



Offre n°2025-08811

PhD Position F/M Accelerating Hardware Coherence Using Programmer Input in Multi/Manycore Systems

Type de contrat : Fixed-term contract

Niveau de diplôme exigé : Graduate degree or equivalent

Fonction : PhD Position

Niveau d'expérience souhaité : Recently graduated

A propos du centre ou de la direction fonctionnelle

The Centre Inria de l'Université de Grenoble groups together almost 600 people in 24 research teams and 9 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 PhD is funded by the Défi Inria Cocorisco (<https://project.inria.fr/cocorisco/>), which brings several Inria teams and CEA together to build high performance platforms based on RISC-V through HW/SW interactions. As a result, this posting

will entail regular travel for project meetings (around once a year).

Mission confiée

To improve performance, a general purpose processor associates a private cache memory to each of its cores, in order to keep a subset of data close to the execution units of the core and thereby accelerate data access. The chip also features a larger cache shared by all cores. This architecture is depicted in Figure \ref{fig:coh}.

To facilitate the development of parallel applications, the various cache memories implemented on chip provide data coherency. That is, a given memory address may be cached in several private caches only if the associated cores are only reading the data. Any write to the data requires invalidating all existing copies in the private caches (except for that of the writer). A cache that lost its copy following a write will have to reload the new version of the data (e.g., from memory). In general purpose processors, coherency is handled by hardware and is completely transparent to the programmer. If that were not the case, software would have to manage coherency explicitly for the program to be correct, which would significantly slow parallel application development down.

However, the hardware handling coherency is forced to make sub-optimal choices as it does not have a global vision of access and sharing patterns across the system. For instance, when data is loaded from memory, a core can ask to insert it into its private cache either with read permission or both read and write permissions. Depending on the access and sharing patterns, the correct decision (for performance or chip traffic) is not always the same. If the data is not shared, the core should obtain write permission in order to avoid sending a second request asking for write permission in the future. If the data is shared but the core is only going to read it, then, the core should ask for read permission only. Asking for write permission would imply invalidating the other copies and would be wasteful in this case. In addition, generally speaking, depending on the number of cycles between when a data is read by a core and when it is written by that core, it can be more interesting to obtain the write permission early (when reading) in order to not slow down the write if it is close in time, or late (when performing the write), in order to allow other cores to keep their copies as long as possible.

Those access and sharing patterns are known by the developer. It would therefore be interesting to express those patterns in the source code to help the hardware make correct decisions at runtime, rather than just trying to guess what that decision might be. The RISC-V instruction set being open source and extensible, it provides us with an opportunity to study this technique, by adding instructions conveying with what access and sharing patterns a data is being manipulated.

Principales activités

The thesis is built around three items:

- First, reviewing the literature on sharing patterns in multicore programs as well as hardware techniques to identify them will allow the candidate to identify patterns that we would want to express via dedicated instructions.
- Second, quantifying the frequency at which such patterns occur at runtime in typical multicore workloads (PARSEC) will be needed, in order to confirm the usefulness of such patterns and prioritize which patterns to support.
- Finally, the candidate will study the performance gain that the introduction of new instructions will bring by i) Adding support for those instructions in gcc or LLVM (through intrinsic) ii) Adding those instructions in the PARSEC benchmarks and iii) By simulating PARSEC benchmarks on a multicore processor model to which the candidate will have added support for the new RISC-V instructions, both in the processor core and the blocks handling cache coherency. For this last step, a high level simulator will be used (gem5), and hardware will not be developed using VHDL or Verilog.

Compétences

Personal Skills

- The candidate should expect to be autonomous in developing software, experiments, and analyzing results.
- The candidate should be able to clearly express their ideas and conclusions, and to motivate their research directions.
- The candidate should be open to constructive criticism from their peers and supervisors.

Technical Skills

- Programming: C/C++. Strong knowledge and understanding of data structures, testing and debugging tools.
Linux scripting: Python, bash or other.
- At least basic level in computer architecture (caches, virtual memory, pipelining) and Instruction Set Architecture concepts. Advanced level is a plus.
- At least basic level in cache coherence concepts (coherence protocols, snooping vs. directory)
- Understanding of synchronization concepts for parallel programming (threads, locks...)
- Strong english level (B2) is required as scientific articles are written and presented in english.

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) 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

Rémunération

2200€ gross salary per months

Informations générales

- **Ville :** Grenoble
- **Centre Inria :** [Centre Inria de l'Université Grenoble Alpes](#)
- **Date de prise de fonction souhaitée :** 2025-09-01
- **Durée de contrat :** 3 years
- **Date limite pour postuler :** 2025-07-31

Contacts

- **Équipe Inria :** AT-EQUIPE-GRA (DIR-RAL)
- **Directeur de thèse :**
Fuguet Tortolero Cesar / cesar.fuguet-tortolero@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 should be aware that a PhD programme is vastly different from a BS or MS programme. Pursuing a PhD requires strong motivation and the ability to focus on a specific topic for three years.

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

Applications must be submitted online via the Inria website. Processing of applications submitted via other channels is not guaranteed.

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.