Virtual machine migration: Sometimes a physical machine gets overloaded and there arises a need to transfer VMs to another physical server to balance the load. While designing load balancing algorithms it is important to consider the time of migration and attack possibilities that are associated with VM migration.

Energy consumption: Load balancing algorithms in cloud computing should be designed in a way that minimizes energy consumption and therefore minimizes carbon emissions of virtual machines in data centers.

Qos performance metrics

Scalability: optimal resource allocation and usage

Overhead: communication between nodes requires user bandwidth, overhead

Performance: evaluation of system effectiveness

1. Static algorithms

Properties of Static algorithms:

• Their decision is based on a fixed rule for instance input load

• Static algorithms are not flexible

• Prior knowledge about the system is required

Optimal: optimal techniques used to gather resource info and allocate resources

Suboptimal: suboptimal techniques used

2. Dynamic
   Their decision is made on the current state of the system.
   • Dynamic algorithms are flexible
   • Performance of the system is improved

   Ant colony:
   Foraging Pheromone (FP): In our algorithm the ant would lay down foraging pheromone after encountering under loaded nodes for searching overloaded nodes.
   Trailing Pheromone (TP)-In a typical ACO the ant uses trailing pheromone to discover its path back to the nest.

   Make span: the time difference between the start and finish of a sequence of jobs or tasks

Fairness in machine learning refers to the various attempts at correcting algorithmic bias in automated decision processes based on machine learning models.

When all of the servers have comparable or identical performance and are operating with equal loads, the round robin method functions effectively. Performance suffers as the demand on the servers changes since the server with the fewest resources gets the next job even though it isn't yet ready to handle the previous one. It is necessary to create an algorithm that uses cutting-edge task allocation models to address this flaw.

The Min-Min algorithm does not provide fault tolerance or fairness. Its performance is good in small tasks. The algorithm also struggles with the starvation issue and is unconcerned about energy usage. The main problem for cloud providers is load imbalance, which is the biggest drawback. Future work on this will require the creation of an algorithm that shortens the processing time while improving resource efficiency.

In the Max-Min algorithm, as the requirements are known beforehand, the performance is better and throughput is high.

The processing of the OLB algorithm for static environments with centralised balancing is considered to be slow because it fails to calculate the current execution time. In the future projects, we must create an algorithm that determines the current execution time. In the future, we can use different crossover and selection procedures in genetic algorithms to provide outcomes that are more precise and effective. Ant colony algorithm is simple and less complex. In order to significantly reduce the time spent looking for candidate nodes, future work on the ant colony needs to examine the process for ant generation triggering and the strategy for pheromone update. The Carton algorithm requires low communication and its working is fair. The advantages and disadvantage of the discussed algorithms are presented in Table 1. Comparisons between these algorithms on the basis of the evaluation metrics is discussed in Table 2