



**Offer #2023-06335**

## **PhD Position F/M Contention-Aware Scheduling of Storage Resources on Exascale Systems**

**Contract type** : Fixed-term contract

**Level of qualifications required** : Graduate degree or equivalent

**Other valued qualifications** : Master's degree

**Fonction** : PhD Position

### **About the research centre or Inria department**

The Inria Rennes - Bretagne Atlantique Centre is one of Inria's eight centres and has more than thirty research teams. The Inria Center is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative SMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

### **Context**

#### **Context**

This thesis is placed in the context of the PEPR NumPEX (<https://numpex.fr/>), whose goal is to co-design the exascale software stack and prepare applications for the exascale era. This thesis will be co-supervised by [Inria](#) and [CEA](#), respectively the Inria center at the University of Rennes and the CEA center at Bruyères-Le-Châtel, near Paris. Beyond the supervision, collaborations within the PEPR with the different laboratories of the consortium are to be expected.

#### **PhD Advisors**

- François Tessier (Inria KerData team)
- Gabriel Antoniu (Inria KerData team)
- Philippe Deniel (CEA)
- Thomas Leibovici (CEA)

#### **Location and Mobility**

The thesis, which will be co-supervised by Inria and CEA, will be hosted by the KerData team at Inria Rennes Bretagne Atlantique and will include regular visits at the CEA Center of Bruyères-le-Châtel. It may also include collaborations with European or/and international partners such as University of Madrid (Spain), University of Bristol (UK) or Argonne National Lab (USA) to name a few. Rennes is the capital city of Brittany, in the western part of France. It is easy to reach thanks to the high-speed train line to Paris. Rennes is a dynamic, lively city and a major center for higher education and research: 25% of its population are students.

#### **The KerData team in a nutshell for candidates**

- As a PhD student hosted in the KerData team, you will join a dynamic and enthusiastic group, committed to top-level research in the areas of High-Performance Computing and Big Data Analytics. Check the team's web site: <https://team.inria.fr/kerdata/>.

- The team is leading multiple projects in top-level national and international collaborative environments, e.g., the JLESC international Laboratory on Extreme-Scale Computing: <https://jlesc.github.io>. It has active collaborations with high-profile academic institutions all around the world (including the USA, Spain, Germany, Japan, Romania, etc.). The team has close connections with the industry (e.g., ATOS, DDN, Cray-HPE).

- The KerData team's publication policy targets the best-level international journals and conferences of its scientific area. The team also strongly favors experimental research, validated by implementation and experimentation of software prototypes with real-world applications on real-world platforms, e.g., clouds such as Microsoft Azure and some of the most powerful supercomputers in the world.

#### **Why joining the KerData team is an opportunity for you**

- The team's collaborations strongly favor successful PhD theses dedicated to solving challenging problems at the edge of knowledge, in close interaction with top-level experts from both academia and industry.

- To follow the career of our former PhD students, have a look here: <https://team.inria.fr/kerdata/team-members/>.

- The KerData team is committed to personalized advising and coaching, to help PhD candidates train and grow in all directions that are critical in the process of becoming successful researchers.

- You will have the opportunity to present your work in high-ranking venues where you will meet the best experts in the field.

- What you will learn. Beyond learning how to perform meaningful and impactful research, you will acquire useful skills for communication both in written form (how to write a good paper, how to design a convincing poster) and in oral form (how to present their work in a clear, well-structured and convincing way). This is how some of our PhD students received awards in recognition to the quality of their research. Have a look here: <https://team.inria.fr/kerdata/awards/>.

- Additional complementary training will be available, with the goal of preparing the PhD candidates for their postdoctoral career, should it be envisioned in academia, industry or in an entrepreneurial context, to create a startup company.

## Assignment

### Introduction

Nowadays, there are many scientific fields where the need for computing power and data processing capacity goes beyond what current machines can provide. In radio astronomy, for example, the international [SKA](#) project aims to create the largest telescope in the world in order to observe a part of the Universe. A very large volume of data is generated at the telescope level and then transits to geo-distributed data centers to be pre-processed (filtering, reduction) in real time at a rate of 10TB/s. The output data is then sent to a supercomputer to be saved and fed into numerical simulations. At this stage, the computing power and storage resources required are such that machines capable of reaching the exascale become necessary. To date, only a few supercomputers such as Frontier at Oak Ridge National Laboratory (USA) have this capability, but in the coming months, new systems will be deployed. However, the efficient use of these systems raises new challenges, especially regarding data management.

Indeed, even though HPC systems are increasingly powerful, there has been a relative decline in I/O bandwidth. Over the past ten years, the ratio of I/O bandwidth to computing power of the top three supercomputers has been divided by 10 while in some scientific computing centers the volume of data stored has been multiplied by 41 [1]. An aspect that accentuates this gap comes from the design of the machines themselves: while it is common for HPC systems to provide exclusive and dynamic access to compute nodes through a batch scheduler, storage resources are usually global and shared by concurrent applications leading to congestion and performance variability [2,3]. To mitigate this congestion, new tiers of memory and storage have been added to recently deployed supercomputers, increasing their complexity. These new tiers can take the form of node-local SSDs, burst buffers or dedicated storage nodes with network-attached storage technologies, to name a few. Harnessing this additional storage capacity is an active research topic but little has been done about how to efficiently provisioning it [4,5].

