Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

SIMULATOR OF AMAZON EC2 SPOT MARKET

Przemysław Szufel
Bogumił Kamiński

Warsaw School of Economics
Al. Niepodległości 162
02-554 Warszawa, Poland

ABSTRACT

Running simulation experiments often requires significant amounts of computations. Public computing clouds are a cost-efficient way to purchase computing power on-demand. In particular Amazon offers a spot pricing mechanism for its public cloud service called EC2. This model has a very attractive pricing but does not guarantee uninterrupted computations. Moreover, in order to maximize their benefits the users have to be active players. In our work we develop a simulator of Amazon EC2 spot price market. We use the simulator to show that there is a trade-off between the computation cost and the computation time.

1 INTRODUCTION

Amazon is the largest provider of public cloud services. The Amazon Elastic Compute Cloud (Amazon EC2) service is particularly useful for running large scale simulation as parallel processes on a computing cluster.

Several analyses of Amazon spot pricing model have already been performed including statistical analysis (Javadi, Thulasiram, and Buyya 2013) or optimal bidding strategies – e.g. (Tang, Yuan, and Li 2012). However some results such as "do not bid to low" are not optimal any more, because Amazon is constantly changing its pricing regime including the recent price reduction in April 2014. Moreover, the exiting reports do not take into account the value of "free lunch" that can be accrued due to not billing for a partial hour terminated by Amazon. It is also important to consider the fact that the user has to pay the booting time of the virtual machine.

2 SIMULATING EC2 SPOT PRICING

Amazon offers 24 different machine types on 21 availability zones within 8 regions on 4 continents. Since the prices are lowest in the USA and the EU we consider only 11 availability zones in those regions. We perform our analysis starting from the latest EC2 price regime change that took place April 1^{st}, 2014.

The users utilize the spot mechanism by placing a bid for an instance of a particular machine type in a selected availability zone. If the bid price exceeds the a price the instance is being launched. While an instance is running, the spot prices are continuously changing reflecting changes in supply and demand for the computing power across different zones. However, a user is always charged the price from the beginning of each full hour. Whenever the spot price exceeds the bid price the instance is being *terminated* without any warning by Amazon. In our analysis we assume that the user is employing some check pointing mechanism (e.g. Amazon EBS, S3, RDS) to store simulation results after each simulation run. The user is not charged for a partial hour when it ended due to termination – we call this situation "free lunch".

We have created a EC2 billing simulator in Python. Our tool considers server booting times and possible terminations due to spot price fluctuation. The simulator enables testing various bidding strategies against historical price data. In Section 3 we present how it can be used to test different strategies against historical data.

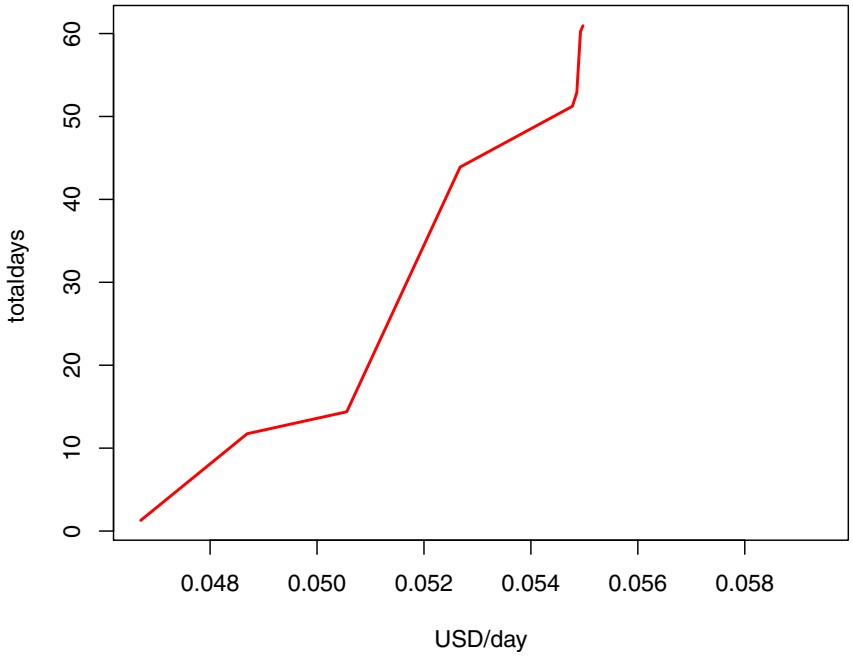

Figure 1: Scatter plot of costs vs running time for different combinations of machines and zones.

3 AMAZON EC2 SPOT BIDDING STRATEGY COMPARISON

Various machine types offered by Amazon differ by their computational power measured in ECU (EC2 Compute Unit). We standardize the machine spot prices by dividing them with their ECU (thus we analyze price per hour per ECU for each machine).

The minimum observed price on Amazon EC2 market is 0.0022857 USD per hour×ECU – spot prices never fall below this level across all machine types. Hence, the minimum reference price is 0.05485 USD per day×ECU. Our analyzes is limited to computational (c3.*) and general (m3.*) instance types (9 in total) across all available EU and USA zones (11 in total). For the simulation scenario we want to compare results of bidding at price 0.00251427 USD per hour×ECU (10% above theoretical minimum) across different machines and zones (99 options in total). We measure two outcomes for a bidding strategy:

(1) average daily costs per one ECU of effective computing power;
(2) computing time that was actually available during 61 days of our simulation experiment.

Figure 1 presents a scatter plot of USD/day and total time of all offers. The red curve denotes the Pareto frontier. We can see there is a clear trade off between the computation time and the average cost. Additionally a lot of machine-zone combinations are far from being efficient. The efficient startegies in bottom left cornet take the benefit from "free lunch" possibility (notice that their average price is lower then the minimal reference spot price). However, such strategy significantly decreases the effective simulation time.

REFERENCES

Javadi, B., R. K. Thulasiram, and R. Buyya. 2013. "Characterizing spot price dynamics in public cloud environments". *Future Generation Computer Systems* 29 (4): 988–999.

Tang, S., J. Yuan, and X.-Y. Li. 2012. "Towards Optimal Bidding Strategy for Amazon EC2 Cloud Spot Instance". In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on*, 91–98. IEEE.