Curriculum Vitæ

Arnaud Legrand

Laboratoire d'Informatique de Grenoble
UMR 5217 (CNRS, INRIA, INPG, UJF, UPMF)

December, 2021

Personal Information

Last, First, and Middle Names: Legrand, Arnaud, Régis

Date of Birth: 02-20-1977

Place of Birth: Créteil (Val de Marne 94), France

Nationality: French

Marital Status: Married, five children

Business Address: LIG - Bâtiment IMAG
791 rue des Résidences, Domaine Universitaire
38401 St Martin d'Hères, France

Business Phone Number: (+33) 4 57 42 15 22

E-Mail Address: Arnaud.Legrand@imag.fr

Web Page: https://team.inria.fr/polaris/members/arnaud-legrand/

Contents

Education and Professional Experience

Oct. 2018 : Tenured Senior Researcher for the CNRS (*Directeur de Recherche*) at Laboratoire d'Informatique de Grenoble

Jan. 2016 : Leader of the POLARIS research team

Nov. 2015 : Habilitation à diriger des recherches, Université Grenoble Alpes
Thesis: *Scheduling for Large Scale Distributed Computing Systems: Approaches and Performance Evaluation Issues.*

Oct. 2004-. . . : Tenured Researcher for the CNRS (*Chargé de Recherche*) at Laboratoire d'Informatique de Grenoble (formerly known as "Informatique et Distribution" laboratory).

2003-2004: Post-Doctoral Research Associate, École Normale Supérieure de Lyon (France).

2000-Dec. 2003: Ph.D. Computer Science, École Normale Supérieure de Lyon. Laboratoire de l'Informatique du Parallélisme.
Thesis: *Heterogeneous parallel algorithms and scheduling : static and dynamic approaches*
Advisors: Prof. Olivier BEAUMONT and Prof. Yves ROBERT.

1999-2000; Diplôme d'études approfondies (DEA/M.Sc.) in Fondamental Computer Science (*Operating Systems, Networks and Parallel Algorithms* section), École Normale Supérieure de Lyon, France.
Thesis: *Parallel linear algebra kernels: heterogeneous and non-dedicated environments*.
Advisors: Prof. Olivier BEAUMONT and Prof. Yves ROBERT.

June-July 1999: Two month internship at University of California, San Diego (USA).
Thesis: *Scheduling Heuristics for Parameter-Sweep Applications on a Grid computing Platform*.
Advisors: Prof. Fran BERMAN and Prof. Henri CASANOVA.

1997-2000 Magistère d'Informatique et de Modélisation (B.Sc. in Computer Science and Modeling); École Normale Supérieure de Lyon, France.

Research Activities

My research takes place in the context of large scale distributed computing infrastructures such as supercomputers/clusters, computing/data grids, volunteer computing platforms, clouds,. . . when used for scientific computing. Such platforms are heterogeneous, dynamic, rely on a complex network infrastructures, are strongly constrained by technology evolution, and are shared by several users. Such complexity raises many technical challenges but also hides very fundamental problems in terms of both exploitation and performance evaluation.
My research activity in the last five years has been organized around the following main three axis:

1. Going beyond classical simulations with SimGrid
2. Multi-scale analysis and visualization of traces of very large computing systems
3. Improving experimental research methodology in computer science : Open Science & Reproducible Research

I will not detail the motivation and the articulation between these different research directions but rather organize my latest contributions along them and underline with which PhD students these contributions have been done.

Last, a fourth "historical" line of research is related to the optimization (load-balancing, scheduling, etc.) in such infrastructures has been revived by the recent developments in online, reinforcement and statistical learning. I have a few publications in this direction, which I will briefly describe in the last paragraph.

Going beyond classical simulations with SimGrid

In practice the scalability of algorithms is often evaluated through very basic models that leave out most of the complexity of real implementations. Most HPC simulators rely on very unrealistic assumptions and models that prevent any hope to faithfully predict the performance of the complex dynamic applications that are now commonly encountered in large HPC infrastructures. SimGrid is a generic toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. It has been been used at CERN for capacity planning and optimization of file replication strategies, in the biomed organization of the EGI grid to optimize task scheduling and checkpointing, in Japan (NII) to design HPC optical networks, by Brazilian researchers to improve map-reduce scheduling strategies, etc. Most of my work related to SimGrid over the last five years has been to study how to predict the performance of HPC applications, in particular of MPI applications, MPI being the communication standard for HPC applications.

- SMPI builds upon SimGrid to allow seamless simulation of arbitrary MPI applications, which allows to capture the full complexity of real HPC applications. There are several ways to achieve such emulation and we have presented an in-depth presentation of SMPI and of the trade-offs between the different approaches in [5].
- Our constant efforts have allowed SMPI to reach a maturity and a stability allowing to use it for teaching purposes [36]. In this context, only simple applications (written by students) are simulated but the simulation capability allows the students to better understand at very low experimental cost the influence of bandwidth, latency or of various algorithmic optimizations.
- We have shown that SMPI could be used to faithfully predict the performance of complex applications such as HPL [32] or Ondes3D [39, 4], a geophysics application developed at the *Bureau de Recherche Géologique et Minière*. Both applications are particularly challenging: HPL implements most parallel algorithm tricks (pipelining, ad hoc collective communication algorithms, look-ahead) found in parallel applications while Ondes3D suffers from non-trivial spatial and temporal load-imbalance that can only be mitigated through smart load-balancing strategies. In every setup, we managed to predict the performance within a few percents, which allows us to conduct most performance related studies fully through simulation. These two studies have been done with two of my PhD students, Rafael Keller Tesser and Tom Cornebize.
- While improving the application support side and studying how we could improve the quality of computation and communication time prediction, we have also added support in SimGrid for other kind of aspects, namely the energy consumption [38] and the performance of I/O operations [42]. Regarding energy consumption, we have proved that a careful calibration of platforms allowed us to predict within a few percents the total energy consumption of a few MPI applications. This work on energy modeling and optimization has mostly been done with my PhD student Christian Heinrich.
- Finally, if MPI is the standard library for parallel applications, it does not offer any support to handle GPU and does not allow to efficiently exploit multi-core architectures. This is why many task-based programming libraries have been proposed in the last decade and in particular StarPU, which is developed at LaBRI and is among the most efficient HPC

runtime. Predicting the performance of applications using StarPU is potentially both (1) easy as computations and communications are well identified through the task semantic and (2) very complex as every scheduling decision is taken dynamically, which results in fully non-deterministic execution paths. Thanks to a tight coupling between StarPU and SimGrid, we have been able to prove that we could systematically and faithfully predict the performance of dense and sparse [30] linear algebra operations running on hybrid (CPU+GPU) nodes. This work has mostly been done with my PhD student Luka Stanisic.

All the previous results are quite unique as none of the other available MPI simulators pretends to accurately predict execution time. The few who published validation results either consider simple computation-bound applications (, which is thus almost trivial) or report predictions that are $\pm 30\%$ away from actual measurements. Having an accurate estimation of execution time is yet essential to procure HPC infrastructures and to allow users to evaluate how much resources they should ask for.
Even though this research direction is far from being terminated (obtaining accurate predictions still requires an important expertise of both SimGrid and the application to evaluate), all these studies brought me to the conclusion that the main challenge does not reside in the modeling part but more in the measurement and model calibration part. An other challenge when validating models and comparing real executions with simulation is related to the difficulty of identifying where the difference comes from and when two executions can be considered to be sufficiently close to each others. These two challenges naturally lead me to consider the following two research directions.

Multi-scale analysis and visualization of traces of very large computing systems

When studying the validity of simulations, we regularly had to compare application traces obtained through simulation with those obtained through real experiments. Regardless of whether the total duration of the execution was similar or not, identifying the most important differences always proved very difficult. One of the main difficulty comes from the fact that such traces comprise an inordinate amount of information, which is both very structured by the application and stochastic. It quickly becomes necessary to conduct multi-scale analysis and aggregation, which requires to put back the a priori knowledge of the analyst in the visualization Although the structure of MPI applications is relatively easy to exploit, the structure of StarPU applications or of dynamically load-balanced MPI applications (e.g., the aforementioned Ondes3D application) is very challenging to exploit.
With colleagues from Brazil, Bordeaux and my PhD student Vinicius Garcia Pinto, we have proposed visualization workflows that build on modern statistics and visualization tools and that can easily be extended [3]. This work is done in collaboration with application/runtime/tracing developers and not only simplified our simulation validation studies but also allowed us to identify several non-trivial performance issues in StarPU applications.

Improving experimental research methodology in computer science: Open Science & Reproducible Research

This theme comprises two main directions:

1. Promoting better research practices: I have co-authored a book on reproducible research [95], which explains the main origins of lack of reproducibility and how to address some of these issues. This book targets a general audience of scientists and aims at guiding the readers through the different concepts and technologies to consider. Another practice, which can be deleterious is related to the authorship and how to credit contributors in articles and software. I have thus co-authored an article with best practices and recommendations when attributing and referencing research software [2].

In the last five years, I have been particularly involved in *promoting Reproducible Research*, better experimental practices and scientific methodology through numerous tutorials and keynotes in conferences and summer schools. To facilitate the dissemination of ideas, I have also organized several webinars (all the slides and links to the videos are available on github[1]) on reproducible research. I have conducted this promotion of laboratory notebooks and reproducible research along with the development of lightweight provenance tracking techniques suited to the distributed and parallel computing community.

2. Developing better research practices: unlike biology or particle physics, experiments in computer science are often considered as cheap. There is thus a lack of experimental culture in computer science whereas I believe computer scientists should build on standard methodological approaches and make them more robust through computer technology. I have therefore particularly explored the notions of laboratory notebooks and of reproducible computational documents to obtain an effective methodology which allows us to efficiently track on a daily basis experimental conditions, provenance, and computational methods. This work is mainly the result of the PhD thesis of Luka Stanisic and I have applied such methodology with all my other master and PhD students ever since, which allows me to publish along every article a computational document with all the data, meta-data and information to reproduce every single figure. This practice is also heavily advertised in my lectures on *Scientific Methodology* and in the *Reproducible Research MOOC*.

   Once information is tracked and made available in an efficient and rigorous way, it is essential to control experiments and plan which ones are worth being conducted. This requires building on solid Design of Experiments for experimental computer science, which I have started exposing in [40] and [31] in the context of the PhD thesis of Pedro Bruel even if it is also particularly present in the thesis of Tom Cornebize.

Optimization and online learning in large infrastructures

Finally, even though in the last five years I have mostly focused on the performance *evaluation* domain, I have also restarted working on the performance *optimization* part. This regain of interest is motivated by the recent development in my team of *online learning techniques*. The work of my PhD student Bruno Luis de Moura Donassolo on workload consolidation in Fog/IoT infrastructures [33] is a first step in evaluating how online learning techniques (bandits, reinforcement learning, etc.) fare in a distributed multi-agent setup where feedback is not only very noisy but also delayed and quite unreliable. Such strategies are indeed supposed to be inherently robust and comparing them with strategies which rely on combinatorial optimization but require faithful information is thus particularly interesting. An other line of work related to online learning has been conducted with my PhD student Salah Zrigui in a batch scheduling context. Indeed, since several scheduling strategies are known to be more or less efficient depending on the workload, it is tempting to dynamically select which one (or which mixture) to use based on the past performance. In [35], we showed that although major performance improvements could be obtained through "overtraining" if workload was unveiled beforehand, in an online context only very limited gain can be obtained because the workload evolves too quickly to efficiently track the optimal configuration.

Administrative and Collective Responsibilities

- I am the leader of the *POLARIS* team, which comprises 4 Inria researchers, 3 CNRS researchers, 4 UGA assistant professors, and 10-15 PhD students. This incurs monthly meet-

[1] https://github.com/alegrand/RR_webinars/blob/master/README.org

ings with both the LIG and the Inria, monthly meetings with the team, yearly interviews of the permanent researchers, and all the management burden of the team related to budget, funding, offices, relational aspects, etc.

- In the last five years, I have been particularly involved in *promoting Reproducible Research*, better experimental practices and scientific methodology through numerous tutorials and keynotes in conferences and summer schools. To facilitate the dissemination of ideas, I have also organized several webinars (all the slides and links to the videos are available on github[2]) on reproducible research. I have conducted this promotion of laboratory notebooks and reproducible research along with the development of lightweight provenance tracking techniques suited to the distributed and parallel computing community. This culminated with the design of a **MOOC** on *Reproducible Research* with Konrad Hinsen (CNRS/Centre de Biophysique Moléculaire) and Christophe Pouzat (CNRS/ Mathématiques Appliquées à Paris 5) with the support of the Inria Learninglab. This MOOC is hosted on the FUN platform and targets graduate students, PhD students, post-doc, engineers and researchers working in any scientific domain relying on computations. In this MOOC, some modern and reliable tools are presented: GitLab for version control and collaborative working, Computational notebooks (Jupyter, RStudio, and Org-Mode) for efficiently combining the computation, presentation, and analysis of data. More than 3,400 people have registered to the first edition in oct–dec 2018 and about 2,000 people have registered to the second edition april-june 2019. The third edition is opened until March 2021 and has already attracted more than 8,500 people.

  In my application domain, computer science shares many features with experimental sciences like physics or biology and I believe, although we can inspire from other domains, that specific approaches also have to be invented. This is why I have recently been proposed to coordinate the new "Reproducible Experiments" action of the GDR RSD.

