



Offre n°2024-08416

Development & Research Engineer @ Alpes-Grenoble: Deep and Shallow Parallel Data Processing on Supercomputers

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Contrat renouvelable : Oui

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Ingénieur scientifique contractuel

Niveau d'expérience souhaité : Jeune diplômé

A propos du centre ou de la direction fonctionnelle

The Centre Inria de l'Université de Grenoble groups together almost 600 people in 22 research teams and 7 research support departments.

Staff is present on three campuses in Grenoble, in close collaboration with other research and higher education institutions (Université Grenoble Alpes, CNRS, CEA, INRAE, ...), but also with key economic players in the area.

The Centre Inria de l'Université Grenoble Alpes is active in the fields of high-performance computing, verification and embedded systems, modeling of the environment at multiple levels, and data science and artificial intelligence. The center is a top-level scientific institute with an extensive network of international collaborations in Europe and the rest of the world.

Contexte et atouts du poste

The candidate will join the [DataMove](#) INRIA team located on the campus of the Univ. Grenoble Alpes near Grenoble. The DataMove team is a friendly and stimulating group with a strong international visibility, gathering Professors, Researchers, PhD and Master students all pursuing research on High Performance Computing.

This work experience will bring you skills from high performance computing up to deep learning that are in high demand.

This work is part of a joint collaboration with international academic partners, giving you the opportunity to work in an international context.

Hiring date is flexible, starting as early as February 2025. Initial contract will last up to the end of 2026, with possibilities for extension.

The city of Grenoble is surrounded by the Alps mountains, offering a high quality of life, and where you can experience all kinds of mountain related outdoors activities and more.

Principales activités

Dask (<https://www.dask.org/>) and Ray (<https://www.ray.io/>) are open source frameworks to distribute the execution of Python tasks as actors on supercomputer and cloud. They provide seamless parallelization of classical data processing libraries like Numpy (<https://numpy.org/>), Panda (<https://pandas.pydata.org/>) or Scikit-learn (<https://scikit-learn.org>). Ray and Dask also enable to deploy the classical AI stacks like Pytorch (<https://pytorch.org/>) or Jax (<https://jax.readthedocs.io/>), through actors on multiple GPUs for training or inference. This makes these frameworks very popular for advanced high performance data processing in the scientific and machine learning communities.

Classical numerical solvers for scientific computing have been central to the development of supercomputers as they can require up to millions of cores for simulating high resolution systems or phenomena.

Today there is a strong need to mix both large scale solvers and data processing tools and run them in a coupled mode on supercomputers.

We developed the open source library `Deisa` (code: <https://github.com/GueroudjiAmal/deisa> - PhD: <https://theses.hal.science/tel-04194958>) to extend Dask with classical parallel solvers based on MPI. The data produced by each MPI process of the solver are routed as soon as available to Dask workers that can execute tasks to process these data. This data are exposed to the user as Dask Arrays a distributed extension of Numpy arrays, that can then conveniently rely on the classical Python Numpy API to process the data in parallel (operations on Dask Arrays are split into tasks distributed automatically to the workers).

We are looking for an engineer that will join our team to extend this work into a consolidated framework and participate to the development of advanced analysis scenarios:

- Performance improvement. We target to deploy `Deisa` with large applications on the european Exascale supercomputers.
- Support novel features by integrating AI frameworks like JAX (<https://github.com/google/jax>) and Pytorch (<https://pytorch.org/>) for instance.
- Refine the programming environment by developing new APIs or algorithms for easing code coupling.
- Develop prototype data processing pipelines for two specific applications: Gyselax (plasma simulation code - <https://gyselax.github.io/>), and Parflow (water flow simulation - <https://parflow.org/>). Required data processing ranges from classical linear algebra to shallow machine learning or deep neural networks.
- Run experiments on a variety of supercomputers
- Participate to the research activity, possibly leading to publications.
- Collaborate with other European partners as this work is part of the European project Eocoe-III (<https://www.eocoe.eu/>).

Through this work the candidate will gain strong expertise in high performance computing and high performance data analysis. She/he will integrate a dynamics research team and have the opportunity to work in an international context.

Compétences

We welcome candidates with a master (or equivalent title) in computer science, experience with parallel programming, distributed data processing, deep learning or numerical solvers.

Expected technical skills include Linux, Python and some C/C++ programming practice, mastering of development processes is a plus (git, continuous integration, containers, etc.).

No previous work experience required as long as your are motivated and ready to train yourself to complement your skills.

Experienced candidates are very welcome with income adjusted to your experience.

Candidates with a PhD that are looking to complement their experience are also welcome.

A reasonable level of English is required. French is not mandatory and INRIA will provide French classes if needed.

To apply submit you CV, references, recent marks, and if available your last Internship/Master Thesis manuscript. With your application provide any element (github account, code snippets, etc.) that could help us assess you skills beyond your academic record, as well as a few references of persons we can contact to get some feedback on your qualities.

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 (90 days / year) and flexible organization of working hours (except for internship)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage under conditions

Rémunération

From 2,692 € (depending on experience and qualifications).

Informations générales

- **Thème/Domaine** : Calcul distribué et à haute performance
Calcul Scientifique (BAP E)
- **Ville** : Saint Martin d'Herès
- **Centre Inria** : [Centre Inria de l'Université Grenoble Alpes](#)
- **Date de prise de fonction souhaitée** : 2025-02-01
- **Durée de contrat** : 1 an, 11 mois
- **Date limite pour postuler** : 2025-01-07

Contacts

- **Équipe Inria** : [DATAMOVE](#)
- **Recruteur** :
Raffin Bruno / bruno.raffin@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.

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

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.