International Journal of Enhanced Research in Management & Computer Applications,
ISSN: 2319-7471, Volume 15 Issue 5, May 2026

# Task Consolidation Algorithm for Heterogeneous Cloud Computing

Dr. Ravinder Kumar[1], Monika[2]

[1]Assistant Professor, Department of Computer Science Engineering, Rattan Institute of Technology and Management, Haryana, India
[2]Research Scholar, Department of Computer Science Engineering, Rattan Institute of Technology and Management, Haryana, India

**ABSTRACT**

**As the recent advancements are going in the field of computer technologies like network devices, hardware capacities and software applications, cloud computing has emerged as an important paradigm that provides scalable and dynamic virtual resources to the users through the internet. Energy consumed by modern computer systems, particularly by servers in a cloud has almost reached at an unacceptable level. Also the energy consumed due to underutilization of resources accounts almost 60% of the energy con-summed at peak load. It has resulted into reduced system reliability, extremely large electricity bills and environmental concerns because of resulting carbon emission. Therefore, there is a great need to optimize energy consumption. Methods like memory compression, request discrimination, task consolidation among virtual machines are developed to enhance resource utilization. Task consolidation problem has been addressed as an op-timization problem in heterogeneous cloud computing environment. Task consolidation maps user service requests to appropriate resources in cloud computing environment. The resource allocation problem in cloud computing environment is NP- complete. This thesis presents resource allocation problem as LPP to optimize energy consumed by the com-putting resources in cloud computing environment.**

## INTRODUCTION

As the recent advancements are going in the field of computer technologies like network devices, hardware capacities and software applications, Cloud computing has emerged as an important paradigm that provides scalable and dynamic virtual resources to the users through the internet. Cloud environment is a delivery model that delivered the on-demand computational resources to the applications running in data centers over internet according to pay-for-use basis. Some of the definitions given by well-known people and organization in this area include:

**1)** According to Buyya et al. (8), (7)- " *A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing re-sources based on service-level agreements established through negotiation between the service provider and consumers"*.

**2)** According to the various researchers in (12),(9),(17)- " *The Cloud is a model for enabling service users to have convenient, ubiquitous and on-demand network ac-cess to a shared pool of configurable computing resources (e.g., servers, networks, services and applications), that can be rapidly provisioned and released with mini-mal management effort or service-provider interaction*.

Resources in cloud are widely distributed and aims to provide reliable Qos while meeting SLA.
*SLA:* It is an agreement between the consumer and the service provider regarding the tech-nical performance promises that are made to the consumers. SLA includes the remedies for performance failures. The SLA is usually composed of three parts:

1) A collection of promises made to subscribers.
2) A collection of promises explicitly not made to subscribers, i.e., limitations.
3) A set of obligations that subscribers must accept.

Cloud Services may have different type of SLA. The main advantage of cloud environ-ment is that it reduces the hardware cost and users can access high quality of services at a low cost.

## LITERATURE REVIEW

Many researchers have investigated and demonstrated the impact of various task allocation and performance techniques in cloud data centers, but few have conducted systematic literature reviews in these areas. Furthermore, to the best of the author's knowledge, no study has combined task allocation and performance management techniques. Several landmarks for task allocation and performance management have been provided by authors such as Banga et al

Banga et al. [10] recommended a conventional cost-based scheduling approach that assigns suitably preferred resources, lowering the total cost of implementation and operation. Several tasks/cloudlets have also been divided and assigned appropriate resources to complete the task in the shortest possible time and at the lowest possible cost, based on their computational capability. Rodriguez et al.

[11] Conducted research on modern cloud computing technologies. They recommended classifying these systems according to different workload kinds, architectures, levels of complexity, and objectives. This research examined the 10 organizational business frameworks "Borg, Kubernetes, Swarm, Mesos, Marathon, Yarn, Omega, Apollo, and Fuxi". Additionally, Weerasiri et al.

Ramezani et al. [20] developed an optimization framework that used Particle Swarm Optimization (PSO) as a task scheduling model to migrate the additional jobs to the new host VMs. The outcomes demonstrated that the load-balancing method was quicker than conventional load-balancing algorithms. Reduced VM downtime and the possibility of users losing vital data were the main goals of the study. This led to lower memory use, lower migration and makespan, and increased data center efficiency, which enhanced the QoS that cloud users experienced. Although the approach emphasizes autonomous activities and homogeneous virtual machines, it has limited scalability.

Dubey et al. [31] stated that hybrid algorithms are a combination of heuristic and metaheuristic approaches. Heuristic methods are used for initial VM placement, and metaheuristic techniques are used to optimize VM placement during migration. According to Mohanty et al. [32], a metaheuristic technique may be used to develop a collection of solutions, and a heuristic approach is then used to choose the best option.

### Heterogeneous Cloud Computing Architecture

Cloud computing environment is designed to offer the on demand scalable services to the users over the internet through web browser or other devices.One of the vital features of cloud computing is to provide a desired level of QoS. QoS also called as Quality of Service can be defined using the term SLA that describes various characteristics like min-imizing response time or latency, maximal throughput, makespan minimization etc. by the deployed system. To meet the growing demand for large volumes of data, DC's host high performance storage devices and computing servers. These servers consume the major part of energy in data centers. As a result, CSP's have to deal with energy performance trade-off of minimizing energy consumption while meeting QoS requirements. Energy usage in large scale computer systems like cloud may yield many other serious issues like carbon emission and system reliability. The recent advancement of the term green or sus-tainable computing is not limited to the main computing components (processors, storage device etc.), but it can be extended to a much larger range of resources associated with computing facilities including auxiliary equipments like water used for cooling and even physical floor space used by these resources. This calls for the development of various software energy saving techniques including scheduling and virtualization. In response to poor utilization of resources in a DC, task consolidation is an effective technique to increase resource utilization. This technique is enabled by virtualization that facilitates the running of several tasks on a single physical resource concurrently.

### Current State of Work

The researchers have developed various task consolidation algorithms that vary greatly in different parameters. These parameters can be the approach used for developing eg., greedy or genetic or some other approach, in terms of the resources they have considered, objective functions, resources considered, simulator used and SLA parameter used for performance evaluation. I have surveyed some of the recently developed task consolidation algorithms. The comparison is shown in table 2.1.the blank field indicates that the required information is not discussed in that paper.

Table 2.1: Task Consolidation Approaches

| Approach | Resource Utilization | Energy Minimiza-tion | SLA | Resources | Simulator |
|---|---|---|---|---|---|
| ECTC (23) | No | Yes | No | CPU | |
| MaxUtil (23) | Yes | Yes | No | CPU | |
| EWRR (3) | Yes | Yes | Execution Deadline | CPU | Cloud Re-port |
| EAA (36) | Yes | Yes | No | CPU, Disk | |
| e-STAB (20) | No | Yes | Load Balancing | CPU, Net-work Band-width | Green Cloud |
| GBEAS (21) | No | Yes | Makespan | CPU | HyperSim-G |
| GBEAS (11) | No | Yes | Makespan | CPU | |
| PBSA (32) | Yes | Yes | Makespan | CPU, Memory | |
| MESFA (24) | No | Yes | Makespan, average waiting time | CPU | Matlab |

**Task Consolidation Problem in Heterogeneous Cloud Computing**

As the workload submitted by a user may vary greatly in terms of their complexity, re-source requirements and other parameters. Thus, the workload model should be flexible enough in defining the task requirements. The user may submit the workload in form of multiple independent jobs or time and precedence constrained jobs. Each job is further divided into a number of tasks that can be dependent on each other. Independent tasks can be executed concurrently leading to faster execution of the submitted job while as for precedence-constrained task, execution takes place on the basis of DAG. The tasks can be preemptive as well as non-preemptive. A single task undergoes various phases from the time of its arrival and until it gets completely executed.

**Scheduling Architecture**

For the above described host model, virtual machine model and task model, figure 3.2 describes the scheduling architecture. The architecture consists of a service scheduler and a VM controller. Service scheduler can be both central and distributed depending upon the requirements. In our work we have taken a centralized scheduler. The job of scheduler is to assign tasks to VM's. It also decided when VM's are to be added or removed to meet the demands. VM controller keep track of the availability of VM's and their available resources. It is also the in charge of migrating VM's across physical machines. When a task arrives, the scheduling process follows the following steps:

1. The scheduler checks the system status information about running task remaining execution time, active hosts, currently allocated VM's.
2. The tasks are sorted according to their arrival time.

