



Offer #2024-07903

Post-Doctoral Research Visit F/M Event detection for Large-Scale Physical Simulations

Contract type : Fixed-term contract

Level of qualifications required : PhD or equivalent

Fonction : Post-Doctoral Research Visit

About the research centre or Inria department

The Inria Saclay-Île-de-France Research Centre was established in 2008. It has developed as part of the Saclay site in partnership with Paris-Saclay University and with the Institut Polytechnique de Paris since 2021.

The centre has 39 project teams , 27 of which operate jointly with Paris-Saclay University and the Institut Polytechnique de Paris. Its activities occupy over 600 scientists and research and innovation support staff, including 54 different nationalities.

Context

While artificial intelligence is growing at a fast pace, the bulk of the world's computing power remains targeted at modeling and predicting physical phenomena, such as climate models, weather forecasting, or nuclear physics. These simulations are run on highly parallel supercomputers on which both the hardware and the software are optimized for the task at hand. While the computing power of each processing unit is still increasing, the communication networks and the storage capabilities in these clusters do not follow such fast trends. As a result, computing nodes produce outputs faster than what can be stored or sent to process elsewhere: These **simulations are IO-bound**.

To reduce the communication burden, a promising venue is **in situ computations**, meaning that most of the data is processed locally by the nodes, and only meaningful aggregates are stored or sent over the network. However, this is a difficult problem in general since to detect meaningful information for the global simulation, one often needs to detect particular events, requiring non-local information that depends on the other nodes' output. The goal of this postdoc is to **leverage machine learning techniques to bypass IO bottlenecks in the context of physics simulation on high-performance computing (HPC) clusters** and help steer them optimally. This work is thus placed in a broader "Machine Learning for Science" context, which aims at using ML to solve key problems arising in traditional sciences. More specifically, we will focus on distributed event detection techniques, which critically inform where the communication and storage budget needs to be spent to retain most of the information.

Environment. The postdoc will take place in Inria Saclay, in the MIND team. This is a large team working focused on mathematical methods for statistical modeling of brain function using neuroimaging data (fMRI, MEG, EEG). Particular topics of interest include machine learning techniques, numerical and parallel optimization, applications to human cognitive neuroscience, event detection, and scientific software development. A particular emphasis is put on interdisciplinary projects. The postdoc will include frequent interaction with Virginie Grandgirard and Julien Bigot, from the CEA, who are experts in the Gysela code, a fusion plasma simulation code used in the ITER project. The supervisors are dynamic researchers, with a strong track record in machine learning, HPC and physical simulation.

This project also takes place in the PEPR NumPEX, an initiative to improve the use of supercomputers for physical simulations. The results from the postdoc will thus be integrated into the software stack for these applications. This position thus provides the unique opportunity to discuss with scientists from other fields and to improve their workflows through IA research. Interaction with scientists developing computational simulations in various fields will be encouraged.

Assignment

This postdoc intends to provide efficient statistical modeling tools for large distributed HPC outputs. The goal of these tools will be to help identify anomalies or errors in the data, which can indicate problems with the HPC system or the underlying algorithms. This information can then be used to adjust the computations and ensure that the HPC system is running smoothly and efficiently. Moreover, statistical processing can also be used to analyze the outputs and detect events of interest to select the right

snapshots of the simulation to allow post-processing of the results. By properly selecting the information to save, we can ensure that the HPC system is able to handle large volumes of output without running into storage or bandwidth limitations.

As specified above, the project will first focus on event and anomaly detection techniques for large scale simulation. The two main objectives are the following:

- Benchmark existing methods: This will require a thorough state-of-the-art review, as well as defining the relevant metrics for evaluating event detection in physics simulations (precision/recall, distance to the event, distribution of the events...). The benchmark will be realized with benchopt [5] and will benefit from the event detection expertise of the supervisors, and from their expert knowledge in the simulations.
- Designing efficient implementation for existing methods: To account for the structure of physic simulations, we propose to investigate how to efficiently leverage the distributed in-situ framework to efficiently leverage existing event detection methods [3, 4, 8, 1].
- Looking for events with self-supervised learning: Finally, the postdoc will consider developing novel unsupervised event detection algorithms, either based on recurring patterns for common events or based on anomaly detection for rare events. Methods based both on optimization –such as convolutional dictionary learning [2] and event modeling [9]– and on deep learning –such as self-supervised object detection [7]– will be considered.

These event detection methods will be evaluated on the previously defined benchmarks, to test their efficiency in highlighting known events. They will then be integrated into the in-situ framework to test their efficiency in a real-world setting. A strong emphasis will be put on the interpretability of the detected events, to ensure that the detected events are meaningful for the scientists. To fit the requirements imposed by the HPC setting, we will consider distributed methods, that work with streaming data split over many computing nodes. An important consideration in our context is that, unlike classical data stream, the data is not i.i.d. on the nodes, but stems from the domain partitioning imposed by the physic of the problem. These considerations will benefit from the supervisors expertise in distributed optimization, for instance from the DICODILE algorithm used to scale up dictionary learning [6].

Main activities

Main activities :

- Read papers and state of the art
- Benchmark existing algorithms
- Write problem formulation and proofs of convergence.
- Adapt the formulation to the target scenario.
- Propose a new dedicated algorithm.
- Program, run, and analyze simulation results.

Complementary activities

- Participate to the teams activities : scientific meetings, seminars, scientific presentations.

Skills

- Strong mathematical background. Knowledge in machine learning is a plus.
- Good programming skills in Python. Knowledge of a deep learning framework is a plus.
- The candidate should be proficient in English. Knowing French is not necessary, as daily communication in the team is mostly in English due to the strong international environment.

Benefits package

- 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

Remuneration

2788 € gross/month

General Information

- **Theme/Domain :** Distributed and High Performance Computing

- Statistics (Big data) (BAP E)
- **Town/city** : Palaiseau
- **Inria Center** : [Centre Inria de Saclay](#)
- **Starting date** : 2024-09-01
- **Duration of contract** : 2 years
- **Deadline to apply** : 2024-08-31

Contacts

- **Inria Team** : [MIND](#)
- **Recruiter** :
Moreau Thomas / thomas.moreau@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

We seek candidates strongly motivated by challenging research topics in machine learning for science. Applicants should have a strong mathematical background with knowledge of numerical optimization and machine learning. With regards to software engineering, proficiency in Python is expected and experience in applying ML to large scale data is a plus.

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

- CV
- Cover letter
- Letter(s) of recommendation, where applicable

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.