A propos du centre ou de la direction fonctionnelle

The Inria research centre in Lyon is the 9th Inria research centre, formally created in January 2023. It brings together approximately 300 people in 16 research teams and research support services.

Its staff are distributed at this stage on 2 campuses: in Villeurbanne La Doua (Centre / INSA Lyon / UCBL) on the one hand, and Lyon Gerland (ENS de Lyon) on the other.

The Lyon centre is active in the fields of software, distributed and high-performance computing, embedded systems, quantum computing and privacy in the digital world, but also in digital health and computational biology.

Contexte et atouts du poste

Numerical simulation is a key technology for many application domains. Thanks to the democratization of high performance computers (HPC), complex physics and more general complex systems can now be simulated routinely. Numerical simulation is considered as the third pillar of sciences (with experiment and theory) and is critical to gain competitive position.

One critical aspect to achieve high performance is the management of the partitioned data that is closely related to the parallelization itself as well as to hierarchy memory optimizations in particular with respect to memory layout and accesses. Unfortunately, different configurations are needed to efficiently exploit different hardware. Hence, a portability layer is usually used to abstract the actual management of partitioned data from their usage. Therefore, it enables the implementation of optimization strategies behind this layer.

The situation is getting more complex as the partitioned data is accessed from multiple points of view. First of all, application developer needs to access it to actually use it, ideally from an application abstract view. Second, as an efficient memory layout may differ for example between the CPU and GPU space, task oriented runtime systems such as StarPU [1] need to have access to the implementation detail of the partitioned data to be able to efficiently handle data transfer and task scheduling. Third, programming models that aim at simplifying the development of HPC and increasing portability such as COMET [2] also need a specific access to the internal of partitioned data to be able to generate tasks and their dependencies.

This work is part of the Exa-Soft projet of NumPEx, a PEPR on the Exascale. Exa-Soft aims at consolidating the exascale software ecosystem by providing a coherent, exascale-ready software stack featuring breakthrough research advances enabled by multidisciplinary collaborations between researchers. This position is part of the efficient and composable programming model workpackage but include joint work with the runtime system workpackage.

Mission confiée

The recruited person will participate to the definition of a model that support the multiple views that need to access to a partitionned data. The model shall address scientific challenges related to productivity, performance portability, and heterogeneity support.

First, the recruit person will conduct a state-of-the-art analysis in particular with respect to data partitioning, programming models (such as COMET [1]) and runtimes (such as StarPU) as well as to selected applications. Second, she will propose a model that support data views from application developers, StarPU, and Comet needs. Third, she will implement and evaluate a proof-of-concept to validate the feasibility of the approach. The PoC will be applied to a use case to be selected during the analysis phase.

The recruited person will work in an stimulating environment in the AVALON team located in the LIP at ENS de Lyon. She will be in connection with the STORM team that develops StarPU as well as many partner of NumPEx to define the actual needs from the various perspectives. In particular, she will collaborate with the BEAGLE team that brings the Avelo use case [3].


**Principales activités**
- Study the state-of-the-art of partitioned data management
- Propose a model combining various views on partitioned data
- Implement a PoC that support StarPU and COMET
- Conduct an evaluation of the proof of concept on selected use cases

**Compétences**

Technical skills and level required:
- very good knowledge of HPC programming models and runtimes (e.g. OpenMP)
- good programming skills (Python and C++)
- knowledge in (meta-)modeling will be appreciated

**Avantages**
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage