



Customized Task Scheduling Algorithm for Scientific Application using Aneka

Prof. S.Y. Inamdar¹, Pratima N. Ghalsasi²

Assistant Professor, Computer Science and Engineering, DACOE, Karad, Maharashtra, India¹

Student, Computer Science and Engineering, DACOE, Karad, Maharashtra, India²

Abstract: Cloud computing is emanate technology in IT domain. The scheduling of the cloud services to the consumers by service providers significance the cost benefit of these computing paradigms. Cloud computing is a recent technology that are used for online distribution of computing resources also services on pay- per- use basis. Task scheduling as well as provision of resources are main problem areas in cloud computing. There are many algorithms like Min-Min, Max-Min, Suffrage, Shortest Cloudlet to Fastest Processor (SCFP), Longest Cloudlet to Fastest Processor (LCFP), Ant-Colony Optimization. Priority Based, First Come First Serve (FCFS) etc. These dynamically scalable resources within a cloud are handled by cloud service provider and distributed among the number of users according to the contract known as Service Level Agreement (SLA). After recognition of benefits of cloud computing, a large number of users using cloud services are increasing tremendously. Therefore task scheduling plays important role in allocating and scheduling the cloud resources among the users efficiently. A capable task scheduling policy provides proper resource utilization, load balancing and optimization of execution time and cost. In this paper, we have given an overview of research work done by several researchers in the area of cloud task scheduling.

Keywords: Cloud computing, job scheduling, scheduling algorithm, cloud services, Service Level Agreement, task scheduling, resource utilization.

INTRODUCTION

Big data is most important technology that has the potential for frequently changing the way organizations use information to enhance the customer experience and transform their business models. Big data is complex market. It is a combination of data-management technologies that have evolved over time. Big data enables organizations to storing, handle, and manipulate vast amounts of data at the right speed and at the right time to gain the right insights. The term "cloud" originates from the world of telecommunications that providers launch using virtual private network (VPN) services for data communications.

Cloud computing is a type of Internet-based computing that provides shared computer resources and data to computers and other devices on demand. It is a model for enabling pervasive, on-demand access to a shared pool of configurable computing resources examples are computer networks, servers, storage, applications and services which can be rapidly provisioned and released with minimal management effort.

It deals with computation, software, and data access also storage services that may not require end-user knowledge of the physical location & the configuration of the system that is delivering the services. Cloud computing is a modern trend in IT that turn computing and data away from desktop and portable PCs into large data centres.[3]

There are following four cloud deployment models which are used to show the ways through which cloud services are used by end users.

Private Cloud: Private cloud is build for the exclusionary use by single organization. The private cloud provides the resources accesses as well as used only by users organization who own that cloud. Main advantage of private cloud is its security since its resources are used within different users of same organization. Another benefit of private cloud is its capacity to provide customization which allows organization to design it according to specification. Private cloud is providing the less scalability.

Community Cloud: Community Cloud enable for sharing its resources among the users of multiple organizations which are having similar specification and objectives. This cloud divides initial formulation cost among several organizations. These clouds provide somewhat more scalability of resources than private cloud.

Community cloud in computing is a collective efforts in which infrastructure shared among the number of organization from particular community with common concerns that is security compliance etc.it is internally managed by third party and hosted internally or externally.

Public Cloud: Public cloud provides huge amount of storage, services and computing environment to the all users over world through web on pay-per-use basis. Public clouds are built and managed by third party agencies. It provides more scalability, availability and flexibility than private clouds. But less security is a major problem with



the public cloud since the resources provided by public cloud is shared between large numbers of worldwide users from different organizations.

Hybrid Cloud: Hybrid clouds are built by linking the private and public clouds. Hybrid cloud thus whole the properties of both private and public clouds such as scalability, flexibility also security. In hybrid cloud, users of private clouds access the resources of public cloud when its own resources become inadequate. The further required resources are taken from public cloud on pay-per-use basis.[3]

Services

There are usually three services of cloud

1. IaaS
2. PaaS
3. SaaS

Infrastructure as a Service(IaaS) is a form of cloud computing that provides virtualized computing resources over internet. Infrastructure as a Service is one of the 3 main categories of cloud services.

Platform as a Service (PaaS) is a category of cloud computing services that provides platform and allowing customers to develop, run as well as manage the applications without complexity of building and maintaining the infrastructure.

Software as a Service (SaaS) is a cloud service. It is a software distribution model in which third party provider host applications and makes them available to customer o Infrastructure as a Service over internet.

What is Aneka?

Aneka is a market oriented Cloud development & management platform with fast application development and workload distribution capabilities. It is an combined package which allows you to mercilessly build and manage an interconnected network in addition to forward development, management and deployment of distributed applications using Microsoft .NET frameworks on these networks.Aneka provides to developers rich set of APIs for transparently employ such resources as well as expressing the business logic of applications by using the preferred programming abstractions.[2]

Task

The Task Model defines an application as a collection of tasks. Tasks are different work units that can be executed by the Schedule. Within the Task Model a taskcomprises all the components required for its execution on a grid node.

The Task Model is the correct solution to use when the distributed application consists of a collection of independent jobs that are executed on a grid and whose results are collected and composed together by the end user. In this concept the user creates a set of tasks, submits them to Aneka platform, and waits for the results.[1]

Scheduling

There have been different types of scheduling algorithm applied to the cloud environment with some modification. The main goal of job scheduling algorithm is to increase the performance also quality of service and at the same time preserving the efficiency and fairness among the jobs and reduce the cost of execution. Traditional job scheduling algorithms are enabling to provide scheduling in the cloud environments. Based on a simple classification, job scheduling algorithms in cloud computing can be classified into two main types. Batch mode heuristic scheduling algorithms (BMHA) and online mode heuristic algorithms. In first algorithm, Jobs are taken into a collection when they come in the system. The scheduling algorithm starts its work after a fixed amount of time. The Round Robin algorithm (RR) is the example of BMHA based algorithms.

In On-line mode heuristic scheduling algorithm, Jobs are scheduled individually as soon as they arriving in the system. These scheduling algorithms are more suitable for a cloud environment Because of the cloud environment's divergence and varying speed of processor.

Architecture:

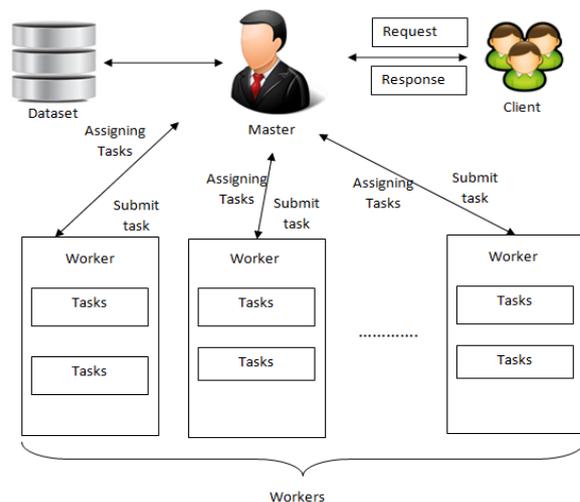


Fig. Architecture of Task Scheduling Algorithm

Following figure shows the basic architecture of Task Scheduling Algorithm. The system includes key components, including Master Node, Worker Node, Dataset and User(Client).

The Master is responsible to perform Scheduling, Storage Device, Accounting and Reporting .Master control and manages the workers and its task. Worker is responsible for executing the task and submitted the result to the master container.

Example of this scheduling algorithm is most fit task scheduling algorithm.

A. First Come First Serve Algorithm(FCFS)

Jobs are prepared in queue as they come. FCFS algorithm is simple and fast. [1]



B. Round Robin algorithm (RR)

In this algorithm, processes are given a bounded amount of CPU time called a time-slice in FIFO manner. If a process does not complete execution before its CPU-time run out, the CPU is pre-empted & given to the next process waiting in a queue. And the appropriate process is placed at the end of the ready queue.

C. Random Algorithm

In random algorithm, the selected jobs are however selected for execution and assigned to Virtual Machine. The algorithm doesn't take into considerations the status of the Virtual machine, which will either be under excessive or low load.

D. Max – Min algorithm

This algorithm selected larger tasks to be executed firstly, which manufacture small task delays for long time.

E. Priority scheduling algorithm

Every process is assigned a priority, then based on priority processes are grant to be executed. Equal-Priority processes are scheduled in FCFS manner.

F. Most fit task scheduling algorithm

Job which fit best in queue based on few parameters is executed first.[1]

Customized Task Scheduling Algorithm

Step 1: START

Step 2: Submit job to master using scientific application.

Step 3: Find out count of available local resources and workers available in private cloud.

Step 4: Calculate time required to run the application using

$$\text{Total Time required} = \frac{\text{Number of task} * \text{Average Time}}{\text{Available resource}}$$

Step 5: Calculate Resource required using

$$\text{Resources required} = \frac{\text{Total tasks} * \text{Average Time}}{\text{Time remaining}}$$

Step 6: if Total Time required > Time Available

Go to step 8.

Step 7: Depending upon the application, customized scheduling algorithm decides which scheduling sub algorithm is choose so that application execution time is reduced.

Step 8: If Extra Resources are required then customized scheduling algorithm choose resources from other pool.

Goto Step 10.

Step 9: Select number of Resources required from Local pool.

Step10: Release Selected Resources after tasks completion.

Step 11: STOP

CONCLUSION

Task scheduling plays important role in managing also sharing cloud resources between the different cloud users. Therefore today task scheduling is important research topic in the area of cloud computing. In this paper, we

have highlighted on the various scenario proposed by the researchers regarding to task scheduling in cloud computing. In this paper various scheduling algorithms, namely FCFS Scheduling, Job Scheduling Model based on Multi-Objective Genetic Algorithm, Priority based Job Scheduling Algorithm, SLA-Tree and Enhanced Max-min Task Scheduling Algorithm have been studied and analysed. Scheduling algorithm which can solve problem of various types of quality services. Because selection of job scheduling algorithms depends on personal reference and on problem need to be solved.

ACKNOWLEDGMENT

It is our privilege to acknowledge with deep sense of gratitude to our project guide **Prof. Inamdar S.Y.** and our H.O.D **Prof. Patil Ashish N.** whose supervision, inspiration and valuable discussion has helped us tremendously to complete our project. Their guidance proved to be the most valuable to overcome all the hurdles in the fulfilment of this mega project on "Customized task scheduling algorithm using ANEKA".

We are grateful to Principal **Dr. Jalindar R. Patil** for direct or indirect help in the completion of this project. Last but not least, this acknowledgement would be incomplete without rendering my sincere gratitude to all those who have helped us in the completion of this project.

REFERENCES

- [1] Rohit O. Gupta, TusharChampaneria "A Survey of Proposed Job Scheduling Algorithms in Cloud Computing Environment", International Journal of Advanced Research in Computer Science and Software Engineering.
- [2] Dr. RajkumarBuyyaKarthikSukumar 'Manjrasoft Market Oriented Cloud Computing Platform 'Manjrasoft Pty Ltd.
- [3]. Mahesh S. Shinde, AnilkumarKadam "Survey of Recent Task Scheduling Strategies in Cloud Computing", International Journal of Advanced Research in Computer Science and Software Engineering.
- [4] Mohammad Hamdaqa and LadanTahvildari , "Cloud Computing Uncovered: A Research Landscape". Elsevier Press. pp. 41–85. ISBN 0-12-396535-7.
- [5] IsamAzawiMohialdeen, "Comparative Study Of Scheduling. Algorithms In Cloud Computing Environment", Journal of Computer Science ISSN 1549-3636
- [6] The NIST definition of cloud computing, NIST special publication 800-145
- [7] Poonam Devi, "Implementation of Cloud Computing By Using Short Job scheduling", International Journal of Advanced Research in Computer Science and Software Engineering, July – 2013, ISSN: 2277-128X.
- [8] AV.Karthick, Dr.E.Ramaraj and R.Kannan, 'An Efficient Tri Queue Job Scheduling using Dynamic Quantum Time for Cloud Environment', International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), IEEE 978-1-4673-6126-2/13, 2013, PP: 871-876.
- [9] ShivaniDubey, Vismay Jain and ShailendraShrivastava, 'An Innovative Approach for Scheduling of Tasks in Cloud Environment', 4th ICCCNT, IEEE-31661, July 4-6, 2013.
- [10] Zhi Yang , Changqin Yin and Yan Liu, 'A Cost-based Resource Scheduling Paradigm in Cloud Computing', 12th International Conference on Parallel and Distributed Computing, Applications and Technologies, IEEE 978-0-7695-4564-6/11, 2011, PP: 417-422.



BIOGRAPHIES



Mr. S. Y. Inamdar Assistant professor of
Computer Science & Engineering,
DACOE, Karad.



Ms. Pratima N. Ghalsasi. Student of BE
Computer Science & Engineering,
DACOE, Karad.