- Responsable d'axe "Distributed Systems, Parallel Computing and Networking" of the LIG (2020-...)
- Coordinator of the "Reproducible Research" challenge (1 out of 22) of the 2018-2022 Inria Strategic Plan (2017).
- Leader of the Inria Project Laboratory HAC SPECIS (2016-2020)
- Coordination of the evaluation of the Parallel and Distributed Computing theme (9 research teams) at Inria (2016).
- Chargé de Mission by the CNRS of the "Distributed Systems, Parallel Computing and Networking" theme of the LIG (supervision of 2nd year PhD students, scientific animation) (2012-2015)
- Adjunct responsible of the Parallel, Distributed and Embedded Systems option of the Master Of Science in Informatics at Grenoble (2011-2016)
- Member of the CUMI (commission des moyens informatiques) of the LIG (2006-2013)
- Member of the webmasters working group for the LIG laboratory (2007-2013)
- Responsible of the seminars in the ID/LIG laboratory for the MESCAL/MOAIS/NANOSIM teams (2005 - 2015)

[2] https://github.com/alegrand/RR_webinars/blob/master/README.org

Scientific Evaluation Responsibilities

I participated several times to the evaluation of projects for the ANR or for the CAPES/COFECUB. I used to regularly do numerous reviews (around 25 to 35 reviews a year) for various journals (JPDC, ParCo, TPDS, JoGC, Cluster Computing) and conferences (IPDPS, CCGrid, Grid, Europar, BDA, ISPA, ICCS, HiPC, ...) but over the last years, I tend to decline more and more as I lack time.
I served in the program committee of the following conferences:

- 2020: PRECS, Europar
- 2019: ISC
- 2018: SBAC-PAD (head of the performance evaluation track), SuperComputing, RescueHPC
- 2017: CCgrid, Cluster, PPAM
- 2016: HIPC, EuroPar (head of the performance evaluation track), ICPP, COMPAS
- 2015: HIPC, COMPAS, PPAM
- 2014: ICPP, HIPC, COMPAS
- 2013: ICPP, IPDPS, PPAM, COMPAS
- 2012: IPDPS, Renpar
- 2011: CCgrid
- 2010: PCGrid
- 2009: HPDC, Renpar, PCGrid (Workshop on Desktop Grids and Volunteer Computing Systems)
- 2008: Grid, IPDPS, Renpar, PCGrid
- 2004: Grid

I also organized the following workshops:

- 2014-2017: International Workshop on Reproducibility in Parallel Computing (RepPar) in conjunction with Euro-Par
- 2007: Workshop on Programming Models for Grid Computing, held in conjonction with CCGrid'07 (7th IEEE International Symposium on Cluster Computing and the Grid)

Thesis Committees

I have been member of the jury for the following PhD thesis.

- Yiqin Gao (rapporteur): Doctorat de l'Université de Lyon opéré par l'École Normale Supérieure de Lyon. Oct. 2021. *Scheduling independent tasks under budget and time constraints*.
- Arthur Chevalier (rapporteur): Doctorat de l'Université de Lyon opéré par l'École Normale Supérieure de Lyon. Nov. 2020. *Optimisation du placement des licences logicielles dans le Cloud pour un déploiement économique et efficient*.

- Mohamad El Sayah (Rapporteur): Doctorat de l'Université Bourgogne Franche-Comté. Nov. 2019. *Random Generation for the Performance Evaluation of Scheduling Algorithms*.
- Louis Poirel (Président): Doctorat de l'Université de Bordeaux. Dec. 2018. *Méthodes de décomposition de domaine algébriques pour solveurs hybrides (direct/iteratif)*.
- Nicolas Denoyelle (Président): Doctorat de l'Université de Bordeaux. Nov. 2018. *De la localité logicielle à la localité matérielle sur les architectures à mémoire partagée, hétérogène et non-uniforme*.
- Marcos Tulio Amarís González: Doctorat de l'Université de Sao Paulo. Oct. 2018. *Performance prediction of application executed on GPUs using a simple analytical model and machine learning techniques*.
- Valentin Reis (Président): Doctorat informatique de l'Université de Grenoble. Sep. 2018. *Learning to control large-scale parallel platforms*.
- Marcos Amaris Gonzales: PhD in Computer Science from University of São Paulo. June 2018. *Performance Prediction of Applications Executed on GPUs using a Simple Analytical Model and Machine Learning Techniques*
- Tien-Dat Phan (Rapporteur): Doctorat informatique de l'École Normale Supérieure de Rennes. Nov. 2017. *Energy-efficient Straggler Mitigation for Big Data Applications on the Clouds.*
- Rafife Nheili: Doctorat informatique de l'Université de Perpignan Via Domitia. Dec. 2016. *How to improve the numerical reproducibility of hydrodynamics simulations: analysis and solutions for one open-source HPC software*
- Harald Servat, Doctorat informatique de l'Universitat Politècnica de Catalunya, Apr. 2015. Towards instantaneous performance analysis using coarse-grain sampled and instrumented data.
- Sorina Camarasu Pop, Doctorat informatique de l'INSA Lyon, novembre 2013. *Exploitation d'infrastructures hétérogènes de calcul distribué pour la simulation Monte-Carlo dans le domaine médical.*
- Javier Celaya, Doctorat informatique de l'Université de Zaragoza, Dec. 2013. *STaRS: A Scalable Task Routing Approach to Distributed Scheduling.*
- Matthieu Pérotin, Doctorat informatique de l'Université François Rabelais, Tours, Dec. 2008. *Calcul haute performance sur matériel générique.*
- Maxime Martinasso, Doctorat informatique système et communication, Université Joseph Fourrier, Grenoble (LIG), 2007. *Analyse et modélisation des communications concurrentes dans les réseaux haute-performance.*
- Feryal Moulai, Doctorat informatique système et communication, Laboratoire d'Informatique de Grenoble (LIG), 2007. *Nouvelles approches pour l'ordonnancement d'applications parallèles sous contraintes de déploiement d'environnements sur grappes.*

Hiring Committees (for Professor and Assistant Professor positions)

- 2021: PU ENS Rennes 0043.
- 2012: Mcf ENSEIRB 0136, Bordeaux University.
- 2010: Chaire CNRS/G-SCOP 1163, Grenoble University.

