

A Literature Survey on Resource Management Techniques, Issues and Challenges in Cloud Computing

Frederic Nzanywayingoma*, Yang Yang

Department of Computer Science and Communication Engineering,
University of Science and Technology Beijing, Beijing, 100083-China

*Corresponding author: nzanywafre@yahoo.fr

Abstract

Cloud computing is a large scale distributed computing which provides on demand services for clients. Cloud Clients use web browsers, mobile apps, thin clients, or terminal emulators to request and control their cloud resources at any time and anywhere through the network. As many companies are shifting their data to cloud and as many people are being aware of the advantages of storing data to cloud, there is increasing number of cloud computing infrastructure and large amount of data which lead to the complexity management for cloud providers. We surveyed the state-of-the-art resource management techniques for IaaS (infrastructure as a service) in cloud computing. Then we put forward different major issues in the deployment of the cloud infrastructure in order to avoid poor service delivery in cloud computing.

Keywords: cloud computing; resource management techniques; resource management challenges; migration techniques; load balancing

Copyright © 2017 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Cloud computing is a modern paradigm to provide services through the Internet. The cloud infrastructure often sit in geographically dispersed data centers around the world and the customers obtain services through internet. A data center is a physical area where IT equipment reside such as disk enclosures, controller enclosure, routers, switches, servers, load balancers, firewalls, power delivery devices, cooling machines, unmanaged devices, and other security devices. Those IT equipment can be remotely managed or monitored IT Engineers using monitoring software. We refer to the figure 1 below to show the cloud resources in data center where a number of racks contain a certain number of servers, swithes or routers stacked one above the other.

The virtualization technologies enabled the realization of large-scale data centers resources to be leased and released by third parties through pay per use model [1]. Since cloud computing resources are scattered across various places and its use continues to grow, companies, universities and research centers have high-performance needs which lead to the resource management problems. According to the NIST [2], cloud computing resources are classified as IaaS (infrastructure as a Service), SaaS (Software as a Service) and PaaS (Platform as a Service). One of the key problem in cloud computing is the management of how to procure and release the resources [3]. Other cloud characteristics such as the interdependancies of heterogeneous resource types, the variation and unpredictability of the workload all contribute to the complexity of cloud environment management.

To address the above mentioned problems, we referred to the works done by other researchers on this topic. Currently, we found that there exist few papers and most of them lack detailed working on each of the resource management technique and demonstrate limited critical discussions on the existing issues and challenges. Author in [4] shows the state-of-the-art resource provisioning technique. Authors have based on some metrics such as cost, utility, and service quality. Author in [5] also has studied dynamic resource provisioning technique using the online auction framework. Authors in [6] have proposed an IEDA based on auction protocol. Another technique of the resource management was studied in [7] using batch

scheduling strategies. For most switable resources, paper [8] has suggested MADM(multi-attribute decision making) technique for multiple resource discovery. Authors have focused on the methods of matching resources at a super-peer nodes based on matchmaker algorithm.

