Mat-2.4177
Seminar on case studies in operations research

Simulating Insurance Portfolios using Cloud Computing

Project Plan
February 23, 2011

Client: Model IT

Jimmy Forsman	69074L (Project Manager)
Tuomo Paavilainen	69255A
Topi Sikanen	55670A
Taneli Silvonen	69318H

Contents

Introduction

The classic interpretation of probability defines the probability of an event to be its volume relative to the possible outcomes. The Monte Carlo method in its simplest form generates random inputs from a specified distribution, and uses these to compute the outcomes. Statistical parameters, such as the expectation value and variance, can be calculated from the outcome distribution. Monte Carlo simulation is useful when modeling events with significant uncertainty in inputs, so an exact deterministic algorithm can't be constructed. This includes for example the calculation of risk in finance.

Model IT has developed a product called mSII for simulating insurance portfolios. The objective is to help insurance companies to fulfill the demands of Solvency II, a set of regulatory requirements set by the European Union. It uses a Monte Carlo method for the simulations. This requires a lot of computing power since the events in every contract for all the clients are simulated, usually a 100 times for every time period, to get a probability distribution. For example the lifetime of a life insurance contract can be over 60 years. The advantage of the Monte Carlo method is that it provides the total distribution, not just the expectation or average values. [1]

The term cloud computing refers to offering hardware and software as a network service, usually in the Internet. It originated from the abundance of redundant capacity. For example, Amazon's IT infrastructure must handle the Christmas season orders, but most of the year the traffic is much lower. Now they offer the excess capacity for customers, and bill by the usage of computing, storage and network services. One type of cloud service is called Infrastructure as a Service (henceforth referred to as IaaS). This means using virtualized components of the service provider, for example processors, which reduces the user's need to own computing infrastructure.

Objectives

The main objective of this project is to find out whether cloud computing is financially and technically a feasible solution to simulate insurance portfolios.

Insurance simulations aren't needed constantly, only occasionally to support the decision-making of an insurance company. This means heavy investments in computing power aren't necessary because the big grid infrastructure would run in idle mode most of the time. The idea of total scalability in computing power, where you pay by the hour per computing unit, makes cloud computing an attractive choice for the simulation needs. Since cloud computing is a rather recent concept, there aren't any ready-made solutions to be found in literature. This project's goal is thus to examine the suitability of the cloud in practice by testing and analyzing the results.

To answer at least some questions about cloud computing, it's essential to concentrate on a clearly bounded part of the problem. Our team has together with Model IT decided to focus on one type of an insurance and one service provider for the test simulations which are presented in more detail in the next chapter. The main IaaS providers are Microsoft and

1. http://www.modelit.fi/web/fi/Vakuutusanalyysi

Amazon of which the first is Model IT's partner, so we will concentrate more on their Azure service. Amazon's EC2 will be included in the pricing comparison.

Providing the testing with Azure doesn't face insurmountable problems, we want to provide a detailed guideline on how to simulate insurance portfolios in the cloud with MATLAB. This includes examining technical implementation and measuring performance with various tests. Since we're talking about a commercial product, there's also usability and pricing issues to be considered. For example, finding out how different data transfer and storage options affect the cost and performance, will be a point of interest.

In addition to Model IT and their clients this project's interest groups are MathWorks (the producer of MATLAB) and Microsoft. Although we're conducting tests and cost-benefit analysis concerning only insurance simulations, the results will apply also to high performance cloud computing in general. Thus one of the project's objectives is to examine the general feasibility of the cloud in heavy number crunching.

Tools and methods

mSII

mSII models insurance risk by using contract level simulation. The risk isn't trivialized to deterministic cash flows at any stage, but the cash flows depend on stochastic events. Numerical simulations are performed with full information. There are no deterministic assumptions, no aggregation of contracts and no simplifications on contract terms. The data obtained through simulations can be analyzed with MATLAB or exported. mSII takes the insurance portfolio, economic scenarios, balance sheet formulas and the company decision rules as inputs. The output includes information for the company's decision-makers and shareholders. Contract level simulation makes the model more simple but heavier to calculate.

After importing all the external variables, such as economic scenarios, the simulation goes through all the years. In the next sub-loop, which is the biggest, all the customers are gone through. Random events are simulated for each customer and the simulation goes through all the contracts the customer has. Random events are simulated for each contract as well, and they consist of surrenders or payments for example. This procedure is done for all the time steps in a year. After a full year, cash flows are calculated and ending conditions checked for the contracts. From this company balance sheets and decisions can be generated.

We have in our use a MATLAB test license and Model IT's simulation engine which is implemented as a MATLAB Toolbox.

Approach

As mentioned before, cloud computing is a relatively new concept so there aren't many reliable articles or other information sources available. Even so, our approach to this project is to conduct first some literature review on the topic of high performance computing (HPC) and cloud/grid computing. In addition, we go through articles about

how Monte Carlo simulation techniques have been applied in finance. Understanding these concepts provides a stable foundation to continue this project further. After getting a good grip of the topics considered, we start performing the test simulations.

When we have MATLAB and the toolbox configured and fully working, it's time to run the tests. Our idea is to concentrate on one rather simple insurance portfolio with only one type of insurance. For this purpose we use an example case called UnitLink used by Model IT to demonstrate the mSII. UnitLink's contract type is a life insurance which has four random variables: yearly fees, withdrawal of the contract by the customer, the customer's death and external financial factors. The portfolio is simulated with the Monte Carlo method by using some imaginary distributions for the random variables.