Research Supervision

PhD. Students

- Lucas Leandro Nesi (co-tutelle with UFRGS 2020-…): Online Learning Strategies for Distributing Task-Based Applications on Heterogeneous Platforms (co-advised with Lucas Mello Schnorr, funded by the Brazilian Government).
- Pedro Bruel (co-tutelle with USP 2017-2021): Design of experiments and autotuning of HPC computation kernels (co-advised with Alfredo Goldman and Brice Videau, funded by the Brazilian Government).
- Tom Cornebize (2017-2021): High Performance Computing: Towards Better Performance Predictions and Experiments (funded by the French Ministery for Research).
- Salah Zrigui (MENRT 2018-2021): Understan and improving HPC systems performance through machine learning and statistical analysis (co-advised with Denis Trystram).
- Bruno Luis de Moura Donassolo (CIFRE Orange 2017-2020): Decentralized management of applications in Fog computing environments (co-advised with Panayotis Mertikopoulos and Ilhem Fajari, funded by Orange).
- Christian Heinrich (2015-2019): Modeling of performance and energy consumption of HPC systems (funded by Inria).
- Vinicius Garcia Pinto (co-tutelle with UFRGS 2013-2018): Performance analysis and visualization of dynamic task-based applications (co-advised with Lucas Schnorr and Nicolas Maillard, funded by the Brazilian government).
- Rafael Tesser (co-tutelle with UFRGS 2013-2018): Simulation and performance evaluation of dynamical load balancing of an over-decomposed Geophysics application (co-advised with Lucas Schnorr and Philippe Navaux, funded by the Brazilian government).
- Luka Stanisic (2012-2015): Performance evaluation, modeling and simulation of HPC systems; Experimental methodology and reproducible research (co-advised with Jean-François Méhaut, funded by the French Ministery for Research). Current position: Post-doctoral Researcher in HPC at Max Planck Institute.
- Rémi Bertin (2007-2009): Collaboration Mechanisms in Peer-to-Peer and Collaborative Computing Systems (co-advised with Corinne Touati, funded by ANR DOCCA). Did not defend the PhD. Current position: R&D Engineer at Allegro DVT.
- Pedro Velho (2006-2011): Modeling and Simulation of Large Scale Distributed Platforms (co-advised with Jean-François Méhaut, funded by a Brazilian grant). Current position: R&D Senior Software Engineer at ActiveEon.

PostDoc, Engineers

- Bruno Luis de Moura Donassolo (Eng. 2020-2021): Experiment engines for the SILECS platform.
- Lucas Schnorr (Invited Professor 2016-2017): Tracing, observation and visualization of large scale distributed systems.
- Augustin Degomme (Eng. 2012-2015): Simulation/performance prediction of MPI applications. Current position: Research Engineer at Basel University, Switzerland.

- Sascha Hunold (Post-doc 2011-2012): Design of Experiments, Reproducible Research, Fair Scheduling of Bag-of-Tasks Applications Using Distributed Lagrangian Optimization. Current position: Professor at TU Vienna, Austria.

- Lucas Schnorr (Post-doc 2009-2012): Tracing, observation and visualization of large scale distributed systems. Current position: Professor at UFRGS, Brazil.

- Pierre Navarro (Eng. 2010-2012): Improvement of the SimGrid Framework (scalability, robustness, new features, ...)

- Lionel Eyraud (Post-doc 2007): Automatically Building Sound Network Representations. Current position: Inria researcher.

Software Development

SimGrid

SimGrid is a toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. SimGrid is a 17 years-old free software project whose specific goal is to facilitate research in the area of distributed and parallel application scheduling on distributed computing platforms ranging from simple network of workstations to Computational Grids. It is is highly scalable and can simulate hundreds of thousands up to millions of nodes on a single machine while using realistic network models, whose soundness was assessed through thorough (in)validation studies. SimGrid can also be used as an MPI Simulator to realistically simulate unmodified MPI programs. I am one of the main architects (together with Henri Casanova, Martin Quinson, and Frédéric Suter) of this project. The SimGrid project is open-source and hosted on the INRIA gforge[3] and on github[4].
My most notable scientific contributions in this software are related to the efficient implementation and to the assessment of the validity of fluid network and CPU models. I have used this software for my own research on scheduling in grid/volunteer computing and I am particularly active in making it evolve so that it can be used to predict the performance of complex HPC applications.
SimGrid is recognized in the HPC community as one of the most prominent simulation environments as shown by its large community of users and the number of publications that use it: in the past ten years, SimGrid has been the basis of at least 210 articles. Besides, the main four articles on SimGrid have been cited more than 1,550 times according to Google Scholar. There are currently about 220 members in the `simgrid-user` mailing list.
Among other success stories, SimGrid has been used at CERN for capacity planning and optimization of file replication strategies, in the biomed organization of the EGI grid to optimize task scheduling and checkpointing, by colleagues from Inria Bordeaux to perform daily non regression testing of their software, in Japan (NII) to design HPC optical networks, by Brazilian researchers to improve map-reduce scheduling strategies, etc.

The core of SimGrid represents 90,000 single lines of code (in C and C++) and more than 100,000 single lines of codes for examples and regression tests. Here is a rough evaluation of SimGrid's maturity following CNRS/Inria standards:

- Audience: Wide audience, large user's community (5/5)
- Originality: Original software implementing a fair number of original ideas (4/4)
- Maturity: Major software project, strong software engineering (4/5)

[3] http://simgrid.gforge.inria.fr
[4] https://github.com/simgrid/simgrid

- Evolution and Maintenance: Well-defined and implemented plans for future maintenance and evolution (4/4)
- Software distribution and licensing: External packaging and distribution, as part of a popular open source distribution (5/5)

Although I was the main developper in 2002-2009, I have become now the fourth contributor in terms of lines of code as I have gradually handed over development to the community, focusing on experimental validation, on coordination of scientific developments and on potential technological transfer. Here is a rough evaluation of my own contribution to SimGrid following CNRS/Inria standards:

- Design and Architecture: A main contributor (4/4)
- Coding and Debugging: A regular contributor (3/4)
- Maintenance and Support: An occasional contributor (2/4).
- Team/Project Management: A main contributor (4/4)

Publications

My publication strategy depends on the domain and whether PhD students and postdocs are involved or not. I particularly value high quality (e.g., JPDC, TPDS, CCPE) journal publications as it allows to publish results with deeper content and better interactions with reviewers. A few recent journals have a B ranking as it was the best way to target a particular audience (e.g., through a special issue on reproducible research for Operating Systems Review from ACM SIGOPS) or reviewers with a particular expertise (e.g., for ACM TOMACS). Regarding conferences, I mostly target rank A conferences (e.g., CCgrid, Cluster, Euro-Par) but also a few specialized workshops (e.g., PMBS or VPA at the SuperComputing conference, or LSAP) that allow a fast and efficient dissemination of ideas toward a particular audience.

Since it may be difficult to know the different venue in computer science, I indicate between parenthesis the ERA[5] and QUALIS[6] (for conferences) ranking.
Regarding authorship, the author order depends on when and with who the article was written. When contributions are fairly shared, we favor alphabetical order. When the article involves PhD students or postdoc who need to build a reputation, I generally insist on having their name to appear first and my name rather to appear in the end as a supervisor.

Journal articles

[1] B. Donassolo, A. Legrand, P. Mertikopoulos, and I. Fajjari. Online Reconfiguration of IoT Applications in the Fog: The Information-Coordination Trade-off. *IEEE Transactions on Parallel and Distributed Systems*, 33(5):1156–1172, 2022. (**A***).

[2] P. Alliez, R. Di Cosmo, B. Guedj, A. Girault, M.-S. Hacid, A. Legrand, and N. P. Rougier. Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria. *Computing in Science and Engineering*, pages 1–14, 2019.

[3] V. Garcia Pinto, L. M. Schnorr, L. Stanisic, A. Legrand, S. Thibault, and V. Danjean. A Visual Performance Analysis Framework for Task-based Parallel Applications running on Hybrid Clusters. *Concurrency and Computation: Practice and Experience*, 30(18):1–31, April 2018. (**A**).

[5] http://www.core.edu.au/conference-portal
[6] http://www.conferenceranks.com/

[4] R. Keller Tesser, L. Mello Schnorr, A. Legrand, C. Heinrich, F. Dupros, and P. O. Alexandre Navaux. Performance Modeling of a Geophysics Application to Accelerate the Tuning of Over-decomposition Parameters through Simulation. *Concurrency and Computation: Practice and Experience*, pages 1–21, 2018. (**A**).

[5] A. Degomme, A. Legrand, G. Markomanolis, M. Quinson, M. L. Stillwell, and F. Suter. Simulating MPI applications: the SMPI approach. *IEEE Transactions on Parallel and Distributed Systems*, 28(8):14, August 2017. (**A***).

[6] L. Stanisic, A. Legrand, and V. Danjean. An Effective Git And Org-Mode Based Workflow For Reproducible Research. *Operating Systems Review*, 49:61 – 70, 2015.

[7] L. Stanisic, S. Thibault, A. Legrand, B. Videau, and J.-F. Méhaut. Faithful Performance Prediction of a Dynamic Task-Based Runtime System for Heterogeneous Multi-Core Architectures. *Concurrency and Computation: Practice and Experience*, page 16, May 2015. (**A**).

[8] H. Casanova, A. Giersch, A. Legrand, M. Quinson, and F. Suter. Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms. *Journal of Parallel and Distributed Computing*, 74(10):2899–2917, June 2014. (**A***).

[9] R. Bertin, S. Hunold, A. Legrand, and C. Touati. Fair scheduling of bag-of-tasks applications using distributed Lagrangian optimization. *Journal of Parallel and Distributed Computing*, August 2013. (**A***).

[10] P. Velho, L. Schnorr, H. Casanova, and A. Legrand. On the Validity of Flow-level TCP Network Models for Grid and Cloud Simulations. *ACM Transactions on Modeling and Computer Simulation*, 23(4), October 2013. (**B**).

[11] H. Kameda, E. Altman, C. Touati, and A. Legrand. Nash Equilibrium Based Fairness. *Mathematical Methods of Operations Research*, 76(1), 2012. (**B**).

[12] L. M. Schnorr, A. Legrand, and J.-M. Vincent. Detection and analysis of resource usage anomalies in large distributed systems through multi-scale visualization. *Concurrency and Computation: Practice and Experience*, 24:1792–1816, 2011. (**A**).

[13] O. Beaumont, L. Carter, J. Ferrante, A. Legrand, L. Marchal, and Y. Robert. Centralized Versus Distributed Schedulers for Multiple Bag-of-Tasks Applications. *IEEE Transactions on Parallel and Distributed Systems*, 19:698–709, 2008. (**A***).

[14] A. Legrand, A. Su, and F. Vivien. Minimizing the Stretch When Scheduling Flows of Divisible Requests. *Journal of Scheduling*, 2008. (**B**).

[15] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Pipelining broadcasts on heterogeneous platforms. *IEEE Transactions on Parallel and Distributed Systems*, 2005. (**A***).

[16] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Steady-state scheduling on heterogeneous clusters. *International Journal of Foundations of Computer Science*, 2005. (**B**).

[17] A. Legrand, L. Marchal, and Y. Robert. Optimizing the Steady-State Throughput of Scatter and Reduce Operations on Heterogeneous Platforms. *Journal of Parallel and Distributed Computing*, 65:1497–1514, 2005. (**A***).

[18] C. Banino, O. Beaumont, L. Carter, J. Ferrante, A. Legrand, and Y. Robert. Scheduling Strategies for Master-Slave Tasking on Heterogeneous Processor Platforms. *IEEE Transactions on Parallel and Distributed Systems*, 15:319–330, 2004. (**A***).

[19] A. Legrand, H. Renard, Y. Robert, and F. Vivien. Mapping and Load-Balancing Iterative Computations on Heterogeneous Clusters with Shared Links. *IEEE Transactions on Parallel and Distributed Systems*, 15:546–558, 2004. (**A***).

[20] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Scheduling Strategies for Mixed Data and Task Parallelism on Heterogeneous Clusters. *Parallel Processing Letters*, 13:225–244, 2003. (**B**).

[21] O. Beaumont, A. Legrand, and Y. Robert. Scheduling Divisible Workloads on Heterogeneous Platforms. *Parallel Computing*, 29:1121–1152, 2003. (**A**).

[22] O. Beaumont, A. Legrand, and Y. Robert. The Master-Slave Paradigm with Heterogeneous Processors. *IEEE Transactions on Parallel and Distributed Systems*, 14:897–908, 2003. (**A***).

[23] O. Beaumont, A. Legrand, F. Rastello, and Y. Robert. Dense Linear Algebra Kernels on Heterogeneous Platforms: Redistribution Issues. *Parallel Computing*, 28:155–185, 2002. (**A**).

[24] O. Beaumont, A. Legrand, and Y. Robert. Static Scheduling Strategies for Heterogeneous Systems. *Computing and Informatics*, 21:413–430, 2002.

[25] A. Legrand. Équilibrage de Charge Statique Pour Noyaux Dálgèbre Linéaire Sur Plate-Forme Hétérogène. *Revue des Sciences et Technologies de l'Information - Série TSI : Technique et Science Informatiques*, 21(5):711–734, 2002.

[26] O. Beaumont, A. Legrand, F. Rastello, and Y. Robert. Static LU Decomposition on Heterogeneous Platforms. *International Journal of High Performance Computing Applications*, 15:310–323, 2001. (**B**).

Conference articles

[27] L. L. Nesi, A. Legrand, and L. Mello Schnorr. Exploiting system level heterogeneity to improve the performance of a GeoStatistics multi-phase task-based application. In *ICPP 2021 - 50th International Conference on Parallel Processing*, pages 1–10, Lemont, United States, August 2021. (**A**/**A2**).

[28] L. L. Nesi, L. Mello Schnorr, and A. Legrand. Communication-Aware Load Balancing of the LU Factorization over Heterogeneous Clusters. In *IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, Hong Kong, France, December 2020. (**B**/**B1**).

[29] S. Thibault, L. Stanisic, and A. Legrand. Faithful Performance Prediction of a Dynamic Task-based Runtime System, an Opportunity for Task Graph Scheduling. In *SIAM PP 2020 - SIAM Conference on Parallel Processing for Scientific Computing*, Seattle, United States, February 2020. (/**B4**).

[30] E. Agullo, A. Buttari, A. Guermouche, A. Legrand, I. Masliah, and L. Stanisic. Simulation of a Sparse Direct Solver on Heterogeneous Systems using Starpu and Simgrid. In *CSE 2019 - SIAM Conference on Computational Science and Engineering*, Spokane, United States, February 2019. SIAM. (/**B4**).