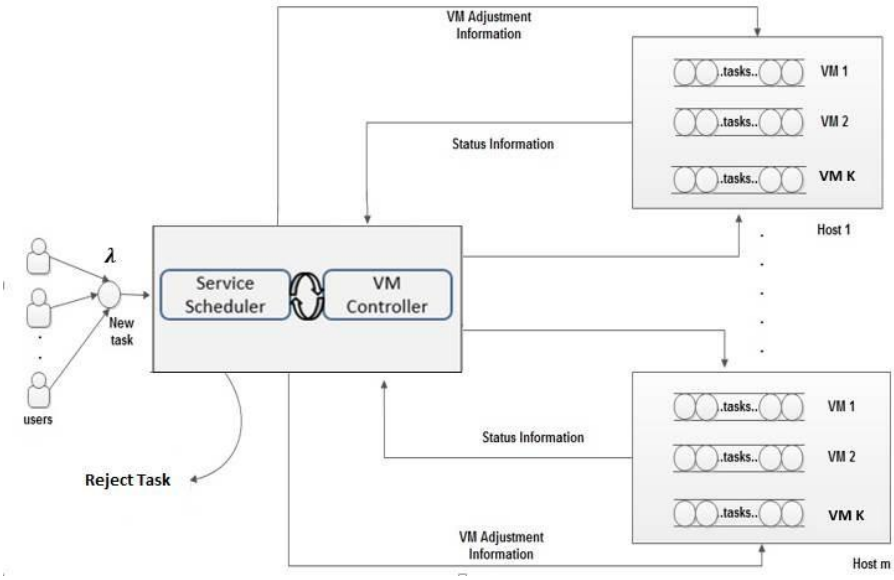


**Figure 3.2: Simulation Architecture**

1. The scheduler checks if the tasks can be allocated or not. If not, scheduler informs VM controller about it. To meet the task requirements, controller adds virtual ma-chines. If schedule is found, task is allocated else task is rejected.
2. Allocated VM's, active hosts,available resources, task completion time etc. are updated.

**Greedy Algorithms for Task Consolidation Problem**

The rapid growth of cloud computing environment with the virtualized DCs has made serious issues including energy consumption, cooling infrastructures and air condition-ing concerns in terms of increasing operational costs (34). The increasing size of cloud infrastructure along with poor resource utilization is coming as a great challenge. The energy consumption varies with the number of blade servers and the incoming workload and has emerged as one of the biggest challenges for cloud computing.

Among the major reasons of energy inefficiency, one is the idle power wastage. Even at very low utilization (10%) the energy consumed is 50-60% of the peak energy (23),(36),(31), (39),(1),(16). This has resulted into reduced system reliability, extremely large electric-ity bills and environmental issues generating due to emission of carbon in large quantity. Thus an energy efficient task consolidation strategy that maximizes the utilization of re-sources and in turn reduces the energy consumption is required. The task consolidation problem is a NP-Complete problem and requires the heuristics technique to solve. In a homogeneous cloud the problem is a bit easy to solve because of the similar resource capabilities and capacities of servers. But in heterogeneous cloud, the problem becomes more complex as all the servers vary in their processor capabilities and capacities.

## CONCLUSION

Task consolidation problem in cloud has been addressed as an optimization problem. Also, due to heterogeneous nature of physical servers and incoming tasks, this problem becomes more complex. In this thesis we have discussed about different task consoli-dation strategies proposed by various researchers. Most of the existing work is focused for homogeneous cloud environment and only a little work is done for addressing the task and machine heterogeneity. As it is a NP-complete problem, heuristics techniques are preferred by the researchers to address the problem. To model the heterogeneity, we have used the ETC model (2). We have also developed a generalized system and workload model to handle a variety of tasks. For the proposed model, we have used the greedy algo-rithms for task consolidation problem. The developed algorithm tries to make optimized use of cloud resources in order to reduce energy consumption. Simulation experiments were conducted to examine the performance of developed EATC algorithm to optimize the energy consumption in cloud computing system. The performance was compared against two other algorithms named Max Util (15) and a randomized algorithm. The re-sults showed a significant improvement in energy savings of EATC over the other two heuristics.

## REFERENCE

[1]. Abdulrahman Alahmadi, Abdulaziz Alnowiser, Michelle M Zhu, Dunren Che, and Parisa Ghodous. Enhanced first-fit decreasing algorithm for energy-aware job scheduling in cloud. In *Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*, volume 2, pages 69–74. IEEE, 2014.

[2]. Shoukat Ali, Howard Jay Siegel, Muthucumaru Maheswaran, and Debra Hensgen. Task execution time modeling for heterogeneous computing systems. In *Heteroge-neous Computing Workshop, 2000.(HCW 2000) Proceedings. 9th*, pages 185–199. IEEE, 2000.

[3]. Abdulaziz Alnowiser, Eman Aldhahri, Abdulrahman Alahmadi, and Michelle M Zhu. Enhanced weighted round robin (ewrr) with dvfs technology in cloud energy-aware. In *Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*, volume 1, pages 320–326. IEEE, 2014.

[4]. Anton Beloglazov and Rajkumar Buyya. Energy efficient resource management in virtualized cloud data centers. In *Proceedings of the 2010 10th IEEE/ACM Inter-national Conference on Cluster, Cloud and Grid Computing*, pages 826–831. IEEE Computer Society, 2010.

[5]. Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya. A taxonomy and survey of energy-efficient data centers and cloud computing systems. *arXiv preprint arXiv:1007.0066*, 2010.

[6]. Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, Albert Zomaya, et al. A taxonomy and survey of energy-efficient data centers and cloud computing systems. *Advances in Computers*, 82(2):47–111, 2011.

[7]. Rajkumar Buyya, Anton Beloglazov, and Jemal Abawajy. Energy-efficient manage-

[8]. ment of data center resources for cloud computing: A vision, architectural elements, and open challenges. *arXiv preprint arXiv:1006.0308*, 2010.