Offre n°2025-08575

PhD Position F/M Reliability Enhancement of Post-Von Neumann Hardware Accelerators

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

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

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 PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

Mission confiée

Artificial Intelligence (AI) is increasingly indispensable across various society sectors due to its potential to transform conventional applications, from smart homes to safety-critical systems like autonomous driving and space exploration. Deep neural networks (DNNs) are state-of-the-art AI methods that outperform other approaches in language processing, image and video classification, audio and radar processing, and instance segmentation [1–3]. Notably, DNNs such as OpenAI GPT-4, Meta LLaMA2, and Mistral Mixture of Experts have captivated public interest with their high accuracy.

Due to their resource-intensive nature, DNNs require powerful dedicated hardware accelerators, such as GPUs and TPUs. However, large hardware accelerators are unsuitable for embedded safety-critical systems due to their high energy consumption. New unconventional accelerator architectures like the ones based on PIM [4] and neuromorphic computing [5] have been proposed for complex DNN deployment in critical applications where power and performance are critical requirements, offering energy-efficient alternatives to traditional GPUs and TPUs. However, their reliability, particularly against radiation-induced faults, remains to be fully assessed.

In this Ph.D., we will identify hardware and software vulnerabilities in PIM accelerators for DNNs and propose fault mitigation techniques.

Principales activités

Activities

The Ph.D. student will characterize the radiation-induced impact on system reliability for different DNN model architectures and how the acceleration that PIM enables impacts the final error rate. The results will be combined with software simulation data for a detailed fault propagation analysis, aiming at deploying effective hardening solutions tailored for PIM executing DNNs.

The Ph.D. student will participate in international experiments and internships at laboratories like Rutherford Appleton Laboratory in the UK and Los Alamos National Laboratory in the USA. The student will participate in conferences and international projects and have their research published in prestigious scientific venues. This will help them develop their research skills and network with professionals in their field.

Additional information about the city and the university

Rennes is a vibrant and student-friendly city in northwestern France. The city has a thriving student culture, with plenty of bars, restaurants, cultural events, and an affordable cost of living. Additionally, *Rennes is evaluated as one of the best cities to live in Europe* [6].

Rennes is home to the University of Rennes, one of the largest universities in France. The University of Rennes has a strong focus on innovation and technology. It is home to many world-renowned research institutes, including INSA, IRISA, and INRIA Rennes. These institutes offer a wide range of Ph.D. programs in computer science, covering various topics such as artificial intelligence, machine learning, data science, and hardware and software engineering. Ph.D. students in Rennes benefit from close relationships with faculty and access to state-of-the-art facilities. The students also have the opportunity to collaborate with leading researchers worldwide.

Team's LinkedIn page: TARAN's LinkedIn

References

[1] Xiaohua Zhai et al., Scaling Vision Transformers, IEEE/CVF CVPR, 2022

[2] Chong Chen, et al., Compound fault diagnosis for industrial robots based on dual-transformer networks, Journal of Manufacturing Systems, 2023

[3] Yuxin Fang, et al., EVA-02: A Visual Representation for Neon Genesis, CVPR 2023

[4] Laguna A. F. et al., In-Memory Computing based Accelerator for Transformer Networks for Long Sequences, IEEE DATE 2021

[5] Yao M. et al., Spike-driven Transformer V2: Meta Spiking Neural Network Architecture Inspiring the Design of Next-generation Neuromorphic Chips, ICLR 2024

[6] European Commission, Quality of life in European cities, 2024, https://ec.europa.eu/regional_policy/information-sources/maps/quality-of-life_en

Compétences

Mandatory skills (the candidate must have):

- *Strong knowledge of computer architecture*
- *Strong programming knowledge of C/C++ and Python*

Optional skills:

- HW design: VHDL/Verilog basics, HW synthesis flow
- Basics of Machine Learning
- Experience with High-Level Synthesis (HLS) is a plus
- Experience in fault-tolerant architectures is a plus
- Knowledge of compilers and LLVM is a plus

Languages: proficiency in written English and fluency in spoken English. *The interviews for the PhD will be in English.*

Relational skills: the candidate will work in a research team, where regular meetings will be set up. The candidate has to be able to present the progress of their work in a clear and detailed manner.

Other values appreciated: Open-mindedness, strong integration skills, and team spirit.

Most importantly, we seek highly motivated candidates.

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 2200€

Informations générales

- **Thème/Domaine :** Architecture, langages et compilation
  Système & réseaux (BAP E)
- **Ville :** Rennes
- **Centre Inria :** Centre Inria de l'Université de Rennes
- **Date de prise de fonction souhaitée :** 2025-06-01
- **Durée de contrat :** 3 ans, 2 mois
- **Date limite pour postuler :** 2025-07-31

Contacts

- **Équipe Inria :** TARAN
- **Directeur de thèse :**
  Fernandes Dos Santos Fernando / fernando.fernandes-dos-santos@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'e?orce 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

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

For more information, please contact Fernando Fernandes dos santos: fernando.fernandes-dos-santos@inria.fr

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.