[31] P. Bruel, S. Quinito Masnada, B. Videau, A. Legrand, J.-M. Vincent, and A. Goldman. Autotuning under Tight Budget Constraints: A Transparent Design of Experiments Approach. In *CCGrid 2019 - International Symposium in Cluster, Cloud, and Grid Computing*, pages 1–10, Larcana, Cyprus, May 2019. IEEE. (**A**/**A2**).

[32] T. Cornebize, A. Legrand, and F. C. Heinrich. Fast and Faithful Performance Prediction of MPI Applications: the HPL Case Study. In *2019 IEEE International Conference on Cluster Computing (CLUSTER)*, 2019 IEEE International Conference on Cluster Computing (CLUSTER), Albuquerque, United States, September 2019. (**A**/**A2**).

[33] B. Donassolo, I. Fajjari, A. Legrand, and P. Mertikopoulos. Fog Based Framework for IoT Service Provisioning. In *CCNC 2019 - IEEE Consumer Communications Networking Conference*, pages 1–6, Las Vegas, United States, January 2019. IEEE. (**B**/**B1**).

[34] B. Donassolo, I. Fajjari, A. Legrand, and P. Mertikopoulos. Load Aware Provisioning of IoT Services on Fog Computing Platform. In *IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019. IEEE. (/**B1**).

[35] A. Legrand, D. Trystram, and S. Zrigui. Adapting Batch Scheduling to Workload Characteristics: What can we expect From Online Learning? In *IPDPS 2019 - 33rd IEEE International Parallel Distributed Processing Symposium*, pages 686–695, Rio de Janeiro, Brazil, May 2019. IEEE. (**A**/**A1**).

[36] H. Casanova, A. Legrand, M. Quinson, and F. Suter. SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation. In *EduHPC-18 - Workshop on Education for High-Performance Computing*, pages 1–10, Dallas, United States, November 2018.

[37] V. G. Pinto, L. Mello Schnorr, A. Legrand, S. Thibault, L. Stanisic, and V. Danjean. Detecção de Anomalias de Desempenho em Aplicações de Alto Desempenho baseadas em Tarefas em Clusters Híbridos. In *WPerformance 2018 - 17º Workshop em Desempenho de Sistemas Computacionais e de Comunicação*, pages 1–14, Natal, Brazil, July 2018.

[38] F. C. Heinrich, T. Cornebize, A. Degomme, A. Legrand, A. Carpen-Amarie, S. Hunold, A.-C. Orgerie, and M. Quinson. Predicting the Energy Consumption of MPI Applications at Scale Using a Single Node. In *Cluster 2017*, Hawaii, United States, September 2017. IEEE. (**A**/**A2**).

[39] R. Keller Tesser, L. Mello Schnorr, A. Legrand, F. Dupros, and P. O. A. Navaux. Using Simulation to Evaluate and Tune the Performance of Dynamic Load Balancing of an Over-decomposed Geophysics Application. In *Euro-Par 2017: 23rd International European Conference on Parallel and Distributed Computing*, page 15, Santiago de Compostela, Spain, August 2017. (**A**/**A2**).

[40] L. Stanisic, L. Mello Schnorr, A. Degomme, F. Heinrich, A. Legrand, and B. Videau. Characterizing the Performance of Modern Architectures Through Opaque Benchmarks: Pitfalls Learned the Hard Way. In *IPDPS 2017 - 31st IEEE International Parallel Distributed Processing Symposium (RepPar workshop)*, Orlando, United States, May 2017. (**A**/**A1**).

[41] V. Garcia Pinto, L. Stanisic, A. Legrand, L. Mello Schnorr, S. Thibault, and V. Danjean. Analyzing Dynamic Task-Based Applications on Hybrid Platforms: An Agile Scripting Approach. In *VPA 2016 - 3rd Workshop on Visual Performance Analysis*, pages 17–24, Salt Lake City, United States, November 2016. IEEE. Held in conjunction with SC16.

[42] A. Lebre, A. Legrand, F. Suter, and P. Veyre. Adding Storage Simulation Capacities to the SimGrid Toolkit: Concepts, Models, and API. In *CCGrid 2015 - Proceedings of the 15th IEEE/ACM Symposium on Cluster, Cloud and Grid Computing*, pages 251–260, Shenzhen, China, May 2015. IEEE/ACM. (**A**/**A2**).

[43] L. Stanisic, E. Agullo, A. Buttari, A. Guermouche, A. Legrand, F. Lopez, and B. Videau. Fast and Accurate Simulation of Multithreaded Sparse Linear Algebra Solvers. In *The 21st IEEE International Conference on Parallel and Distributed Systems*, Melbourne, Australia, December 2015. (**B**/**B1**).

[44] L. Stanisic and A. Legrand. Effective Reproducible Research with Org-Mode and Git. In *1st International Workshop on Reproducibility in Parallel Computing*, Porto, Portugal, August 2014.

[45] L. Stanisic, S. Thibault, A. Legrand, B. Videau, and J.-F. Méhaut. Modeling and Simulation of a Dynamic Task-Based Runtime System for Heterogeneous Multi-Core Architectures. In *Euro-par - 20th International Conference on Parallel Processing*, Euro-Par 2014, LNCS 8632, pages 50–62, Porto, Portugal, August 2014. Springer International Publishing Switzerland. (**A**/**A2**).

[46] P. Bedaride, A. Degomme, S. Genaud, A. Legrand, G. Markomanolis, M. Quinson, M. L. Stillwell, F. Suter, and 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.

[47] L. Mello Schnorr, A. Legrand, and J.-M. Vincent. Interactive Analysis of Large Distributed Systems with Scalable Topology-based Visualization. In *International Symposium on Performance Analysis of Systems and Software (ISPASS'13)*, Austin, Texas, United States, 2013. IEEE Computer Society Press. (**C**/**B1**).

[48] L. Stanisic, B. Videau, J. Cronsioe, A. Degomme, V. Marangozova-Martin, A. Legrand, and J.-F. Mehaut. Performance Analysis of HPC Applications on Low-Power Embedded Platforms. In *DATE - Design, Automation Test in Europe*, pages 475–480, Grenoble, France, March 2013. (**B**/**A1**).

[49] L. Bobelin, A. Legrand, M. A. G. David, P. Navarro, M. Quinson, F. Suter, and 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*, page 19, Ottawa, Canada, May 2012. (**A**/**A2**).

[50] B. De Moura Donassolo, A. Legrand, and C. Geyer. Non-Cooperative Scheduling Considered Harmful in Collaborative Volunteer Computing Environments. In *CCGrid'11 - 11th IEEE International Symposium on Cluster Computing and the Grid*, pages 144–153, Newport Beach, United States, 2011. IEEE Computer Society Press. (**A**/**A2**).

[51] L. Mello Schnorr, A. Legrand, and J.-M. Vincent. Multi-scale analysis of large distributed computing systems. In *Proceedings of the third international workshop on Large-scale system and application performance*, pages 27–34, San Jose, CA, United States, June 2011. ACM.