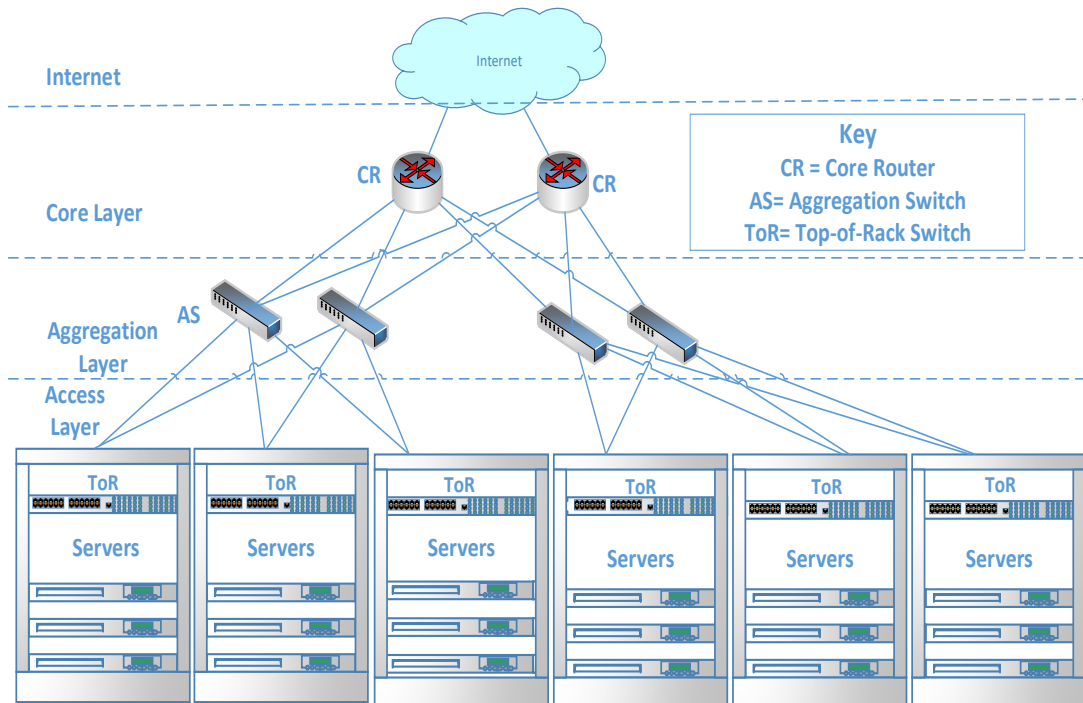


Figure 1. Data Center Network Topology

According to [9], the survey study on VMs placement for cloud computing shows that virtual machines (VMs) can be mapped onto physical machines (PMs) to optimize cloud resources. Authors in [10] have outlined the use of self-organizing algorithms in resource management such as bio-inspired computing, multi-agent system, bacterial foraging algorithm (BFA), shuffle flog algorithm leaping (SFLA), evolutionary techniques and have highlighted few of the main open challenges related to the area such as energy efficiency, and security issues. The security issues in cloud resources also were highlighted in [11].

Our contributions in this survey study is to conduct a comprehensive study on different techniques that can be used to solve resource management problems. For this paper, we reviewed 300+ recent research papers, which we selected about a third for this article. The rest of the paper is divided into the following section: Section 2 outlined the scope of the resource management. Section 3 presented resource management techniques. Section 4 discussed some issues and solutions for resource management. Section 5 showed the remaining open problems, and suggestions. Section 6 concluded the paper.

2. The Scope of Resource Management in Cloud Computing

According to Peter Mell and Timothy Grance from NIST [2], cloud computing is defined as a model for providing services to configurable computing resources which can be rapidly leased and released with cloud user's minimal management effort or with a service provider interaction. The cloud computing service models are: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS).

2.1 Classifications of Cloud Computing Resources

Cloud computing resources are classified into physical resources, virtual resources, and logical resources. The physical resources are processors, memory, disk drive, Network

Interface Controller (NIC), peripheral devices (floppy, keyboard), network resources (networking products), storage media, and other physical components. The physical resources are distributed around the world in data centers for faster serving the cloud clients. Virtualization is a technique to divide the computer resources into execution environments to make the physical machine to be more productive. It allows multiple instances of an operating system and associated application to run on single physical machine. The virtual resources (VR) are dynamically shared, scheduled, and scaled according to consumer demand. Virtual resources are virtual processor (vCPU), virtual memory, virtual Switch (vSwitch), virtual storage (vSAN).

Through the virtualization technology, multiple resources are dynamically assigned and reassigned. Those resources are managed with resource management system (RMS) such that when a user orders resources, the RMS checks the resources availability and decides if they satisfy users' requests. The RMS is responsible to monitor continuously the user's allocated resources while optimizing the system [10]. Virtual and Physical servers in data-centers host many applications types such as webservers, databases, customer business apps, etc. Virtual resource management tools and technologies in cloud data-center has the cloud computing platforms shifts toward providing infrastructure as service(IaaS) where computing resources are provisioned on a per-demand resource basis regardless of location of the available computing resources [12]. Logical resources management are system abstractions which have temporary control over physical resources [13] (i.e operating system, energy, network throughput/bandwidth, information security, protocols, APIs, network loads, and delays).

2.2. The Requirements of Cloud Computing Resource Management

Cloud computing presents five characteristics and their requirements: on-demand self-service, broadband network access, resource pooling, rapid elasticity, and measured service.

