



Offre n°2024-07491

PhD Position F/M Enabling Scientific Workflow Composition in Large-Scale Distributed Infrastructures

Type de contrat : Fixed-term contract

Niveau de diplôme exigé : Graduate degree or equivalent

Autre diplôme apprécié : Master's degree

Fonction : PhD Position

A propos du centre ou de la direction fonctionnelle

The Inria Centre at Rennes University is one of Inria's eight centres and has more than thirty research teams. The Inria Centre 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 PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

Contexte et atouts du poste

Supervisory Team

- Silvina Caino-Lores, PhD (Inria, France)
- Gabriel Antoniu, PhD, HDR (Inria, France)

Location and Mobility

The thesis will be hosted by the KerData team at the Inria research center of Rennes. 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.

This thesis will likely include collaborations with international partners from Europe or the USA, thus research visits to and from the collaborator's teams are expected.

The KerData team in a nutshell for candidates

- KerData is a human-sized team currently comprising 5 permanent researchers, 2 contract researchers, 1 engineer and 5 PhD students. You will work in a caring environment, offering a good work-life balance.
- KerData is leading multiple projects in top-level national and international collaborative environments such as within the Joint-Laboratory on Extreme-Scale Computing: <https://jlesc.github.io>. Our team has active collaboration with high-profile academic institutions all around the world (including the USA, Spain, Germany or Japan) and with industry.
- Our team strongly favors experimental research, validated by implementation and experimentation of software prototypes with real-world applications on real-world platforms including some of the most powerful supercomputers worldwide.
- 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.
- Check our website for more about the KerData team here <https://team.inria.fr/kerdata/>

Mission confiée

Context and Overview

As witnessed in industry and science and highlighted in strategic documents such as the European ETP4HPC Strategic Research Agenda [MCS+22], there is a clear trend to combine numerical computations, large-scale data analytics and AI techniques to improve the results and efficiency of traditional HPC applications, and to advance new applications in strategic scientific domains (e.g., high-

energy physics, materials science, biophysics, AI) and industrial sectors (e.g., finance, pharmaceutical, automotive, urbanism). A typical scenario consists in Edge devices creating streams of input data, which are processed by data analytics and machine learning applications in the Cloud; alternatively (or in parallel) they can feed simulations on large, specialised HPC systems, to provide insights and help for prediction of some future system state [BSAPLM17, dSFP+17]. Such emerging applications typically need to be implemented as complex workflows and require the coordinated use of supercomputers, Cloud data centres and Edge-processing devices. This assembly is called the Computing Continuum (CC).

In the current state, multitudes of software development stacks are tailored to specific use cases, with no guarantee of interoperability between them. This greatly impedes application software development for integrated CC use cases. Moreover, existing software stacks have been developed specifically for HPC, data analytics and AI with very different requirements for their initial execution infrastructures, and cannot be integrated efficiently to support CC workflows. Programming the workflow at the highest level requires the ability to consistently combine all these components ad hoc. In this scenario, there is a need to efficiently integrate simulations, data analytics and learning, which first requires interoperable solutions for data processing in the CC [MCS+22]. Existing works on workflow composition and deployment in the CC focus on task-flow control and are disconnected from data patterns and structures beyond domain-specific applications [BTRZ+19, AVHK21]. Moreover, general approaches for representing knowledge and provenance in the form of metadata are also lacking for these converged workflows, and common interfaces for data management in the CC are necessary [RSS+20, GWW+22a]. Unified data abstractions can enable the interoperability of data storage and processing across the continuum and facilitate data analytics at all levels [BGBSC22], alleviating the disconnect between application- and storage-oriented approaches to interoperability. However, no unified data modeling approaches exist for how to structure and represent data on a logical level across the CC.

Research Objectives

This project has the overarching goal of researching new data-centric approaches for scientific workflow composition in the full spectrum of the existing computing continuum, combining large-scale and distributed computing paradigms (e.g., HPC, edge-to-cloud computing) and methods in scientific computing, data science and ML/AI. The project is structured into three primary objectives:

