10th Seminar on Modeling  
4 February 2014

Generative programming and active libraries for cosmological simulations

Vincent Reverdy  
Ph.D. student at Observatoire de Paris (LUTH)

Since 1998, the question of the origin of the accelerated expansion of the Universe has become one of the most fundamental open problem in cosmology. To investigate this question, numerical simulations of large scale structure formation such as clusters of galaxies and filaments are a particularly relevant tool. In this presentation, I will review the current status of these simulations and the related computational problems. I will focus on the full universe run that has been carried out on the entire Curie supercomputer in 2012. Then, I will highlight the lessons taken from this simulation and the ongoing library development based on generative programming to address most of the issues of current cosmological codes. The goal is to avoid the intertwining of physics, algorithms, and parallelization, and to make the best of supercomputers by creating a "compiler inside a compiler". I will describe these techniques and explain their substantial advantages for large cosmological simulations.

Efficient failure containment for large scale MPI HPC applications

Thomas Ropars  
Researcher at École polytechnique fédérale de Lausanne EPFL

As HPC systems keep growing in scale, providing efficient fault tolerance mechanisms becomes a major issue. Studies on future exascale systems highlight that, considering that the expected meantime between failures will range from one day to a few hours, simple solutions based on coordinated checkpoints saved to a parallel file system will not work: more time will be spent dealing with failures than doing useful computation. As a consequence, new checkpointing techniques should be designed. A checkpointing protocol for large scale HPC systems should provide good performance in failure-free execution and in recovery while limiting the amount of resources used for fault tolerance. Designing a solution that can achieve all these conflicting goals is a hard task. In this talk, I will introduce hybrid rollback-recovery protocols as a solution to this problem: A hybrid rollback-recovery protocol combines coordinated checkpointing with some message logging to provide failure containment. In a first part, I will explain why hybrid protocols can be efficiently applied to most MPI HPC applications. In a second part, I will present SPBC, our new checkpointing solution based on an hybrid protocol. SPBC is the first checkpointing solution that provides failure containment without logging any information reliably apart from process checkpoints, and this, without penalizing recovery performance. To achieve this result, we used an original approach: Instead of designing a protocol that works for all message-passing applications, we identified properties common to our target applications, namely MPI HPC applications, and we leveraged these properties to design a fault tolerant solution that can be more efficient than existing protocols at large scale.