

**OPTIMIZATION OF RESOURCE
ALLOCATION IN CLOUD COMPUTING USING
SERVICE BASED ALGORITHMS**

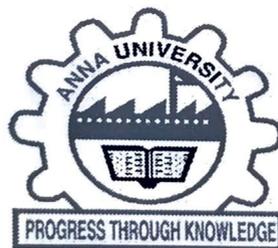
A THESIS

Submitted by

DIVYA R

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY



**FACULTY OF INFORMATION AND
COMMUNICATION ENGINEERING**

ANNA UNIVERSITY

CHENNAI 600 025

AUGUST 2021

ABSTRACT

Cloud computing offers a lot of computing, storage and memory based resources to the users on the 'pay for use' basis. Cloud Service Provider (CSP), users and the resources are entangled in the service cycle of cloud computing. All the Cloud computing services aim at profit and Quality of Service. The Profit of CSP depends on various factors. The foremost factor is the efficient usage of resources in cloud and it involves the process of scheduling and providing them to appropriate cloud server. Scheduling the CR (Cloud Request) is done based on the cost factors, to achieve best profit. The pricing model of cloud comes as the second factor. It is decided by the CSP, on their own standard terms and investment. Hence, optimizing the cloud based on resources and investment, plays an important role in achieving profit and efficiency.

Initial work deals with the deadlock prevention in Software as a Service (SaaS) cloud. An improved queuing model with Wait for resource algorithm is proposed as a solution of deadlock free cloud. For better resource utilization, the proposed model is designed to queue the arriving CRs in the first step and then a session is created between the CR and the servers with constraint on the number of CRs per session. Wait for resource algorithm pre-determines the resource needs of CR before allocation, due to this the deadlock is avoided. The proposed model gives 33 sec of average make span, 60% average resource utilization and a throughput factor of 1.3. It also gives least average waiting time and outperforms the load balancing algorithm, which produced varying response time for random applications running on cloud, without deadlock.

Next, the above proposed queuing cloud model is extended to support Platform as a Service (PaaS) incorporating virtualization to produce quick response time. To achieve this, nearest server allocation algorithm is proposed for cluster selection in the PaaS cloud. This improves the response time by scheduling the nearest server to the arriving CRs. The Nearest server allocation algorithm is written in java to perform resource scheduling, considering the bandwidth delay between the server and the arriving CRs. To support PaaS services, the above algorithm will create the Virtual Machine (VM) in appropriate location depending on the request from the cloud user. The algorithm provides better throughput factor of 2 with least data transfer cost and response time of 330.84 sec. It can be used for many cloud service providers irrespective of the applications they serve.

Finally, in order to support the extended level of Infrastructure as a Service (IaaS) based cloud services and to increase the energy, speed and transfer rate, Modified Heuristic Search Algorithm (MHSA) is proposed for profit based resource allocation at the provider side. Profit in resource scheduling is achieved using selection of CRs and cluster server allocation. This gives probabilistic guarantee in preventing resource overloading and migration. MHSA has Cost Effective Reliable Resource Allocation (CERRA) scheme for the efficient cluster selection and the heuristic algorithm includes a fitness function for scheduling the CRs based on profit. The algorithm provides maximum speed of 231.42 GB/s, average transfer rate of 29.67 GB/s and average energy of 62.167 KWh.

The MHSA algorithm will be extended to support security features in future. Load balancing is also to be concentrated while dealing with enormous cloud users. These two factors will add benefit to the cloud.