Survey on Virtual Machine Allocation in Cloud Infrastructure

Ashwini Ukarde
CSULB ID – 026328001

Saishree Jayakumar
CSULB ID – 026617056

Neha Bhoi
CSULB ID - 026627560

Progress Report

As part of the survey of Virtual Machine Allocation in Cloud Datacenters, we have selected 6 algorithms for survey under different aspects including but not limited to Energy Efficiency, Resource Allocation and Time efficiency. Out of that, we did a survey on 3 of the algorithms (K Means, Bin Packing - Reordering Grouping Genetic and Interior Search Algorithm) among literary papers. We were able to complete 48% of the survey.

Reordering Grouping Genetic Algorithm (RGGA) comes under energy efficient and optimized resource allocation algorithm (RGGA). RGGA is genetic algorithm to solve the Multi-Capacity Bin Packing Problem (MCBPP). RGGA used first fit ordering which naturally avoids generating infeasible solutions. An experiment is done by creating artificial loads on VMs. Which shows RGGA uses 8.4% less energy than first fit and 6.84% less energy than worst fit. This experiment considers that the loads on virtual machines are deterministic and known.
As we know one of the weakness of GAs is the speed it takes to run them. If we look at the amount of time needed to solve an answer RGGA scales linearly within this range as the problem size increases. Moreover, data centers of even 500 servers can be solved in fewer than 10 minutes on a single core.

Interior Search Algorithm: With increasing demands, cloud datacenters suffer from high energy consumption to satisfy the requirements for the users. It has a butterfly effect of increasing the infrastructure cost and also intoxicating the surroundings with carbon emission. There is an imminent urge to reduce the energy consumption on these datacenters. We did a deep survey on various algorithms which concentrate on energy efficiency and compared it against our Interior Search algorithm (ISA). ISA is a global optimization algorithm uses two groups – composition group being used for finding various solutions and mirror group placed between fittest elements to find improved perfect view.
The experimental results show that energy consumption of Best Fit Decreasing (BFD) and Genetic Algorithm (GA) is around 90 – 95% comparing with energy efficient ISA consuming only 65% which saves nearly 30% of energy. Also, ISA uses only one hyper parameter(α) which is ideal to tune compared to other traditional optimization algorithms. ISA as a global optimization algorithm allows us to solve complex multi-dimensional optimization problems and also has the flexibility and efficiency to adapt to varying types of problems even in NP-hard situations.

K-means algorithm is a time efficient algorithm for virtual machine allocation. In this survey K-means algorithm is used to deal with Infrastructure as a Service (IaaS) cloud service in which resources on the virtual machines can be used like resources on real machines. In K-means algorithm used clustering technique which divide the large dataset into small groups that have similar characteristics. It will select the points as initial centroids K where K is the number of clusters specified by the user. Every point is assigned to nearest centroid and each set of points is called cluster. Centroid of this cluster is then updated based on number of points in the cluster. This process continues till centroid is unchanged. Euclidian distance equation can be used to find the distance between centroid and points. In this survey, number of clusters determined dynamically according to number of data centers to number of hosts. Datacenter information received from Cloud Information Service and it perform a query to cloud provider after this number of K clusters are determined by the broker. Depending upon the number of hosts, number of K clusters in a datacenter are determined. After that clustering against the virtual machines is done using k-means algorithm. According to clustering result broker perform the virtual machine allocation on host and datacenter. Once the VM is allocated broker begins sending cloudlet request to VM. This algorithm is using space shared schedular for cloudlet scheduler where cloudlet1 must complete execution first before executing cloudlet2 and so on. The algorithm results show that it is Time efficient as the idle time is less comparative to traditional FIFO method also CPU utilization is better in K-means algorithm.

Challenges:

Reordering Grouping Genetic Algorithm: One challenge over here is that this algorithm considers workloads for VMs are deterministic. But in reality, these are non-deterministic loads.

Interior Search Algorithm: Even though ISA algorithm solves for energy efficiency and better resource optimization, one of the challenges here is that generally optimization algorithms has to run multiple iterations to converge to a global optimum there affecting the response time for VM creation/migration. There is a very less literary survey about it.

K-means algorithm: Analysis of actual implementation of algorithm

Possible solution:

Reordering Grouping Genetic Algorithm: Perform an experiment with simulation of non-deterministic loads and compare the results

Interior Search Algorithm: Literary survey shows that the optimization algorithms like Gradient Descent can be parallelized to reduce the time complexity with fewer margins of errors, we believe that taking the same approach to ISA will help reduce the latency.

K-means algorithm: Reading more papers to understand the exact behavior of the algorithm.