RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

FRANCE 2030 — PROGRAMME DE RECHERCHE NUMÉRIQUE POUR L'EXASCALE

Fault-tolerant scheduling strategies for iterative algorithms

M2 Internship — starting Spring 2025
PhD thesis — starting Fall 2025

Context: Several error sources may impact the execution of iterative algorithms on large-scale platforms. They include fail-stop errors, that are immediately detected, and silent errors (a..a silent data corruptions), that can be detected through some verification mechanism. Fail-stop errors correspond to permanent failures, e.g., processor crashes. Silent errors are disruptions that strike and stay undetected until they manifest eventually through strange application behavior. Silent errors arise from two main sources: computation errors and memory bit-flips. Protecting algorithms and software libraries from all these errors is a major concern within the HPC community.

The standard way to deal with fail-stop errors is checkpoint-restart, and the optimal checkpointing period is well-known, at least for memoryless IID error inter-arrival times. However, mitigating the impact of silent errors remains an open challenge. On the one hand, replication (or even triplication to avoid a sequential re-execution) does a perfect job but at a prohibitive cost. On the other hand, numerous application-specific detectors have been introduced, such as Algorithm-Based Fault Tolerance (ABFT) checksums, recomputing a residual, checking orthogonality of some vectors, applying space and time filters across a neighborhood, etc.

These detectors are usually limited to a particular error type. A major problem is that they may well either fail to detect some errors, or raise many false alarms. In other words, these detectors are not perfect: their recall and precision are not at 100%. Most, if not all published works assume perfect detectors, which is not realistic.

Objectives: The first (and main) objective of this internship is to design and assess scheduling strategies based upon a combination of checkpoints and imperfect detectors to guarantee protection from a single source of silent-errors with a high probability. This requires to introduce some assumptions, such as upper bounding the latency of the detection, or to introduce randomized tests on the data.

The second step (that may come later during a PhD) is to provide a resilient holistic methodology to protect iterative algorithms from all error types, namely fail-stop errors and all sources of silent errors.

This internship is expected to continue with a PhD thesis, for which funding from the NumPEx PEPR program is already secured.

The internship (and potential PhD thesis) will take place at ENS Lyon in the LIP laboratory and ROMA team (https://www.ens-lyon.fr/LIP/ROMA/).

Prerequisites: Some knowledge in algorithm design, complexity, and probabilities. The work is on the algorithmic side of the problem, with potential simulations to validate the results.

To apply, please send to Anne.Benoit@ens-lyon.fr the following documents, in a single pdf file:
(a) motivation letter; (b) detailed CV; (c) academic transcripts (bachelor/master); (d) one-page summary of the research work conducted so far, and list of publications (if any); (e) at least two recommendation letters.

Contacts:

- Anne Benoit, ENS Lyon, Anne.Benoit@ens-lyon.fr
- Yves Robert, ENS Lyon, Yves.Robert@ens-lyon.fr
- Co-advising also comes from Emmanuel Agullo, Inria Bordeaux, Emmanuel.Agullo@inria.fr

cea cnrs Inria anr agence nationale de la recherche