Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 38 (2012) 2881 – 2888

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference On Modeling Optimization And Computing – (ICMOC-2012)

Survey on Scheduling Issues in Cloud Computing

Vijindra[a], Sudhir Shenai[b], a*

[a]PG Scholar, Department of IT, Karunya University, Coimbatore, India

[b]Assistant Professor, Department of IT, Karunya university, Coimbatore, India

Abstract

Cloud computing has captured the attention of today's CIOs, offering huge potential for more flexible, readily-scalable and cost-effective IT operations. It represents a different way to architect and remotely manage computing resources. Cloud computing deals with different kinds of virtualized resources, hence scheduling places an important role in cloud computing. In cloud, user may use hundreds of thousands virtualized resources for each user task. Hence manual scheduling is not a feasible solution. Focusing scheduling to a cloud environment enables the use of various cloud services to help framework implementation. Thus the comprehensive way of different type of scheduling algorithms in cloud computing environment surveyed which includes the workflow scheduling as well as grid scheduling. This study gives an elaborate idea about grid, cloud, workflow scheduling.

© 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Noorul Islam Centre for Higher Education Open access under CC BY-NC-ND license.

Key words: cloud computing; grid; scheduling; workflows.

1. Introduction

Everyone has an opinion on what is cloud computing. Cloud computing promises to increase the velocity with which applications are deployed, increase innovation, and lower costs, all while increasing business agility. Cloud computing can be the ability to use applications on the Internet that store and protect data while providing a service - anything including email, sales force automation and tax preparation [1]. Virtualization is a key feature of cloud computing. IT organizations have understood for years that virtualization allows them to quickly and easily create copies of existing environments - sometimes involving multiple virtual machines - to support test, development, and staging activities [1].

* Corresponding author. Tel.: +91-9488841133; fax: 0422-2614534.

E-mail address: vijindrarajendran@gmail.com.

1877-7058 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.06.337

Cloud computing has various definitions as per different industries, There are over 20 different definitions from a variety of sources. In this paper, we agree to the definition of cloud computing provided by The National Institute of Standards and Technology (NIST), as it covers, in our Opinion, all the essential aspects of cloud computing: The Cloud computing is a model for enabling 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[2]. Cloud computing has a variety of characteristics such as Commercialization, Virtualization, Shared Infrastructure, Dynamic Provisioning, Network Access, Managed Metering, Self-service based usage mode, self-managed platform, Consumption-based billing, Resource pooling, Rapid elasticity, Multi Tenacity[3]. Cloud computing provides everything as a service, the three basic service layers of cloud computing is Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) [4]. Since cloud computing is in infancy state, different kinds of researches are going on, the research areas which includes cloud systems development and management, resource allocation and scheduling, security issues, cloud storage, elastic scalability, programming models, and so on[5]. In cloud computing, user may face hundreds of thousands of virtualized resources to utilize, it is impossible for anyone to allocate the jobs manually. Due to the commercialization and virtualization properties, cloud computing leaves job scheduling complexity to the virtual machine layer through resource virtualization. Hence to allocate the resources to each job efficiently, scheduling plays more important role in cloud computing [6].

The objective of this paper is to be focus on various scheduling algorithms in cloud, grid and workflows. The rest of the paper is organized as follows. Section II presents scheduling algorithms in cloud, grid, and various workflows. Section III presents scheduling parameters of existing scheduling algorithms with table and section IV concludes the paper with a summary of our contributions.

2. Existing Scheduling Algorithms

The following scheduling algorithms are presently established in the area of grids, clouds and various workflows and these algorithms have been summarized in table1 with the scheduling parameters.

2.1 An energy efficient scheduling application based on private clouds

Hybrid energy efficient scheduling application [7] based on pre-power techniques and least load first algorithm developed. Since private clouds have some unique characteristics and special requirements, it is still a challenging problem to effectively schedule virtual machine requests onto compute nodes, especially with multiple objectives to meet. Specifically two problems of virtual machine scheduling are discussed. Pre-power technique is used to reduce the response time and it uses idle threshold value. Least load first algorithm is used to balance workloads when the data centres are running on low power mode.

2.2 A scheduling algorithm for private cloud

A hybrid energy-efficient scheduling algorithm [8] using dynamic migration was proposed. This paper is based on [7], but by using the threshold value, powering down a busy node is not feasible. Hence, an expected spectrum set for the left capacity is used. It uses power up command to wake the sleep nodes as well as the idle nodes. Hence power efficiency is improved.

2.3 Energy-Efficient Scheduling of HPC Applications in Cloud Computing Environments

proposed a near-optimal scheduling policies[9] that exploits heterogeneity across multiple data centers for a cloud provider. A number of energy efficiency factors such as energy cost, carbon emission rate, workload, and CPU power efficiency which changes across different data center depending on their location, architectural design, and management system were considered.

2.4 Power-aware provisioning of Cloud Resources for Real-time Services

This paper explore power-aware provisioning of virtual machines for real-time services[10]. Energy consumption in a data centre is a critical issue in cloud computing. Three power-aware VM provisioning schemes proposed: Lowest-DVS, $Advanced-DVS, and Adaptive-DVS. A real-time Cloud service framework where each real-time service request is modelled as RT-VM in resource brokers have developed. This Proposed approach is

- to model a real-time service as a real-time virtual machine request; and
- To provision virtual machines of datacenters using DVFS (Dynamic Voltage Frequency Scaling) schemes.

The proposed algorithm shows better reduced power consumption with high performance.

2.5 An ANT colony algorithm for balanced job scheduling in Grids

here proposed a balanced Ant colony algorithm [11] which uses pseudo random proportional rule to balance the entire system load while completing all the jobs at hand as soon as possible according to the environment status. Current scientific problems are very complex and need huge computing power and storage space. To utilize the grid resources efficiently, balanced ant colony algorithm is proposed by balancing the workload as well as minimizing the makespan.

2.6 Job Scheduling Algorithm based on Dynamic Management of Resources Provided by Grid Computing Systems

An algorithm of job scheduling and dynamic adjustment of nodes loading within a grid system proposed[12] . Within a distributed computing system, requests of processing are randomly received from the system's users. A good planning of these requests assumes their assigning towards available processors, so that all requests have to be solved as soon as possible. Considering the resources sharing in grid systems, a job scheduling algorithm are proposed with dynamic load balancing is proposed. The distribution of first come, first served (FCFS) with a round robin mechanism of the execution nodes is proposed.

2.7 Scheduling of scientific workflows using a Chaos genetic algorithm

A Meta heuristic algorithm [13] based on Genetic Algorithms proposed. In a grid environment, numbers of challenges are available like

- sources are shared(competition)
- scheduler is not in control of resources
- Numbers of available resources are constantly changing and so on.

By using the characteristic of chaotic variable in scattering the solutions among the whole search space and thus avoids the precipitate convergence of the solutions and produces better results within a shorter time. Investigation of scheduling workflows considering the QoS constraints (user budget, deadline) has done.

2.8 Evolution of Gang scheduling performance and cost in a cloud computing system

An efficient job scheduling algorithm [14] for time sharing proposed. This paper is to study the performance of a distributed Cloud Computing model, based on the Amazon Elastic Compute Cloud (EC2) architecture and to revise, study, and estimate both the performance and the overall cost of two foremost gang scheduling algorithms. It utilizes the concept of virtual machines which acts as the computational units of the system. The proposed system implemented for adding and removing virtual

machines from the system depending on the systems load at any specific time. Job routing, job scheduling has done, adaptive FCFS and largest job first served are used.

2.9 Heterogeneity aware resource selection & scheduling in the cloud

A metric of share in a heterogeneous cluster to realize a scheduling scheme that achieves high performance and fairness proposed [15]. The heterogeneity of the environment should be developed along with performance and cost-effectiveness. The data analytics system must report for heterogeneity of the situation and workloads. It also needs to provide fairness among jobs when multiple jobs share the cluster. Hence architecture to allocate resources to a data analytics cluster in the cloud proposed.

2.10 Adopting market oriented scheduling polices for cloud computing

By considering the 2 levels of resource provisioning (time and cost) two methodologies were proposed [16]. It uses cheapest resource type which is known as small computational unit. This paper deals with how scheduling policies inside the broker can benefit from resources supplied by the IaaS providers in addition to the local schedulers to get the use of application completed by the requested deadline and provided budget.

2.11 A particle swarm optimization based heuristic for scheduling workflow applications in cloud computing systems

Meta heuristics method based on particle swarm optimization proposed [17]. In grid environment, user applications may incur large amount of data retrieval and execution costs when they are scheduled taking into account only the execution time. Added to that optimizing the execution time, the cost arising from data transfer between resources as well as execution costs must also be taken into account, and focusing to minimize the total execution cost of applications on resources. PSO's ability to find near optimal solutions for mapping all the tasks in the workflow to the given set of computer resources. It takes both computation and communication cost into account, if the resource cost increases PSO minimizes the maximum total cost of assigning all tasks to resources.

2.12 Cloud-DLS: Dynamic trusted scheduling for Cloud computing

A trust dynamic level scheduling algorithm [18] named Cloud-DLS proposed. Because of the characteristics of cloud computing, obtaining trustworthiness in computing resources is difficult. Novel Bayesian method based cognitive trust model, trust relationship models of sociology used. This paper focuses on trustworthiness in cloud computing. Cognitive trust model is used. Two kinds of trust that is direct trust degree; recommendation trust degree is obtained to obtain the trusted scheduling, and extends the traditional DLS algorithm by considering trustworthiness of resource nodes. This algorithm meets the requirement of user tasks in trust, and makes tasks scheduling based on directed acyclic graph (DAG) more reasonable.

2.13 A Framework for Resource Allocation Strategies in Cloud Computing Environment

Framework for resource allocation problem [19] based on online tailored active measurements was developed. Focusing towards network awareness and consistent optimization of resource allocation strategies and identifies the issues which need further investigation by the research community. There is a need for developing new methods of reliable active measurements aimed at capturing global Internet behavior online, and for enabling prediction of the most critical performance parameters. In this paper, the resource allocation problems are discussed that are based on Artificial Intelligence based, Theory of random graphs, Peer to peer based approaches and resource allocation framework is proposed in cloud computing, and considers computing resources, storage resources for allocation.

2.14 A Survey of Various Workflow Scheduling Algorithms in Cloud Environment

A survey of various workflow scheduling algorithms [20] has done. Previous workflow scheduling algorithms does not consider reliability and availability factors. Thus needed a workflow scheduling algorithm that can improve availability and reliability in cloud environment. The main purpose of a workflow management system (WfMS) is to support the definition, execution, registration and control of business processes. Three major components in a workflow enactment engine are the workflow scheduling, data movement and fault management. Due to the reduced performance faced in grids, there is a need to implement workflows in cloud.

2.15 A Compromised-Time-Cost Scheduling Algorithm in SwinDeW-C for Instance-Intensive Cost-Constrained Workflows on Cloud Computing Platform

Novel compromised-time-cost scheduling algorithm [22] which considers the execution time and cost along with the user input proposed. This work focuses on minimizing the cost under user defined deadlines and it provides just in time graph of time cost relationship during workflow execution, so that if the user want to change the schedule, it is possible by using multiple concurrent instances on the dynamic cloud computing platform.

3. Our Contributions

Based on the survey a scheduling framework can be implemented by using the different parameters. Good scheduling framework should contain the following characteristics. It must focus on

- Energy efficiency and Load balancing of the data centers
- QoS parameters determined by the user which includes execution time, cost, and so on
- It should satisfy the security features.
- Fairness resource allocation places a vital role in scheduling

Considering all the parameters in a single scheduling framework is not a feasible solution, hence it increases the complexity of the design, depending on the nature and size of job, environment, resource availability, the scheduling framework can be implemented.

Table 1.Scheduling Parameters considered by existing scheduling algorithms

Paper	Execution Time	Response Time	Cost	Make span	Scalability	Trust	Reliability	Resource utilization	Energy consumption	Load balancing	fairness
Hybrid energy efficient scheduling algorithm [7]	✓	×	×	×	×	×	×	×	✓	✓	×
Energy efficient algorithm using migration [8]	×	✓	×	×	×	×	×	×	✓	✓	×
Near optimal scheduling policy [9]	×	×	✓	×	×	×	×	✓	×	✓	×

Adaptive-DVS, §-Advanced-DVS [10]	×	✓	×	×	×	×	×	×	✓	×	×
Ant Colony Algorithm [11]	×	×	×	✓	×	×	×	×	×	✓	×
An algorithm of scheduling & dynamic load balancing [12]	✓	×	×	×	×	×	×	×	×	×	×
Chaos genetic algorithm [13]	✓	×	✓	×	×	×	×	×	×	×	×
Gang scheduling algorithm [14]	×	✓	✓	×	×	×	×	×	×	×	×
Data Analytic Cloud Architecture [15]	×	✓	✓	×	×	×	×	×	×	×	✓
Deadline Budget Constraint Scheduling Policy [16]	✓	✓	✓	×	×	×	×	×	×	×	×
particle swarm optimization based meta heuristics method [17]	×	×	✓	×	×	×	×	×	×	×	×
A trust dynamic level scheduling algorithm [18]	×	×	×	×	×	✓	×	×	×	×	×
Resource allocation framework [19]	×	×	×	×	×	×	×	✓	×	✓	×
Workflow scheduling algorithm [20]	×	×	×	×	✓	×	✓	×	×	×	×
Compromised -Time-Cost Scheduling Algorithm [21]	✓	×	✓	✓	×	×	×	×	×	×	×

4. Conclusion

Cloud computing is one of the user oriented technology in which user faces hundreds of thousands of virtualized resources for each task. In this paper we survey various existing scheduling algorithms in cloud, grid and workflows. The table 1 is shown for further reference. Since cloud computing is in infancy state, a scheduling framework should be implemented to improve the user satisfaction along with the service providers. The scheduling parameters can be coupled to prepare a framework for resource allocation and scheduling in cloud computing. The scheduling framework should consider the user input constraints (execution cost, deadlines, transmission cost, energy efficiency, performance issues, and makespan) and so on.

References

[1] Introduction To Cloud Computing Architecture white paper, *Sun Microsystems*,1st edition, June 2009.

[2] NIST Definition of Cloud Computing v15, csrc.nist.gov/groups/SNS/cloud-computing/cloud-def v15.doc

[3] Essential characteristics of Cloud Computing, white paper.

[4] Introduction to cloud computing, *Dialogic making innovation thrive*, july 2010 white paper.

[5] The future of cloud computing, opportunities for european cloud computing beyond 2010, public version1.0.

[6] Baomin Xu , Chunyan Zhao,Enzhao Hua,Bin Hu, "Job scheduling algorithm based on Berger model in cloud environment", *Elsevier publications*, march 2011.

[7] Jiandun Li, Junjie Peng, Zhou Lei, Wu Zhang, "An Energy-efficient Scheduling Approach Based on Private Clouds", *Journal of Information & Computational Science*, april 2011.

[8] Jiandun Li, Junjie Peng, Zhou Lei, Wu Zhang, "A Scheduling Algorithm for Private Clouds", *Journal of Convergence Information Technology*, Volume6, Number 7, July 2011.

[9] Saurabh Kumar Garg, Chee Shin Yeo,Arun Anandasivam, Rajkumar Buyya, "Energy-Efficient Scheduling of HPC Applications in Cloud Computing Environments", *Elsevier publications*, September 2009.

[10] Kyong Hoon Kim, Anton Beloglazov, Rajkumar Buyya, "Power-aware Provisioning of Cloud Resources for Real-time Services", *ACM Publications*, December 2009.

[11] Ruay-Shiung Chang, Jih-Sheng Chang, Po-Sheng Lin, "An ant algorithm for balanced job scheduling in grids", *Future Generation Computer Systems*, june 2008.

[12] Ungurean, "Job Scheduling Algorithm based on Dynamic Management of Resources Provided by Grid Computing Systems", feb 2010

[13] Golnar Gharoonifard, Fahime Moeindarbari, Hossein Deldari, Anahita Morvaridi, "Scheduling of scientific workflows using a chaos-genetic algorithm", *Elsevier publications*, 2010.

[14] Ioannis A.Moschakis, Helen D. Karatza, "Evaluation of gang scheduling performance and cost in a cloud computing system", *Springer publications*, 2010.

[15] Gunho Lee, Byung-Gon Chun, Randy H. Katz, "Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud".

[16] Mohsen Amini Salehi, Rajkumar Buyya, "Adapting Market-Oriented Scheduling Policies for Cloud Computing".

[17] Suraj Pandey, LinlinWu, Siddeswara Mayura Guru, Rajkumar Buyya, "A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments".

[18] Wei Wang , Guosun Zeng , Daizhong Tang , Jing Yao, "Cloud-DLS: Dynamic trusted scheduling for Cloud computing", *Expert Systems with Applications*, 2011.

[19] Muhammad Asad Arfeen, Krzysztof Pawlikowski, Andreas Willig, "A Framework for Resource Allocation Strategies in Cloud Computing Environment"

[20] Anju Bala, Inderveer Chana, "A Survey of Various Workflow Scheduling Algorithms in Cloud Environment", 2nd National Conference on Information and Communication Technology (NCICT) 2011, *International Journal of Computer Applications publications*.

[21] Ke Liu, Hai Jin, Jinjun Chen, Xiao Liu, Dong Yuan, Yun Yang, "A Compromised-Time-Cost Scheduling Algorithm in SwinDeW-C for Instance-Intensive Cost-Constrained Workflows on Cloud Computing Platform".