

AWS Research Proposal

Zichen Xu
xu.925@osu.edu

We propose a framework for geo-diverse data services hosted on EC2 spot instances. Spot instances implement market driven pricing for spare resources within Amazon’s data centers. On average, they can be 78% cheaper than instances provided with fixed, on-demand pricing [1]. However, it is challenging to serve data from spot instances because their prices can change every hour. Instances hosting critical data can be suspended with little warning when spot prices change. We studied a trace of spot prices provided by Amazon and observed that prices change more than 300 times per month. Further, the relative cost of spot instances in different regions changes more frequently. Naively migrating data to sites with low costs would incur prohibitive bandwidth costs. Consistent hashing, a widely used approach for data replication, would also incur significant migration costs. Thus, it is not tailored to geo-diverse settings where latency aware placement is needed.

Our cost-aware data replication framework uses online data replication to reduce migration costs and make wise decisions regarding price volatility. The key insight is that price volatility and non-uniform access rates magnify the cost for poor replication policies on popular data. If we target these heavy hitters, by predicting them and carefully allocating resources, we can significantly reduce the total cost. We have implemented our framework using novel intra- and inter-region data management policies. When considering data replication across regions, the framework forecasts the price at each site and replicate data to sites that combined to yield low cost. Such replication decisions are made online, i.e., when data is created (after a short profiling period), and thus avoids overhead by moving data frequently in response to the changing price. The framework manages the intra-region data replication to meet the dynamic workload. We built a 0+1 raid scheme to spawn new spot instances for workload peaks. That is, we maintain a service mirror in an on-demand instance in case the spot bid fails. We have evaluated our framework at a small scale, using up to 10 spot instances and 1 on-demand instance. The results show that, compared to consistent hashing, our approach reduces cost by 80% while increasing response time by less than 5%. All queries are served without failure reported.

This proposal requests AWS grant to scale up our proposed system. To test the performance of our cost-aware data replication framework in EC2, we plan to deploy our data services into about 20 instances from two tiers (e.g., t2.small and m3.medium) across five regions in US, Europe and Pacific Asia. The inter-region data transfer would be less than 200GB per month, mostly for the data replication transmission. Thus, the total cost for such setup, based on AWS’s calculator, is around \$3665 per month [2]. The whole project planned in EC2 will last for about one month.

[1] Amazon, "Amazon EC2 Spot Instances", <http://aws.amazon.com/ec2/purchasing-options/spot-instances/>, 2015

[2] Amazon, "Amazon AWS simple monthly calculator", <http://calculator.s3.amazonaws.com/index.html#key=calc-09D66061-E507-4F04-97FE-28A5FBAE55ED>, 2015