2SC7694 – Energy optimization and acceleration of a cloud financial calculation graph

Instructors: Stephane Vialle
Department: DOMINANTE - MATHÉMATIQUES, DATA SCIENCES, DOMINANTE - INFORMATIQUE ET NUMÉRIQUE
Language of instruction: ANGLAIS
Campus: CAMPUS DE PARIS - SACLAY
Workload (HEE): 80
On-site hours (HPE): 48,00

Description

Project topic in partnership with ANEO.

Application Context

Modern **insurers** have a highly regulated but at the same time relatively broad field of activity: different types of insurance, banking services, etc. One of the difficulties in assessing the accounts of an insurance company (or bank) lies in the valuation of financial assets (e.g. EDF shares, life or car insurance contracts, etc.) and the underlying risks. The approach generally used consists of evaluating for each asset the cost of a devaluation as well as the associated risk. For simple assets such as shares, the calculation is simple. For more complex assets such as insurance products or derivatives, the calculation is more complex since it is usually based on the consideration of many factors. Depending on the valuation of the risks taken, regulations resulting from various economic crises, such as Solvency II or IFRS17, require the insurance or bank to tie up a certain amount of equity capital. Certain risks may cancel each other out between different assets (e.g. the risk of a life insurance asset based on the euro/dollar rate may be covered by other assets based on the euro/yen and yen/dollar rates). In order to maximise this potential for offsetting through the assets owned, these bodies will consolidate the accounts on the widest possible scale, the group as a whole. That is, they will carry out risk analyses as if all assets belonged to a single entity. One of the difficulties of the exercise then consists in distributing the capital requirement among the different legal entities whose accounts have been consolidated, so this process is actually more complex than a simple pooling of assets followed by a global risk analysis.

The ***process of constructing consolidated accounts for an insurance company therefore generates numerous calculations***. These calculations concern, on the one hand, the modelling of the cost of repayment of

contracts according to various factors and, on the other hand, the modelling of investments made with the money available. As an example for a life insurance contract, risk modelling is based on mortality tables provided by INSEE and taking into account different factors such as geography, socio-professional category, family situation, etc. To do this, the life of the contract is simulated year after year in order to take into account changes in these factors. Different scenarios are played out in order to reflect all possible changes in situations (moves, changes in family situation, etc.). These scenarios are then aggregated. This process is of course a simplified view and does not take into account various elements such as the aggregation of contracts in order to reduce the volume of calculations, which aggregation is in itself the subject of various optimization works. Other elements of the process include the consolidation of asset/liability risks by contract type, consolidation by legal entity and taking into account the specific regulatory requirements of each country, and the use of these simulations to optimise the risk of the contracts offered as well as their price.

Problem addressed in this project

The process of constructing the consolidated financial statements takes several weeks and includes calculation steps as well as manual steps; we will consider the latter here as instantaneous. The calculation steps correspond to the equivalent of 413177 hours of calculation time, i.e. just over 10 full days on an infrastructure of 1700 cores. However, in reality, the process cannot take place in 10 days on such an infrastructure because of the dependencies between the computing tasks: there are times when there are not enough tasks to occupy the grid. A fine analysis of the dependencies shows that the critical path duration is 11h30. This duration would be that of the whole computation if an infrastructure of infinite size was available.

Project subject: In order to optimize costs without investing in a very large computing grid that would ultimately be little used, we want to use on-demand resources available in the cloud. To make the most of this, **we want to optimize the execution of the task graph** by searching for :

- The best strategy for switching compute nodes on and off.
- The best scheduling of tasks on the available nodes.

The study will have to take into account the following elements:

- The dependencies between tasks
- The duration of the tasks, known in advance
- The duration of the transfer of results between tasks (only the relevant files will be listed).

It should be noted here that the cost mentioned can be energy as well as financial, and that the two are closely linked: in use, more than half of the cost of owning a computing infrastructure corresponds to the cost of