
Optimal resilience patterns to cope with fail-stop and silent errors

Aurélien Cavelan^{*†}

¹ROMA (ENS Lyon / CNRS / Inria Grenoble Rhône-Alpes) – CNRS : UMR5668, Laboratoire d'informatique du Parallélisme, École Normale Supérieure (ENS) - Lyon, INRIA – Laboratoire de l'Informatique du Parallélisme 46 Allée d'Italie 69364 Lyon, France

Abstract

This work focuses on resilience techniques at extreme scale. Many papers deal with fail-stop errors. Many others deal with silent errors (or silent data corruptions). But very few papers deal with fail-stop and silent errors simultaneously. However, HPC applications will obviously have to cope with both error sources.

This paper presents a unified framework and optimal algorithmic solutions to this double challenge.

Silent errors are handled via verification mechanisms (either partially or fully accurate) and in-memory checkpoints. Fail-stop errors are processed via disk checkpoints. All verification and checkpoint types are combined into computational patterns. We provide a unified model, and a full characterization of the optimal pattern. Our results nicely extend several published solutions and demonstrate how to make use of different techniques to solve the double threat of fail-stop and silent errors. Extensive simulations based on real data confirm the accuracy of the model, and show that patterns that combine all resilience mechanisms are required to provide acceptable overheads.

*Speaker

†Corresponding author: