M. BERNARDO, R. DE NICOLA, J. HILLSTON (editors), June 2016, vol. Programming and Software Engineering, 16th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2016, Bertinoro, Italy, June 20-24, 2016, Advanced Lectures [*DOI :* 10.1007/978-3-319-34096-8_3], https://hal.inria.fr/hal-01334358

Research Reports

[30] D. BENIAMINE, G. HUARD. *Moca: An efficient Memory trace collection system*, Inria Grenoble Rhône-Alpes, Université de Grenoble, July 2016, n^o RR-8931, 16 p. , https://hal.inria.fr/hal-01342679

[31] S. DURAND, B. GAUJAL. *Complexity and Optimality of the Best Response Algorithm in Random Potential Games*, Inria - Research Centre Grenoble – Rhône-Alpes ; Grenoble 1 UGA - Université Grenoble Alpe, June 2016, n^o RR-8925, 30 p. , https://hal.inria.fr/hal-01330805

[32] B. JONGLEZ, B. GAUJAL. *Distributed Adaptive Routing in Communication Networks*, Inria ; Univ. Grenoble Alpes, October 2016, n^o RR-8959, 25 p. , https://hal.inria.fr/hal-01386832

[33] A. MARTIN, V. MARANGOZOVA-MARTIN. *Automatic Benchmark Profiling through Advanced Trace Analysis*, Inria - Research Centre Grenoble – Rhône-Alpes ; Université Grenoble Alpes ; CNRS, March 2016, n^o RR-8889, https://hal.inria.fr/hal-01292618

Other Publications

[34] J. COHEN, A. HÉLIOU, P. MERTIKOPOULOS. *Exponentially fast convergence to (strict) equilibrium via hedging*, 2016, https://arxiv.org/abs/1607.08863, https://hal.archives-ouvertes.fr/hal-01382290

[35] A. DEGOMME, A. LEGRAND, M. S. MARKOMANOLIS, M. QUINSON, M. S. STILLWELL, F. S. SUTER. *Simulating MPI applications: the SMPI approach*, November 2016, Under review in IEEE TPDS, https://hal.inria.fr/hal-01415484

[36] J. DONCEL, N. GAST, B. GAUJAL. *Mean-Field Games with Explicit Interactions*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01277098

[37] N. GAST, B. GAUJAL. *Computing absorbing times via fluid approximations*, June 2016, working paper or preprint, https://hal.inria.fr/hal-01337950

[38] R. KELLER TESSER, L. MELLO SCHNORR, A. LEGRAND, F. DUPROS, P. O. A. NAVAUX. *Using Simulation to Evaluate and Tune the Performance of Dynamic Load Balancing of an Over-decomposed Geophysics Application*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01391401

[39] P. MERTIKOPOULOS, E. V. BELMEGA, R. NEGREL, L. SANGUINETTI. *Distributed stochastic optimization via matrix exponential learning*, 2016, http://arxiv.org/abs/1606.01190, https://hal.archives-ouvertes.fr/hal-01382285

[40] P. MERTIKOPOULOS. *Learning in concave games with imperfect information*, 2016, https://arxiv.org/abs/1608.07310, https://hal.archives-ouvertes.fr/hal-01382282

[41] P. MERTIKOPOULOS, A. L. MOUSTAKAS, A. TZANAKAKI. *Boltzmann meets Nash: Energy-efficient routing in optical networks under uncertainty*, 2016, https://arxiv.org/abs/1605.01451, https://hal.archives-ouvertes.fr/hal-01382287

[42] P. MERTIKOPOULOS, W. H. SANDHOLM. *Riemannian game dynamics*, 2016, http://arxiv.org/abs/1603.09173, https://hal.archives-ouvertes.fr/hal-01382281

[43] P. MERTIKOPOULOS, M. STAUDIGL. *On the convergence of gradient-like flows with noisy gradient input*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01404586

[44] I. STIAKOGIANNAKIS, P. MERTIKOPOULOS, C. TOUATI. *Adaptive Power Allocation and Control in Time-Varying Multi-Carrier MIMO Networks*, 2016, https://arxiv.org/abs/1503.02155, https://hal.archives-ouvertes.fr/hal-01382289

References in notes

[45] R. M. BADIA, J. LABARTA, J. GIMÉNEZ, F. ESCALÉ. *Dimemas: Predicting MPI Applications Behaviour in Grid Environments*, in "Proc. of the Workshop on Grid Applications and Programming Tools", June 2003

[46] C. BAIER, B. HAVERKORT, H. HERMANNS, J.-P. KATOEN. *Model-checking algorithms for continuous-time Markov chains*, in "Software Engineering, IEEE Transactions on", 2003, vol. 29, n^o 6, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1205180

[47] A. BASU, S. FLEMING, J. STANIER, S. NAICKEN, I. WAKEMAN, V. K. GURBANI. *The State of Peer-to-peer Network Simulators*, in "ACM Computing Survey", August 2013, vol. 45, n^o 4

[48] D. BECKER, F. WOLF, W. FRINGS, M. GEIMER, B. WYLIE, B. MOHR. *Automatic Trace-Based Performance Analysis of Metacomputing Applications*, in "Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International", March 2007, http://dx.doi.org/10.1109/IPDPS.2007.370238

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

[50] G. BIANCHI. *Performance analysis of the IEEE 802.11 distributed coordination function*, in "Selected Areas in Communications, IEEE Journal on", 2000, vol. 18, n^o 3

[51] L. Bobelin, A. Legrand, M. A. G. David, P. Navarro, M. Quinson, F. Suter, C. Thiery. *Scalable Multi-Purpose Network Representation for Large Scale Distributed System Simulation*, in "CCGrid 2012 – The 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Ottawa, Canada, May 2012, 19 p. , https://hal.inria.fr/hal-00650233

[52] L. Bortolussi, J. Hillston. *Model checking single agent behaviours by fluid approximation*, in "Information and Computation", 2015, vol. 242, http://dx.doi.org/10.1016/j.ic.2015.03.002

[53] L. Bortolussi, R. Lanciani. *Model Checking Markov Population Models by Central Limit Approximation*, in "Quantitative Evaluation of Systems", Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013, n° 8054

[54] L. Bortolussi, R. Lanciani. *Fluid Model Checking of Timed Properties*, in "Formal Modeling and Analysis of Timed Systems", Springer International Publishing, 2015

[55] H. Brunst, D. Hackenberg, G. Juckeland, H. Rohling. *Comprehensive Performance Tracking with Vampir 7*, in "Tools for High Performance Computing 2009", M. S. Müller, M. M. Resch, A. Schulz, W. E. Nagel (editors), Springer Berlin Heidelberg, 2010, http://dx.doi.org/10.1007/978-3-642-11261-4_2

[56] A. Busic, B. Gaujal, G. Gorgo, J.-M. Vincent. *PSI2 : Envelope Perfect Sampling of Non Monotone Systems*, in "QEST 2010 - International Conference on Quantitative Evaluation of Systems", Williamsburg, VA, United States, IEEE, September 2010, pp. 83-84, https://hal.inria.fr/hal-00788884

[57] A. Busic, B. Gaujal, F. Perronnin. *Perfect Sampling of Networks with Finite and Infinite Capacity Queues*, in "19th International Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA) 2012", Grenoble, France, K. Al-Begain, D. Fiems, J.-M. Vincent (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7314, pp. 136-149 [*DOI :* 10.1007/978-3-642-30782-9_10], https://hal.inria.fr/hal-00788003

[58] S. Böhm, C. Engelmann. *xSim: The Extreme-Scale Simulator*, in "Proceedings of the International Conference on High Performance Computing and Simulation (HPCS) 2011", Istanbul, Turkey, IEEE Computer Society, Los Alamitos, CA, USA, July 2011

[59] H. Casanova, A. Giersch, A. Legrand, M. Quinson, F. Suter. *Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*, in "Journal of Parallel and Distributed Computing", June 2014, vol. 74, n° 10, pp. 2899-2917 [*DOI :* 10.1016/j.jpdc.2014.06.008], https://hal.inria.fr/hal-01017319