### Thesis proposal

Dealing with this high degree of storage heterogeneity a real challenge for scientific workflows and applications. This PhD thesis aims to address this issue through the point of view of the resource provisioning.

## Main activities

Through intelligent scheduling algorithms, the thesis goal is to enable applications and workflows to seamlessly use storage systems [8] on Exascale systems and beyond (Cloud). Multiple criteria can be taken into account further the only resource contention aspect such as financial cost or energy. These algorithms will need to rely on a resource abstraction model that also need to be devised. The evaluation of these algorithms and the implementation of these models will be done in an existing WRENCH-based [6] simulator, called StorAlloc [5], developed in the team. Tools developed by the CEA, including the Robinhood policy engine [7] and the outcomes from the IO-SEA European Project [9] will also be used. For this work, a strong emphasis will be put on international collaborations (University of Manoa (HI, USA) for instance).

The PhD position is mainly based in Rennes, at IRISA/Inria within the KerData research team and regular visits will be organized at the CEA Center near Paris. The selected candidate will have the opportunity to join a very dynamic group in a stimulating work environment with a lot of active national, European and international collaborations as part of cutting-edge international projects in the areas of Exascale Computing, Cloud Computing, Big Data and Artificial Intelligence. The candidate will also have the opportunity to be hosted for 3-6 month internships abroad to strengthen the international visibility of his/her work and benefit from the expertise of other researchers in the field.

### References

[1] GK. Lockwood, D. Hazen, Q. Koziol, RS. Canon, K. Antypas, and J. Balewski. "Storage 2020: A Vision for the Future of HPC Storage". In: Report: LBNL-2001072. Lawrence Berkeley National Laboratory, 2017.

[2] O. Yildiz, M. Dorier, S. Ibrahim, R. Ross, and G. Antoniu. "On the Root Causes of Cross-Application I/O Interference in HPC Storage Systems". In: 2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS). 2016, pp. 750–759

[3] F. Tessier, V. Vishwanath. "Reproducibility and Variability of I/O Performance on BG/Q: Lessons Learned from a Data Aggregation Algorithm". United States: N. p., 2017. Web. doi:10.2172/1414287

[4] F. Tessier, M. Martinasso, M. Chesi, M. Klein, M. Gila. "Dynamic Provisioning of Storage Resources: A Case Study with Burst Buffers". In: IPDPSW 2020 - IEEE International Parallel and Distributed Processing Symposium Workshops, May 2020, New Orleans, United States.

[5] J. Monniot, F. Tessier, M. Robert, G. Antoniu. "StorAlloc: A Simulator for Job Scheduling on Heterogeneous Storage Resources". In: HeteroPar 2022, Aug 2022, Glasgow, United Kingdom.

[6] H. Casanova, R. Ferreira da Silva, R. Tanaka, S. Pandey, G. Jethwani, W. Koch, S. Albrecht, J. Oeth, and F. Suter. "Developing Accurate and Scalable Simulators of Production Workflow Management Systems with WRENCH". In: Future Generation Computer Systems, vol. 112, p. 162-175, 2020.

[7] <https://github.com/cea-hpc/robinhood>

[8] N. Cherièr. "Towards Malleable Distributed Storage Systems: From Models to Practice". Theses. École normale supérieure de Rennes, Nov. 2019.

[9] <https://iosea-project.eu/>

## Skills

- An excellent Master degree in computer science or equivalent
- Strong knowledge of distributed systems
- Knowledge on storage and (distributed) file systems
- Ability and motivation to conduct high-quality research, including publishing the results in relevant venues
- Strong programming skills (Python, C/C++)
- Working experience in the areas of HPC and Big Data management is an advantage
- Very good communication skills in oral and written English.
- Open-mindedness, strong integration skills and team spirit

## Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Possibility of teleworking (90 days per year) and flexible organization of working hours
- Partial payment of insurance costs

## Remuneration

monthly gross salary amounting to 2051 euros for the first and second years and 2158 euros for the third year

## General Information

- **Theme/Domain** : Distributed and High Performance Computing  
Scientific computing (BAP E)
- **Town/city** : Rennes
- **Inria Center** : [Centre Inria de l'Université de Rennes](#)
- **Starting date** : 2023-09-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2023-12-31

## Contacts

- **Inria Team** : [KERDATA](#)
- **PhD Supervisor** :  
Tessier François / [francois.tessier@inria.fr](mailto:francois.tessier@inria.fr)

## About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

## The keys to success

The candidate will have to show motivation, autonomy and an ability to initiate links between the research activities carried out at INRIA and at the CEA center.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

## Instruction to apply

Please submit online : your resume, cover letter and letters of recommendation eventually

For more information, please contact [francois.tessier@inria.fr](mailto:francois.tessier@inria.fr)

### **Defence Security :**

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

### **Recruitment Policy :**

As part of its diversity policy, all Inria positions are accessible to people with disabilities.