1. **Objective 1: gain a deep understanding of the role of data** in modern workflows and how data influences our ability to effectively and efficiently interoperate computing environments. Breakthroughs in data characterization are needed to understand the next steps towards interoperability in the CC, since existing works tend to focus on task characterization and placement. Workflow data and metadata characterization and profiling will be conducted to deliver data patterns for converged workflows and benchmarks.
2. **Objective 2: model data to enable interoperability** of existing programming models across the CC to leverage the diversity of resources efficiently. Data patterns and workflow characterization as a result of completing the first objective will be the base to research what are the essential attributes needed to represent data and metadata (e.g., ML models, simulation data, annotations resulting from analysis) under uniform data abstractions that can be specialized for different programming models coexisting in the CC. The outcome of this objective is the formal definition of unified data abstractions and their implementation to facilitate the integration of heterogeneous data, tasks and compute resources.
3. **Objective 3: enable modular composition across the continuum.** This supports the vision of the workflows community towards a modular approach to workflow composition and management, in which specialized building blocks (e.g., task scheduling, task control flow, data staging, provenance) can work together and can be configured to support the needs of different computing sites, users and applications. With a focus on data interoperability, we will contribute to this vision by providing a data exchange layer connecting established data staging and transport layers, alleviating the disconnect between raw data management and knowledge-based workflow management in the CC (e.g., in anomaly detection, steering, resource balancing, and provenance). This will be a key software deliverable of the overall project and will allow the composition and deployment of workflows across the full spectrum of the CC.

Principales activités

Envisioned Approach

The key novelty of this project is the perspective that places data in a central role for building and managing scientific workflows in the CC. We will build upon previous work that already leveraged this idea in the scope of high-performance computing and cloud computing, we will enrich and extend its reach, and we will increase its impact by establishing new collaborations.

For **Objective 1**, first we will survey the literature to define how existing workflows are currently leveraging the resources in the CC in terms of infrastructure (e.g., HPC with cloud support, edge-to-cloud, and federated infrastructures) and application structure (e.g., ML-powered workflows, workflows including HPC simulations, use cases with large-scale science instruments, and applications of in situ analysis and visualization). A taxonomical analysis of these workflows will be complemented with exhaustive profiling of data volume, production rate, transfer volume and frequency of communication. We will conduct a descriptive statistical analysis of these metrics to characterize the data needs of these workflows and how they impact performance and scalability. In addition, we will qualitatively classify the characteristic features in the most common data models in the CC. In addition, a key aspect to consider is how different data access and transfer patterns at the application level affect the performance of the underlying computing hardware. For example, GPU accelerators are now a common resource in HPC, and much work has been conducted to improve the arrangement of data in memory to facilitate the work of

the GPU in common application scenarios. Similarly, we will extract common workflow motifs and data patterns that must be represented by unified data abstractions and their interfaces. We have established initial connections with teams from Barcelona Supercomputing Centre (Spain) and Oak Ridge National Laboratory (USA) that are currently investigating these aspects [BBEdSL24, GWW+22b, GGX+22].

Objective 2 will borrow inspiration from our previous work that enabled the interoperability of data models tailored for supercomputing applications with Big Data-oriented paradigms like streaming and key-value storage [CLCN+19]. We achieved the interoperability of process- and data-centric programming models by representing the core characteristics of their respective data abstractions. We proposed a unified distributed data abstraction inspired by the data-awareness and task-based parallelism of data-centric abstractions, but with the possibility to preserve state as required by HPC applications. This abstraction represents a distributed collection of data organized in chunks, which can be locally accessible by both process- and data-centric computing units. We will leverage this foundation and the knowledge derived from the characterization of the data patterns in Objective 1 to consolidate the characteristic features of additional computing models into unified data abstractions suitable for programming models coexisting in the CC. Practically, in Objective 2 we will specialize the data abstractions by defining implementations tuned for different parts of the CC, and provide translation methods to interoperate these concrete implementations. These translation methods will take the form of data transformation interfaces and decorators for the specialization of data abstractions to specific infrastructures, thus hiding the details of the different implementations coexisting in the CC from the user of the data abstraction. The implementation will expose a Python interface, which is ubiquitous to most elements of the CC, and allows for a simple but powerful approach to workflow composition through Jupyter notebooks. This approach is increasingly accepted to build workflows as it can encapsulate configuration, composition, deployment, and post hoc analysis in a usable and reproducible manner [BTK+21, CAC+22]. Furthermore, Python programming is a common skill even for domain scientists, which can increase adoption and impact of the resulting software.

Objective 3 requires, first, identifying key enabling technologies for data staging and data management, including domain-specific and generalist solutions. We will study which technologies would give the best support to the unified data abstractions implemented in Objective 2, prioritizing the solutions produced in our team like Damaris [DAC+16], a middleware for I/O management and real-time processing of data from large-scale MPI-based HPC simulations. We will also study related solutions from our network of collaborators (e.g., Oak Ridge National Laboratory ADIOS-2 [GPW+20], University of Utah DataSpaces [DPK10], Argonne National Laboratory BraidDB [WLV+22]) to secure proper support in the adoption of these technologies. ADIOS-2 is a particularly promising enabling technology as it allows applications to express what data is produced, when that data is ready for output, and what data an application needs to read and when. Ultimately, we will design a data exchange layer that connects with the data staging and transport layers, alleviating the disconnect between raw data management and knowledge-based workflow management in the CC. This will take the form of an API that hides the complexity of re-implementing data models for different target infrastructures and connecting with the underlying data staging, transfer, and storage technologies. We have already discussed aspects of design of this data exchange layer with colleagues from ORNL as part of a symposium at the SIAM PP 2024 conference. Finally, we will build demonstrator workflows and end-to-end real-world applications on top of this system, specifically targeting our ongoing efforts towards supporting workflow composition for large-scale scientific projects like the Square-Kilometer Array Observatory (SKAO) in the context of [the ExaDoST project of the NumPEX PEPR program](#).

References

- [AVHK21] Gabriel Antoniu, Patrick Valduriez, Hans-Christian Hoppe, and Jens Krüger. Towards Integrated Hardware/Software Ecosystems for the Edge-Cloud-HPC Continuum, 2021. ETP4HPC White Paper.
- [BBEdSL24] Rosa M. Badia, Laure Berti-Equille, Rafael Ferreira da Silva, and Ulf Leser. Integrating HPC, AI, and Workflows for Scientific Data Analysis (Dagstuhl Seminar 23352). Dagstuhl Reports, 13(8):129–164, 2024.
- [BGBSC22] Pablo Brox, Javier Garcia-Blas, David E Singh, and Jesus Carretero. Dice: Generic data abstraction for enhancing the convergence of hpc and big data. In High Performance Computing: 8th Latin American Conference, CARLA 2021, Guadalajara, Mexico, October 6–8, 2021, Revised Selected Papers, pages 106–119. Springer, 2022.
- [BSAPLM17] Rosa Maria Badia Sala, Eduard Ayguad ´e Parra, and Jes ´us Jos ´e Labarta Mancho. Workflows for science: A challenge when facing the convergence of hpc and big data. Supercomputing frontiers and innovations, 4(1):27–47, 2017.
- [BTK+21] Marijan Beg, Juliette Taka, Thomas Kluyver, Alexander Konovalov, Min Ragan-Kelley, Nicolas M. Thi ´ery, and Hans Fangohr. Using jupyter for reproducible scientific workflows. Computing in Science Engineering, 23(2):36–46, 2021.
- [BTRZ+19] Daniel Balouek-Thomert, Eduard Gibert Renart, Ali Reza Zamani, Anthony Simonet, and Manish Parashar. Towards a computing continuum: Enabling edge-to-cloud integration for data-driven workflows. The International Journal of High Performance Computing Applications, 33(6):1159–1174, 2019.

[CAC+22] Iacopo Colonnelli, Marco Aldinucci, Barbara Cantalupo, Luca Padovani, Sergio Rabellino, Concetto Spampinato, Roberto Morelli, Rosario Di Carlo, Nicolò Magini, and Carlo Cavazzoni. Distributed workflows with jupyter. *Future Generation Computer Systems*, 128:282–298, 2022.

[CLCN+19] Silvina Caino-Lores, Jesus Carretero, Bogdan Nicolae, Orcun Yildiz, and Tom Peterka. Toward high-performance computing and big data analytics convergence: The case of spark-diy. *IEEE Access*, 7:156929–156955, 2019.

[DAC+16] Matthieu Dorier, Gabriel Antoniu, Franck Cappello, Marc Snir, Robert Sisneros, Orcun Yildiz, Shadi Ibrahim, Tom Peterka, and Leigh Orf. Damaris: Addressing performance variability in data management for post-petascale simulations. *ACM Transactions on Parallel Computing (TOPC)*, 3(3):1–43, 2016.

[DPK10] Ciprian Docan, Manish Parashar, and Scott Klasky. Dataspaces: an interaction and coordination framework for coupled simulation workflows. In *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*, pages 25–36, 2010.

[dSFP+17] Rafael Ferreira da Silva, Rosa Filgueira, Ilia Pietri, Ming Jiang, Rizos Sakellariou, and Ewa Deelman. A characterization of workflow management systems for extreme-scale applications. *Future Generation Computer Systems*, 75:228–238, 2017.

[GGX+22] Ana Gainaru, Dmitry Ganyushin, Bing Xie, Tahsin Kurc, Joel Saltz, Sarp Oral, Norbert Podhorszki, Franz Poeschel, Axel Huebl, and Scott Klasky. Understanding and leveraging the i/o patterns of emerging machine learning analytics. In Jeffrey Nichols, Arthur ‘Barney’ Maccabe, James Nutaro, Swaroop Pophale, Pravallika Devineni, Theresa Ahearn, and Becky Verastegui, editors, *Driving Scientific and Engineering Discoveries Through the Integration of Experiment, Big Data, and Modeling and Simulation*, pages 119–138, Cham, 2022. Springer International Publishing.

[GPW+20] William F Godoy, Norbert Podhorszki, Ruonan Wang, Chuck Atkins, Greg Eisenhauer, Junmin Gu, Philip Davis, Jong Choi, Kai Germaschewski, Kevin Huck, et al. Adios 2: The adaptable input output system. a framework for high-performance data management. *SoftwareX*, 12:100561, 2020.

[GWW+22a] Ana Gainaru, Lipeng Wan, Ruonan Wang, Eric Suchyta, Jieyang Chen, Norbert Podhorszki, James Kress, David Pugmire, and Scott Klasky. Understanding the impact of data staging for coupled scientific workflows. *IEEE Transactions on Parallel and Distributed Systems*, 33(12):4134–4147, 2022.

[GWW+22b] Ana Gainaru, Lipeng Wan, Ruonan Wang, Eric Suchyta, Jieyang Chen, Norbert Podhorszki, James Kress, David Pugmire, and Scott Klasky. Understanding the impact of data staging for coupled scientific workflows. *IEEE Transactions on Parallel and Distributed Systems*, 33(12):4134–4147, 2022.

[MCS+22] Michael Malms, Laurent Cargemel, Estela Suarez, Nico Mittenzwey, Marc Duranton, Sakir Sezer, Craig Prunty, Pascale Rossé-Laurent, Maria Pérez-Hernández, Manolis Marazakis, Guy Lonsdale, Paul Carpenter, Gabriel Antoniu, Sai Narasimhamurthy, André Brinkman, Dirk Pleiter, Utz-Uwe Haus, Jens Krueger, Hans-Christian Hoppe, Erwin Laure, Andreas Wierse, Valeria Bartsch, Kristel Michielsen, Cyril Allouche, Tobias Becker, and Robert Haas. ETP4HPC’s SRA 5 - Strategic Research Agenda for High-Performance Computing in Europe - 2022. Zenodo, 2022.

[RSS+20] Daniel Rosendo, Pedro Silva, Matthieu Simonin, Alexandru Costan, and Gabriel Antoniu. E2clab: Exploring the computing continuum through repeatable, replicable and reproducible edge-to-cloud experiments. In *2020 IEEE International Conference on Cluster Computing (CLUSTER)*, pages 176–186. IEEE, 2020.

[WLV+22] Justin M Wozniak, Zhengchun Liu, Rafael Vescovi, Ryan Chard, Bogdan Nicolae, and Ian Foster. Braid-db: Toward ai-driven science with machine learning provenance. In *Driving Scientific and Engineering Discoveries Through the Integration of Experiment, Big Data, and Modeling and Simulation: 21st Smoky Mountains Computational Sciences and Engineering, SMC 2021, Virtual Event, October 18–20, 2021, Revised Selected Papers*, pages 247–261. Springer, 2022.

Compétences

Required:

- An excellent academic record in computer science courses
- Knowledge on distributed systems and data management systems
- Strong programming skills (Python, C/C++)
- Ability and motivation to conduct high-quality research, including publishing the results in relevant venues
- Very good communication skills in oral and written English
- Open-mindedness, strong integration skills and team spirit

Appreciated:

- Knowledge on scientific computing and data analysis methods
- Professional experience in the areas of HPC and Big Data management

Avantages

- - 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

Rémunération

Monthly gross salary amounting to 2100 euros for the first and second years and 2190 euros for the third year

Informations générales

- **Thème/Domaine** : Distributed and High Performance Computing
Scientific computing (BAP E)
- **Ville** : Rennes
- **Centre Inria** : [Centre Inria de l'Université de Rennes](#)
- **Date de prise de fonction souhaitée** : 2024-09-01
- **Durée de contrat** : 3 years
- **Date limite pour postuler** : 2024-06-04

Contacts

- **Équipe Inria** : [KERDATA](#)
- **Directeur de thèse** :
Caino Lores Silvina / silvina.caino-lores@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

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

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

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

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un

poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.