[60] A. Chaintreau, J.-Y. Le Boudec, N. Ristanovic. *The Age of Gossip: Spatial Mean Field Regime*, in "SIGMETRICS Perform. Eval. Rev.", June 2009, vol. 37, n° 1, http://doi.acm.org/10.1145/2492101.1555363

[61] K. Coulomb, M. Faverge, J. Jazeix, O. Lagrasse, J. Marcoueille, P. Noisette, A. Redondy, C. Vuchener. *Visual trace explorer (ViTE)*, October, 2009

[62] J. Doncel, N. Gast, B. Gaujal. *Mean-Field Games with Explicit Interactions*, February 2016, https://hal.inria.fr/hal-01277098

[63] S. Durand, B. Gaujal, F. Perronnin, J.-M. Vincent. *A perfect sampling algorithm of random walks with forbidden arcs*, in "QEST 2014 - 11th International Conference on Quantitative Evaluation of Systems",

Florence, Italy, Springer, September 2014, vol. 8657, pp. 178-193 [*DOI :* 10.1007/978-3-319-10696-0_15], https://hal.inria.fr/hal-01069975

[64] C. FRICKER, N. GAST. *Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity*, in "EURO Journal on Transportation and Logistics", June 2014, 31 p. [*DOI :* 10.1007/S13676-014-0053-5], https://hal.inria.fr/hal-01086009

[65] C. FRICKER, N. GAST, H. MOHAMED. *Mean field analysis for inhomogeneous bike sharing systems*, in "AofA", Montreal, Canada, July 2012, https://hal.inria.fr/hal-01086055

[66] D. FUDENBERG, D. K. LEVINE. *The Theory of Learning in Games*, Economic learning and social evolution, MIT Press, Cambridge, MA, 1998, vol. 2

[67] R. M. FUJIMOTO. *Parallel Discrete Event Simulation*, in "Commun. ACM", October 1990, vol. 33, n^o 10, http://doi.acm.org/10.1145/84537.84545

[68] N. GAST, B. GAUJAL. *Markov chains with discontinuous drifts have differential inclusion limits*, in "Performance Evaluation", 2012, vol. 69, n^o 12, pp. 623-642 [*DOI :* 10.1016/J.PEVA.2012.07.003], https://hal.inria.fr/hal-00787999

[69] N. GAST, B. GAUJAL, J.-Y. LE BOUDEC. *Mean field for Markov Decision Processes: from Discrete to Continuous Optimization*, in "IEEE Transactions on Automatic Control", 2012, vol. 57, n^o 9, pp. 2266 - 2280 [*DOI :* 10.1109/TAC.2012.2186176], https://hal.inria.fr/hal-00787996

[70] N. GAST, J.-Y. LE BOUDEC, D.-C. TOMOZEI. *Impact of Demand-Response on the Efficiency and Prices in Real-Time Electricity Markets*, in "ACM e-Energy 2014", Cambridge, United Kingdom, June 2014 [*DOI :* 10.1145/2602044.2602052], https://hal.inria.fr/hal-01086036

[71] N. GAST, B. VAN HOUDT. *Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms*, in "ACM SIGMETRICS 2015", Portland, United States, June 2015 [*DOI :* 10.1145/2745844.2745850], https://hal.inria.fr/hal-01143838

[72] D. A. GOMES, J. MOHR, R. R. SOUZA. *Discrete time, finite state space mean field games*, in "Journal de Mathématiques Pures et Appliquées", 2010, vol. 93, n^o 3, pp. 308–328

[73] J. GONZALEZ, J. GIMENEZ, J. LABARTA. *Automatic detection of parallel applications computation phases*, in "Parallel and Distributed Processing Symposium, International", 2009, vol. 0, http://doi.ieeecomputersociety.org/10.1109/IPDPS.2009.5161027

[74] M. HEATH, J. ETHERIDGE. *Visualizing the performance of parallel programs*, in "IEEE software", 1991, vol. 8, n^o 5

