A Review on Different Task Scheduling Algorithm in Cloud Computing

[1] *Jugal Trivedi, Department of Computer Science and Engineering, Parul Institute of Technology, Vadodara, India*
[2] *Asst. Prof. Upendra R. Bhoi, Department of Computer Science and Engineering, Parul Institute of Technology, Vadodara, India*

ABSTRACT

Cloud computing is an internet-based Technology in which application, data, and Information services are provided over the internet. Cloud Computing is useful to share data and offer services transparently among its various users. In cloud computing, many tasks need to execute at a time by the available resources in order to achieve better performance, minimum completion time, shortest response time, resource utilization etc. Because of these different factors, we need to design, develop, and propose a scheduling algorithm for the proper allocation of tasks to the resources. Here in this paper, a review of various task scheduling techniques is presented. This study eventual that all the existing techniques focus on reducing makespan, service response time and improving performance etc. There are many parameters can be observed of scheduling problem such as makespan, virtual machine allocation, load balancing, throughput, reliability, resource cost, service utilization.

Keyword: - *Cloud computing, recent technology in computing, On-demand services through internet, issues of cloud computing, Task Scheduling.*

1. INTRODUCTION

Cloud computing has been established in recent years as an important area of research. Cloud computing is nothing but it is a collection/group of integrated and networked hardware, software and Internet infrastructure (called a platform).

In addition, the platform provides on-demand services that are always on anywhere, anytime and anyplace. Cloud computing technology virtualizes and offers many services across the network. It mainly aims at scalability, availability, throughput, and resource utilization.

NIST provides a somewhat more objective and specific definition here. [1]

"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of

five essential characteristics, three service models, and four deployment models."

Cloud computing is the combination of the several concepts from resource pooling, virtualization, dynamic provisioning, utility computing, and on-demand deployment, Internet delivery of services, to enable a more flexible approach to deploy and scaling applications. There are mainly three services provided by cloud computing, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS). [1]

IaaS: Infrastructure as a Service is the virtual delivery of computing resources in the form of hardware, networking, and storage services. It may also include the delivery of operating systems and virtualization technology to manage the resources.Ex. Amazon web services. [10]

PaaS: Platform as a Service (PaaS) as the name suggests, provides you computing platforms which typically includes an operating system, programming language execution environment, database, web server etc. Examples: AWS

Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos. [10]

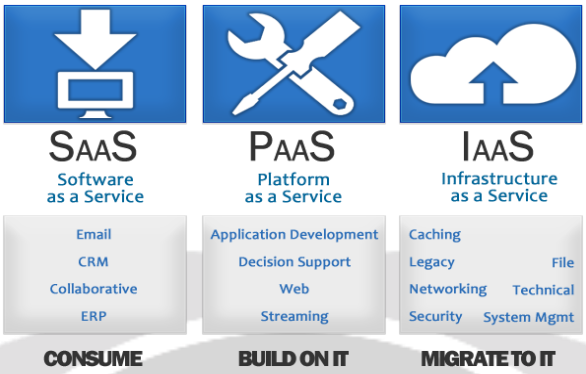

Fig 1: Cloud Computing Services [1]

SaaS: Represent the largest cloud market and are still growing quickly. It uses the web to deliver applications that are managed by a third-party vendor and whose interface is accessed on the clients' side. Most SaaS applications can be run directly from a web browser without any downloads or installations required, although some require plugins.Ex. Google Apps. [10]

There are four deployment models are discovered, namely as follows:

Public cloud: It is freely available for access based on the standard cloud computing. The service provider makes resources, such as applications and storage, available to the general public over the internet. [2]

Private cloud: It allows the usage of services by a single client on a private network. The importance benefits of this model are data security, corporate governance, and reliability concerns. The private cloud is used by the organization when it has a huge, well-run data center having a lot of spare capacity. [2]

Community cloud: It provides a number of benefits, such as privacy and security. This model, which is quite expensive, is used when the organizations having common goals and requirements are ready to share the benefits of the cloud service. [2]

Hybrid cloud: A combination of public cloud storage and private cloud storage where some critical data resides in the enterprise's private cloud while another data is stored and accessible from a public cloud storage provider. [2]

2. TASK SCHEDULING

In Cloud, So many tasks have to be executed by the available resources to achieve Minimal total time for completion, shortest response time and Effective utilization of resources. Because of these distinct intentions, we need to design, develop and propose a scheduling algorithm that is used by the task scheduler to outperform appropriate allocation map of tasks on resources.