[52] B. Donassolo, H. Casanova, A. Legrand, and P. Velho. Fast and scalable simulation of volunteer computing systems using SimGrid. In ACM, editor, *LSAP 2010 : 2nd Workshop on Large-Scale System and Application Performance*, pages 605–612, Chicago, Illinois, United States, June 2010. ACM.

[53] H. Kameda, E. Altman, C. Touati, and A. Legrand. Nash Equilibrium Based Fairness. In *International Conference on Game Theory for Networks (GameNets) - 2009*, Game Theory for Networks, 2009. GameNets '09. International Conference on, pages 533 – 539, Istanbul, Turkey, May 2009. IEEE.

[54] P. Velho and A. Legrand. Accuracy Study and Improvement of Network Simulation in the SimGrid Framework. In *SIMUTools'09, 2nd International Conference on Simulation Tools and Techniques*, Rome, Italy, March 2009.

[55] R. Bertin, A. Legrand, and C. Touati. Toward a Fully Decentralized Algorithm for Multiple Bag-of-tasks Application Scheduling on Grids. In *IEEE/ACM International Conference on Grid Computing (Grid)*, Tsukuba, Japan, 2008. (**A**/**A2**).

[56] H. Casanova, A. Legrand, and M. Quinson. SimGrid: a Generic Framework for Large-Scale Distributed Experiments. In *10th IEEE International Conference on Computer Modeling and Simulation - EUROSIM / UKSIM 2008*, Cambrige, United Kingdom, April 2008. IEEE.

[57] L. Eyraud-Dubois, A. Legrand, M. Quinson, and F. Vivien. A First Step Towards Automatically Building Network Representations. In *13th International Euro-Par Conference - Euro-Par 2007*, volume 4641/2007 of *Lecture Notes in Computer Science*, pages 160–169, Rennes, France, August 2008. Springer Berlin / Heidelberg. (**A**/**A2**).

[58] A. Legrand and C. Touati. How to measure efficiency? In *Proceedings of the 1st International Workshop on Game theory for Communication networks (Game-Comm07)*, 2007.

[59] A. Legrand and C. Touati. Non-Cooperative Scheduling of Multiple Bag-of-Task Applications. In *INFOCOM*, Anchorage, Alaska, May 2007. (**A***/**A1**).

[60] O. Beaumont, L. Carter, J. Ferrante, A. Legrand, L. Marchal, and Y. Robert. Centralized Versus Distributed Schedulers Multiple Bag-of-Task Applications. In *International Parallel and Distributed Processing Symposium IPDPS'2006*, Rhodes Island, Greece, 2006. IEEE Computer Society Press. (**A**/**A1**).

[61] A. Legrand, A. Su, and F. Vivien. Minimizing the Stretch When Scheduling Flows of Biological Requests. In *Symposium on Parallelism in Algorithms and Architectures SPAA2006*. ACM Press, 2006. (**A**/**A2**).

[62] M. Quinson, H. Casanova, A. Legrand, and K. Fujiwara. The SIMGRID Project: Simulation and Deployment of Distributed Applications. In IEEE, editor, *The 15th IEEE International Symposium on High Performance Distributed Computing (HPDC'06)*, Paris, France, June 2006. (**A**/**A1**).

[63] O. Beaumont, V. Boudet, P.-F. Dutot, A. Legrand, and Y. Robert. Fondements théoriques pour la conception d'algorithmes efficaces de gestion de ressources. In *Informatique répartie: architecture, parallélisme et systèmes*. Hermes Sciences, 2005.

[64] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Independent and Divisible Tasks Scheduling on Heterogeneous Star-Schaped Platforms with Limited Memory. In *PDP2005, 13th Euromicro Workshop on Parallel, Distributed and Network-Based Processing*, pages 179–186. IEEE Computer Society Press, 2005. (**C**/**B1**).

[65] A. Legrand, A. Su, and F. Vivien. Off-Line Scheduling of Divisible Requests on an Heterogeneous Collection of Databanks. In *Proceedings of the 14th Heterogeneous Computing Workshop*, Denver, Colorado, USA, 2005. IEEE Computer Society Press. (/**B1**).

[66] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Assessing the Impact and Limits of Steady-State Scheduling for Mixed Task and Data Parallelism on Heterogeneous Platforms. In *HeteroPar2004: International Conference on Heterogeneous Computing, Jointly Published with ISPDC2004: International Symposium on Parallel and Distributed Computing*, pages 296–302. IEEE Computer Society Press, 2004.

[67] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Complexity Results and Heuristics for Pipelined Multicast Operations on Heterogeneous Platforms. In *2004 International Conference on Parallel Processing (ICPP2́004)*, pages 267–274. IEEE Computer Society Press, 2004. (**A**/**A2**).

[68] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Pipelining broadcasts on heterogeneous platforms. In *International Parallel and Distributed Processing Symposium IPDPS'2004*. IEEE Computer Society Press, 2004. (**A**/**A1**).

[69] O. Beaumont, A. Legrand, L. Marchal, and Y. Robert. Steady-State Scheduling on Heterogeneous Clusters: Why and How? In *6th Workshop on Advances in Parallel and Distributed Computational Models APDCM*, page 171a (8 pages). IEEE Computer Society Press, 2004.

[70] E. Caron, P. K. Chouhan, and A. Legrand. Automatic Deployment for Hierarchical Network Enabled Server. In *Heterogeneous Computing Workshop*. IEEE Computer Society Press, 2004. (/**B1**).

[71] A. Legrand, L. Marchal, and Y. Robert. Optimizing the Steady-State Throughput of Scatter and Reduce Operations on Heterogeneous Platforms. In *6th Workshop on Advances in Parallel and Distributed Computational Models APDCM 2004*, page 176a (8 pages). IEEE Computer Society Press, 2004.

[72] A. Legrand and M. Quinson. Automatic deployment of the Network Weather Service using the Effective Network View. In *High-Performance Grid Computing Workshop, associated to IPDPS'04*, Santa Fe, New Mexico, USA, April 2004. (**A**/**A1**).

[73] O. Beaumont, A. Legrand, and Y. Robert. Optimal Algorithms for Scheduling Divisible Workloads on Heterogeneous Systems. In *IPDS'2003, Parallel and Distributed Processing Symposium, 2003*, Nice, France, February 2003. IEEE Computer Society Press. (**A**/**A1**).

[74] O. Beaumont, A. Legrand, and Y. Robert. Scheduling Strategies for Mixed Data and Task Parallelism on Heterogeneous Clusters and Grids. In *PDP'2003, 11th Euromicro Workshop on Parallel, Distributed and Network-Based Processing*, pages 209–216, Gênes, Italy, February 2003. IEEE Computer Society Press. (**C**/**B1**).

[75] H. Casanova, A. Legrand, and L. Marchal. Scheduling Distributed Applications: the SimGrid Simulation Framework. In *Proceedings of the Third IEEE International Symposium on Cluster Computing and the Grid (CCGrid0́3)*. IEEE Computer Society Press, 2003. (**A**/**A2**).