Computing time is an obvious choice for performance measure when conducting simulations. Questions regarding data transfer and usability will also be examined. Further simulations with more complex portfolios are possible assuming that there's enough time, and the team and the client feels that would give us some relevant additional information. It's possible that MATLAB computing is beneficial to compile into .exe format and this possibility may be studied as well.

Simulations will be run with several configurations. The number of clients, the time scale and the number of time steps in a year can be varied for the purpose of getting an idea of how they effect the computing time. Computations need to be heavy enough so that it makes sense to run them in a cloud.

Another issue is whether or not to use cloud provider's services to host data and code. Both the code and the data files must be located at the cluster. For a large code base and large data sets, transmitting code and data every time when using the cluster can be very time consuming. To avoid losing time, there's a possibility to buy data storage services. Usability of the service is something to be assessed based on experiences obtained in the process.

Financial aspects are considered too. Our objective is to construct an estimate for the expenses. After obtaining the estimate we study if it's cost-effective to run these simulations parallel, or are the advantages lost because of e.g. purchasing licenses for MATLAB. Cloud providers claim to obey "pay for what you use" -policy and have listed prices for example for computing ($/h) and for storage ($/GB/month). This information will be used to calculate expense estimates, even though pricing schemes may turn out to be more complicated than appeared in the first place. Running test simulations may provide further information on this matter as well.

Methods planned to be used in this project can be summarized as follows:

- Literature review on HPC and Monte Carlo simulation in finance. Justify the need for using these techniques in insurance contract analysis.
- Conducting test simulations is the backbone of this project. Subjects of interest are for example computing time and usability. Conclusions are made based on results and comparison with possible references.
- Calculations to find out the cost-effectiveness of cloud computing.

Responsibilities and schedule

The project schedule is vital for the successful execution of the project. The time and resources available for the project are relatively limited. Time invested by the team members form a serious constraint, as none of the team members are able to work full time. Computing resources are limited to the publicly available computers at the university and one laptop donated to the team. In addition to these, cloud computing resources will be provided by either Model IT or Microsoft. Unless the project schedule will be followed, it is likely that some of the risks mentioned in further this document may realize.

Based on the preliminary research and the objectives provided by the client, we came up with the following division of tasks and responsibilities. In the division, we emphasize the importance of executing the critical path tasks that need to be completed in order not to hinder the smooth realization of the project.

Task	Weeks	Resources
1. Literature review, simulation (Monte Carlo)	7-10	Jimmy, Tuomo
2. Literature review, High performance computing	7-10	Taneli, Topi
3. Configuring the technological environment for testing	9-10	Tuomo, Taneli
4. Determining test portfolios	11-13	Jimmy, Topi (all)
5. Running test portfolios, gathering results and the preliminary analysis of the results.	13-14	Tuomo, Taneli
6. Comparison of the results with the literature and conclusions	15	Jimmy, Topi
7. Cost analysis of the pricing options based on the results (5.)	16	Topi
8. Reporting findings and results. Communication with the client	All the time	Jimmy
9. Conducting a final report	15-18	Everybody

The purpose of the tasks one and two is to provide the team members with a sufficient insight to the logics and principles of both Monte Carlo simulation and High performance computing. All members will be familiar with both concepts, but two team members will be specializing in simulation and the other two in high performance computing, thus subdividing the workload. Some literature dealing with Solvency II should also be covered.

Task three comprises of setting up MATLAB and its relevant toolboxes. It also includes the configuring of cloud computing interfaces, and finding out how simulation tasks can be sent to the cloud to be calculated. It is also necessary to get to know the syntax of mSII and creating simple test portfolios. Once we have a certain level of understanding of how the simulations behave, the test portfolios will be determined and designed. Although this task is relatively short and simple, it is of great significance, as major results will be based on the decisions made at this stage. It is untenable to return to this task, when the calculations have already been made.

Task five on the other hand comprises of the actual running of the simulations determined in the previous task. The results ought to be gathered and converted into a proper format allowing us to analyze the results of the simulation. Some preliminary analyses may also be done, so that more specific calculi can be performed subsequently. In task six, conclusions will be drawn and potential future research subject identified. Finally in task seven, the cost analysis mentioned in the previous chapter is conducted. The allocation of the tasks to resources is depicted in Figure 1.

Resources	Week											
	7	8	9	10	11	12	13	14	15	16	17	18
Jimmy F	1	1	1	1	4	4	4	4	6	9	9	9
Tuomo P	1	1	3	3	4	4	5	5			9	9
Taneli S	2	2	3	3	4	4	5	5			9	9
Topi S	2	2			4	4	4	4	6	7	9	9
Computer							5	5				

Figure 1. Allocation of resources to tasks

Risks

Technical difficulties

We identify technical difficulties as the main risks in this project. Due to the novel nature of our undertaking there is a high risk of running in to insurmountable technical difficulties. It is therefore imperative that the technical implementations are tested as soon as possible so that possible obstacles can be identified in timely manner. This allows us to redefine our project goals if the original goals turn out to be unattainable in the given time.

Problems finding information

Cloud computing is a relatively young field and very little public information is available on running simulations on the cloud. Finding useful information, not just opinion pieces, may turn out to be difficult.

Too little time

Any technical difficulties encountered with the technical implementation will easily result in major delays. This risk can be managed by starting the implementation process immediately and possibly limiting our study to a single cloud computing environment.

Difficulties in cost analysis

Due to complicated pricing schemes used by cloud computing providers it might be difficult to offer a definite conclusion on the cost-effectiveness of cloud computing for simulation of insurance portfolios. This risk can be also mitigated by focusing on a single service provider. Furthermore, if cost-effectiveness cannot be assessed, at least information on feasibility of such endeavor is gathered.