[75] T. HOEFLER, T. SCHNEIDER, A. LUMSDAINE. *LogGOPSim - Simulating Large-Scale Applications in the LogGOPS Model*, in "Proc. of the ACM Workshop on Large-Scale System and Application Performance", June 2010

[76] L. HU, J.-Y. LE BOUDEC, M. VOJNOVIĆ. *Optimal channel choice for collaborative ad-hoc dissemination*, in "INFOCOM, 2010 Proceedings IEEE", IEEE, 2010

[77] L. V. KALÉ, G. ZHENG, C. W. LEE, S. KUMAR. *Scaling applications to massively parallel machines using Projections performance analysis tool*, in "Future Generation Comp. Syst.", 2006, vol. 22, n° 3

[78] T. G. KURTZ. *Approximation of population processes*, SIAM, 1981, vol. 36

[79] J.-M. LASRY, P.-L. LIONS. *Mean field games*, in "Japanese Journal of Mathematics", 2007, vol. 2, n° 1

[80] Y.-B. LIN, E. D. LAZOWSKA. *A Time-division Algorithm for Parallel Simulation*, in "ACM Trans. Model. Comput. Simul.", January 1991, vol. 1, n° 1, http://doi.acm.org/10.1145/102810.214307

[81] G. LLORT, J. GONZÁLEZ, H. SERVAT, J. GIMÉNEZ, J. LABARTA. *On-line Detection of Large-scale Parallel Application's Structure*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS'2010)", 2010

[82] L. MELLO SCHNORR, A. LEGRAND. *Visualizing More Performance Data Than What Fits on Your Screen*, in "Tools for High Performance Computing 2012", A. CHEPTSOV, S. BRINKMANN, J. GRACIA, M. M. RESCH, W. E. NAGEL (editors), Springer Berlin Heidelberg, 2013, pp. 149-162 [*DOI :* 10.1007/978-3-642-37349-7_10], https://hal.inria.fr/hal-00842761

[83] S. MEYN, P. BAROOAH, A. BUSIC, J. EHREN. *Ancillary service to the grid from deferrable loads: the case for intelligent pool pumps in Florida*, in "Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on", IEEE, 2013

[84] M. MITZENMACHER. *The power of two choices in randomized load balancing*, in "Parallel and Distributed Systems, IEEE Transactions on", 2001, vol. 12, n° 10

[85] K. MOHROR, K. KARAVANIC, A. SNAVELY. *Scalable Event Trace Visualization*, in "Euro-Par 2009 – Parallel Processing Workshops", H.-X. LIN, M. ALEXANDER, M. FORSELL, A. KNÜPFER, R. PRODAN, L. SOUSA, A. STREIT (editors), Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2010, vol. 6043, http://dx.doi.org/10.1007/978-3-642-14122-5_27

[86] W. NAGEL, A. ARNOLD, M. WEBER, H. HOPPE, K. SOLCHENBACH. *VAMPIR: Visualization and Analysis of MPI Resources*, in "Supercomputer", 1996, vol. 12, n° 1

[87] V. PILLET, J. LABARTA, T. CORTES, S. GIRONA. *PARAVER: A tool to visualise and analyze parallel code*, in "Proceedings of Transputer and occam Developments, WOTUG-18", Transputer and Occam Engineering, IOS Press, 1995, vol. 44

[88] J. PROPP, D. WILSON. *Coupling from the past: a user's guide*, in "DIMACS Series on Discrete Mathematics and Theoretical Computer Science", 1998, vol. 41, Microsurveys in discrete probability

[89] M. L. PUTERMAN. *Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons, 2014

[90] D. REED, P. ROTH, R. AYDT, K. SHIELDS, L. TAVERA, R. NOE, B. SCHWARTZ. *Scalable performance analysis: the Pablo performance analysis environment*, in "Scalable Parallel Libraries Conference, 1993., Proceedings of the", 1993

[91] W. H. SANDHOLM. *Population Games and Evolutionary Dynamics*, Economic learning and social evolution, MIT Press, Cambridge, MA, 2010

[92] W. H. SANDHOLM, M. STAUDIGL. *A Sample Path Large Deviation Principle for a Class of Population Processes*, in "arXiv preprint arXiv:1511.07897", 2015

[93] H. SERVAT, G. LLORT, J. GIMÉNEZ, K. HUCK, J. LABARTA. *Folding: detailed analysis with coarse sampling*, in "Tools for High Performance Computing 2011", Springer Berlin Heidelberg, 2012

[94] H. SERVAT, G. LLORT, J. GONZALEZ, J. GIMENEZ, J. LABARTA. *Identifying code phases using piece-wise linear regressions*, in "Parallel and Distributed Processing Symposium, 2014 IEEE 28th International", IEEE, 2014

[95] B. SHNEIDERMAN. *The eyes have it: A task by data type taxonomy for information visualizations*, in "Visual Languages, 1996. Proceedings., IEEE Symposium on", IEEE, 1996

[96] H. TEMBINE, J.-Y. L. BOUDEC, R. EL-AZOUZI, E. ALTMAN. *Mean field asymptotics of Markov decision evolutionary games and teams*, in "GameNets", 2009, pp. 140–150

[97] M. TIKIR, M. LAURENZANO, L. CARRINGTON, A. SNAVELY. *PSINS: An Open Source Event Tracer and Execution Simulator for MPI Applications*, in "Proc. of the 15th International Euro-Par Conference on Parallel Processing", LNCS, Springer, August 2009, n^{o} 5704

[98] B. VAN HOUDT. *A Mean Field Model for a Class of Garbage Collection Algorithms in Flash-based Solid State Drives*, in "Proceedings of the ACM SIGMETRICS", New York, NY, USA, SIGMETRICS '13, ACM, 2013, http://doi.acm.org/10.1145/2465529.2465543

[99] P. VELHO, L. SCHNORR, H. CASANOVA, A. LEGRAND. *On the Validity of Flow-level TCP Network Models for Grid and Cloud Simulations*, in "ACM Transactions on Modeling and Computer Simulation", October 2013, vol. 23, n^{o} 4, https://hal.inria.fr/hal-00872476

[100] J. J. WILKE, K. SARGSYAN, J. P. KENNY, B. DEBUSSCHERE, H. N. NAJM, G. HENDRY. *Validation and Uncertainty Assessment of Extreme-Scale HPC Simulation through Bayesian Inference*, in "Euro-Par 2013 Parallel Processing: 19th International Conference, Aachen, Germany, August 26-30, 2013. Proceedings", Springer Berlin Heidelberg, Berlin, Heidelberg, 2013

[101] F. WOLF, B. MOHR. *Automatic performance analysis of hybrid MPI/OpenMP applications*, in "Journal of Systems Architecture", 2003, vol. 49, n^{o} 10-11

[102] T. YANG, P. G. MEHTA, S. P. MEYN. *A mean-field control-oriented approach to particle filtering*, in "American Control Conference (ACC), 2011", IEEE, 2011

[103] L. YING. *On the Rate of Convergence of Mean-Field Models: Stein's Method Meets the Perturbation Theory*, in "arXiv preprint arXiv:1510.00761", 2015

[104] O. ZAKI, E. LUSK, W. GROPP, D. SWIDER. *Toward Scalable Performance Visualization with Jumpshot*, in "International Journal of High Performance Computing Applications", 1999, vol. 13, n^{o} 3, http://dx.doi.org/10.1177/109434209901300310

[105] G. ZHENG, G. KAKULAPATI, L. KALÉ. *BigSim: A Parallel Simulator for Performance Prediction of Extremely Large Parallel Machines*, in "Proc. of the 18th International Parallel and Distributed Processing Symposium (IPDPS)", April 2004

[106] J. C. DE KERGOMMEAUX, B. STEIN, P. BERNARD. *Paje, an interactive visualization tool for tuning multi-threaded parallel applications*, in "Parallel Computing", 2000, vol. 10, n^{o} 26, pp. 1253–1274