Distinct types of scheduling are based on different criteria, such as Static versus Dynamic, Centralized versus Distributed are described below:

a) Static Scheduling

Pre-Schedule jobs, all details about available resources and jobs/tasks in the application must be known and moreover, a task is assigned once to a resource, so that it's easier to adapt based on scheduler's perspective. [3]

b) Dynamic Scheduling

It is more flexible compare to static scheduling where jobs are dynamically available for scheduling over the time by the scheduler with no issues, to be able of determining run time in advance. It is highly critical to include load balance as the main factor to obtain stable and efficient scheduler algorithm. [3]

e) Immediate / Online Mode

In which scheduler schedules any recently arriving job as soon as it arrives with no waiting for next time interval on available resources at that moment. [4]

f) Batch / Offline Mode

The scheduler holds arriving jobs as group of problems to be solved over successive time intervals so that it is better to map a job for suitable resources depending on its characteristics. [4]

3. VARIOUS TASK SCHEDULING TECHNIQUES IN CLOUD COMPUTING

Following task scheduling techniques are currently prevalent in clouds,

A study on strategic provisioning of cloud computing services [5]

In this paper Md whaiduzzaman et. Illustrate cloud computing is currently emerging as an ever-changing, growing Paradigm that models everything-as-a-service. Virtualized physical resources, infrastructure, and applications are supplied by service provisioning in the cloud. The evolution in the adoption of cloud computing is driven by clear and distinct promising features for both cloud users and cloud providers. However, the increasing number of cloud providers and the variety of service offerings has made it difficult for the customers to choose the best services. By employing successful service provisioning, the essential services required by customers, such as agility and availability, pricing, security and trust, and user metrics can be guaranteed by service provisioning. Hence, continuous service provisioning that satisfies the user requirements is a mandatory feature for the cloud user and vitally important in cloud computing service offerings. Therefore, we aim to review the state-of-the-art service provisioning objectives, essential services, topologies, user requirements, necessary metrics, and pricing mechanisms. We synthesize and summarize different provision techniques, approaches, and models through a comprehensive literature review. A thematic taxonomy of cloud service provisioning is presented after the systematic review. Finally, future research directions and open research issues are identified.

Load Balanced Static Grid Scheduling Using Max-Min Heuristic [6]

In this paper Tarun Kumar Ghosh, Rajmohan Goswami, Sumit Bera and Subhabrata Barman has proposed algorithm is executed in two-phases. LBMM executes Max-Min in the first round. In the second round it chooses the resources with heavy load and reassigns them to the resources with light load. LBMM identifies the resources with heavy load by choosing the resource with high makespan in the schedule produced by Max-Min. It then considers the tasks assigned in that resource and chooses the task with minimum execution time on that resource.

Load Balanced Min-Min Algorithm for Static Meta-Task Scheduling in Grid Computing[7]

In This Research paper T. Kokilavani1, Dr. D.I. George Amalarethinam proposed algorithm is executed in two-phases. LBMM executes Min-Min in the first round. In the second round it chooses the resources with heavy load and reassigns them to the resources with light load. LBMM identifies the resources with heavy load by choosing the resource with high makespan in the schedule produced by Min-Min. It then considers the tasks assigned in that resource and chooses the task with minimum execution time on that resource.

A Three-Phases Scheduling in a Hierarchical In Cloud Computing Network [8]

In this paper Shu-Ching Wang, Kuo-Qin Yan has order to reach the load balance and reduce the execution time of each node in the cloud computing environment, this proposes a three-phases scheduling.
(1) BTO (Best Task Order) scheduling determines the execution order for each task request.
(2) EOLB (Enhanced Opportunistic Load Balancing) scheduling assigns a suitable service manager for allocation of the service node.
(3) EMM (Enhanced Min-Min) scheduling guarantees that a suitable service node will be assigned to execute the task in the minimum execution time. each task will be completed effectively and quickly in the hierarchical cloud computing network.

Improved Max-Min Heuristic Model for Task Scheduling In Cloud [9]

In this paper S.Devipriya and C.Ramesh has Improved Max-Min algorithm calculates the expected completion time of the submitted tasks on each resource. Then the task with the overall maximum expected execution time (Largest Task) is assigned to a resource that has the minimum overall completion time (Slowest Resource). Finally, this scheduled task is removed from meta-tasks and all calculated times are updated and then applying max-min

algorithm on remaining tasks. That means Select task with the overall maximum expected execution time (Largest Task) then assign to be executed by resource with minimum expected completion time (Slowest Resource). It will be achieved shortest makespan of submitted tasks on available resources beside concurrency.

4. CONCLUSION

Task scheduling in cloud computing is highly challenging in cloud computing. To meet the needs of thousands of requests by making best possible use of cloud resources is a challenge for task manager. Traditional methods of scheduling lead to high response time and low throughput. In this paper, I have surveyed the various existing task scheduling algorithms. The efficiency of the cloud computing system depends on the algorithm used during the scheduling task. During scheduling they had considered various techniques and applied constraints but as the cloud computing is too vast that they had not been able to capture all aspects at the same time but they mentioned these facts that there is a chance of modification of algorithms and which part has to be modified.

REFERENCES

[1] MellPeter, GranceTimothy , "The NIST Definition of Cloud Computing", Computer Security Division Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, MD 20899-8930 ,NIST Special Publication 800-145, September 2011.

[2] Huaglory Tianfield, "Cloud Computing Architectures" School of Engineering and Built Environment, Glasgow Caledonian University Cowcaddens Road, Glasgow G4 0BA, United Kingdom, ©2011 IEEE.

[3]Vijayalakshmi A. Lepakshi, Dr. Prashanth C S R, "A Study on Task Scheduling Algorithms in Cloud Computing" International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 11, May 2013.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] Mythry Vuyyuru, Pulipati Annapurna, K.Ganapathi Babu, A.S.K Ratnam, "An Overview of Cloud Computing Technology", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.

[5] Md. Whaiduzzaman, Mohammad nazmul haque, Md rejaul karim chowdhury, and Abdullah gani., ―A Study On Strategic Provisioning Of Cloud Computing Services‖, Hindawi Publishing Corporation, E Scientific World Journal Volume 2014.

[6] Luiz F. Bittencourt, , Edmundo R. M. Madeira, and Nelson L. S. Da Fonseca, ―Scheduling In Hybrid Clouds‖, Institute of Computing – University of Campinas – UNICAMP, Campinas – S˜ao Paulo – Brazil, 13083-852.

[7]Tarun Kumar Ghosh, Rajmohan Goswami, Sumit Bera and Subhabrata Barman, "Load Balanced Static Grid Scheduling Using Max-Min Heuristic", 2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing.

[8] Shu-Ching Wang, Kuo-Qin Yan, Shun-Sheng Wang, Ching-Wei Chen, "A Three-Phases Scheduling in a Hierarchical Cloud Computing Network", IEEE, 2011.

[9] S.Devipriya and C.Ramesh"Improved Max-Min Heuristic Model for Task Scheduling in Cloud" (IEEE) 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE).

[10] Definition "IaaS, PaaS, SaaS" http://searchcloudcomputing.techtarget.com/defin