[76] A. Legrand, H. Renard, Y. Robert, and F. Vivien. Load-Balancing Iterative Computations on Heterogeneous Clusters with Shared Communication Links. In *PPAM-2003: Fifth International Conference on Parallel Processing and Applied Mathematics*, LNCS 3019, pages 930–937. Springer Verlag, 2003.

[77] A. Legrand, H. Renard, Y. Robert, and F. Vivien. Mapping and Load-Balancing Iterative Computations on Heterogeneous Clusters. In *Euro-PVM-MPI-2003: Recent Advances in Parallel Virtual Machine and Message Passing Interface*, LNCS 2840, pages 586–594. Springer Verlag, 2003.

[78] C. Banino, O. Beaumont, A. Legrand, and Y. Robert. Scheduling Strategies for Master-Slave Tasking on Heterogeneous Processor Grids. In *PARA0́2: International Conference on Applied Parallel Computing*, LNCS 2367, pages 423–432, Espoo, Finland, June 2002. Springer Verlag.

[79] O. Beaumont, L. Carter, J. Ferrante, A. Legrand, and Y. Robert. Bandwidth-Centric Allocation of Independent Tasks on Heterogeneous Platforms. In *International Parallel and Distributed Processing Symposium IPDPS'2002*. IEEE Computer Society Press, 2002. (**A**/**A1**).

[80] O. Beaumont, A. Legrand, and Y. Robert. A Polynomial-Time Algorithm for Allocating Independent Tasks on Heterogeneous Fork-Graphs. In *ISCIS XVII, Seventeenth International Symposium On Computer and Information Sciences*, pages 115–119. CRC Press, 2002. (**C**/**B2**).

[81] O. Beaumont, A. Legrand, and Y. Robert. Mixed Task and Data Parallelism. In *Parallel Matrix Algorithms and Applications*. Université de Neuchâtel, 2002.

[82] O. Beaumont, A. Legrand, and Y. Robert. Static Scheduling Strategies for Dense Linear Algebra Kernels on Heterogeneous Clusters. In *Parallel Matrix Algorithms and Applications*. Université de Neuchâtel, 2002.

[83] O. Beaumont, A. Legrand, and Y. Robert. Static Scheduling Strategies for Heterogeneous Systems. In *ISCIS XVII, Seventeenth International Symposium On Computer and Information Sciences*, pages 18–22. CRC Press, 2002. (**C**/**B2**).

[84] O. Beaumont, V. Boudet, A. Legrand, F. Rastello, and Y. Robert. Heterogeneous Matrix-Matrix Multiplication, or Partitioning a Square into Rectangles: NP-Completeness and Approximation Algorithms. In *EuroMicro Workshop on Parallel and Distributed Computing (EuroMicro2́001)*, pages 298–305. IEEE Computer Society Press, 2001. (**C**/**B1**).

[85] O. Beaumont, A. Legrand, and Y. Robert. Master-Slave Tasking with Heterogeneous Processors. In *2001 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2́001)*, pages 857–863. CSREA Press, 2001. (**B**/**B2**).

[86] O. Beaumont, A. Legrand, and Y. Robert. The Master-Slave Paradigm with Heterogeneous Processors. In Katz, D. S., Sterling, T., Baker, M., Bergman, L., Paprzycki, M., Buyya, and R., editors, *Cluster2́001*, pages 419–426. IEEE Computer Society Press, 2001. (**A**/**A2**).

[87] O. Beaumont, V. Boudet, A. Legrand, F. Rastello, and Y. Robert. Dense Linear Algebra Kernels on Heterogeneous Platforms. In *Parallel Matrix Algorithms and Applications*. Université de Neuchâtel, 2000.

[88] O. Beaumont, V. Boudet, A. Legrand, F. Rastello, and Y. Robert. Heterogeneity Considered Harmful to Algorithm Designers. In *Cluster'2000*, pages 403–404. IEEE Computer Society Press, 2000. (**A**/**A2**).

[89] H. Casanova, A. Legrand, D. Zagorodnov, and F. Berman. Heuristics for Scheduling Parameter Sweep Applications in Grid Environments. In *Heterogeneous Computing Workshop*, pages 349–363, Cancun, Mexico, 2000. (/**B1**).

Book chapters

[90] L. Mello Schnorr and A. Legrand. Visualizing More Performance Data Than What Fits on Your Screen. In Cheptsov, Alexey, Brinkmann, Steffen, Gracia, José, Resch, M. M., Nagel, and W. E., editors, *Tools for High Performance Computing 2012*, pages 149–162. Springer Berlin Heidelberg, 2013.

[91] L. Eyraud-Dubois and A. Legrand. The impact of platform models on scheduling techniques. In Y. Robert and F. Vivien, editors, *Introduction to Scheduling*. Taylor and Francis, 2009.

[92] O. Beaumont, V. Boudet, A. Legrand, F. Rastello, and Y. Robert. Static Data Allocation and Load Balancing Techniques for Heterogeneous Systems. In C. K. Yuen, editor, *Annual Review of Scalable Computing*, volume 4, pages 1–37. World Scientific Publishing, 2002.

PhD thesis and Accreditation to Supervise Research (H.D.R.)

[93] A. Legrand. *Scheduling for Large Scale Distributed Computing Systems: Approaches and Performance Evaluation Issues*. Habilitation à diriger des recherches, Université Grenoble Alpes, November 2015.

[94] A. Legrand. *Algorithmique Parallèle Hétérogène Et Techniques D'ordonnancement : Approches Statiques Et Dynamiques*. PhD thesis, École Normale Supérieure de Lyon, December 2003.

Books

[95] L. Desquilbet, S. Granger, B. Hejblum, A. Legrand, P. Pernot, N. P. Rougier, E. de Castro Guerra, M. Courbin-Coulaud, L. Duvaux, P. Gravier, G. Le Campion, S. Roux, and F. Santos. *Vers une recherche reproductible*. Unité régionale de formation à l'information scientifique et technique de Bordeaux, May 2019.

[96] H. Casanova, A. Legrand, and Y. Robert. *Parallel Algorithms*. Chapman Hall, 2008.

[97] A. Legrand and Y. Robert. *Algorithmique Parallèle – Cours Et Exercices Corrigés*. Dunod, 2003.

Pre-prints and other documents

[98] T. Cornebize and A. Legrand. Simulation-based Optimization and Sensibility Analysis of MPI Applications: Variability Matters. working paper or preprint, February 2021.

[99] M. C. Miletto, L. L. Nesi, L. Mello Schnorr, and A. Legrand. Performance Analysis of Irregular Task-Based Applications on Hybrid Platforms: Structure Matters. working paper or preprint, July 2021.

[100] S. Zrigui, R. Y. De Camargo, A. Legrand, and D. Trystram. Improving the Performance of Batch Schedulers Using Online Job Runtime Classification. working paper or preprint, October 2020.

[101] T. Cornebize, F. C. Heinrich, A. Legrand, and J. Vienne. Emulating High Performance Linpack on a Commodity Server at the Scale of a Supercomputer. working paper or preprint, December 2017.