Table 1. Resource Management Characteristics and its Requirements

	Characteristics	Requirements	Objectives
	On-demand self-service	Intelligent and business-related resource management	
	Broad network access	End-to-end resource chain management	
	Resource pooling	Dynamic deployment management of virtual resources.	QoS & Cost optimization
Cloud Computing	Rapid elasticity	Dynamic adaptive resource management.	
	Measured service	Monitoring and reporting of resource usage at appropriate level of abstraction	
Green Cloud computing		Energy-efficient resource management	

2.3 Challenges in Cloud Computing Resource Management

There exists several problems to manage cloud resources. In this section, we surveyed some of the challenges related to cloud resource management. When a cloud users make requests to provide of one or more resources, the cloud provider's resource management system (RMS) schedules the available resources by making them available to the cloud users but some challenges are still available such as the capabilities to minimize energy consumption, minimize the cost, and maximize the performance. The efficient strategies for resource management are as follows [14-17]: energy-efficient, bandwidth cost minimization, storage and performance optimization. We limit ourselves to the brief review of some works addressing power resources, networking resources, compute resources, storage resources challenges.

The power resource is a big issue in cloud resource management [18]. The power consumption rate for idle servers in datacenters is 60% and the rest of power coupled with cooling equipment, and the security system. In order to improve the energy efficiency it is advised to turn off the idle servers and automate the cooling system. Other power management solutions are: Dynamic Component Deactivation (DCD), Dynamic Performance Scaling (DPS), and Dynamic Voltage and Frequency Scaling (DVFS) [19]. The DVFS scheme maximizes energy consumptions and user profits. The study in [20] outlines an energy-aware algorithm that

supports DVFS. Paper [21], proposes a multi-agent-based resource management to minimize the energy consumption.

The network resource also is a big challenge in cloud computing. The study in [22, 23] proposed the VM resource scheduling algorithm based on auction mechanism which considered network bandwidth cost minimization. The compute resource is important and can be optimization. Edwardo, et al.[24] has presented a dynamic programming approach to optimize the computing resources. Traditional storage resource management system faces many challenges in cloud computing. A modern resource management system is needed to manage cloud storage resources.

2.4. Strategies of Cloud Computing Resource Management

Resource management strategies are classified into resource monitoring, resource allocation, resource consolidation, resource discovery, resource scheduling, resource sharing, resource modeling, and resource provisioning. We reviewed various strategies developed so far for resource management.

(1) Automatic resource allocation and dynamic resource allocation are the strategies used for resource management to increase the profit of cloud resource utilization, and minimize the response time.

(2) Network-Aware Resource Management strategy solves efficient network resource sharing issue, and network congestion issues based on SDN (Software-Defined Network) resource management strategy.

(3) Resource scheduling, we adopt two types of resource scheduling: Processor scheduling and server scheduling. Single and multiprocessor scheduling should be used to reduce energy consumption. Server scheduling helps to find the idle machines to be turn off when no services running on them to save the energy. Using this strategy, server can dynamically start, hang, or turned off.

(4) Resource provisioning can solve the overall workflow execution cost and deadline constraints.

(5) Resource discovery helps to locate desired resources without the global knowledge about shared resource information with less human intervention (i.e multi-valued distributed hash tables).

(6) Resource monitoring strategy allows monitoring the resource information such as compute (CPU usage, memory usage, and operations), capacity, network, power, security, and event monitoring.

(7) Resource consolidation strategy allows consolidating resources (server consolidation, data-center consolidation, etc.)

3. Resource Management Techniques

The main objective of resource management is to reduce the service costs, improve the performance, security, and energy efficiency. Here, we considered different resource management techniques based on parameters like monetary parameters (service cost), application performance parameters (response time, execution time, delay, SLA violation, task type, number of processors needed to run tasks, throughput, resource availability, resource utilization, etc), security, and energy-efficiency parameters (overall power, energy consumption) [25]. The algorithms such as load balancing, Round Robin, Bin packing algorithm and gradient search algorithm were mentioned as the optimization methods to improve the performance, reduce the cost, and minimize the energy consumption in IaaS resources [26-29].

The research in [30] considered a stochastic model based on load balancing and scheduling in cloud computing clusters [31, 32]. Another resource management methods were proposed in [33] which use the Self-Organizing Clouds (SOC) to achieve the maximum of the resource utilization and to deliver optimal execution time. Table 2 gives a summary comparison of cloud resources based on some metrics.

3.1. Virtualization Techniques

Virtualization is a part of cloud computing and is very worthwhile concept [34]. This concept has started back in 1964 with IBM project called CP/CMS system and later spreaded all the world [35]. Virtualization technique is a framework to divide the resources of a computer into

execution environments to make the physical machine to be productive for the resource utilization and resource management. There are different kind of virtualization technology such as storage virtualization, server virtualization, network virtualization, client virtualization, and operation framework virtualization. The network virtualization is based on virtualized network switch technology commonly known as vSwitch. Client Virtualization is based on creating a client desktop known as VDI (virtual desktop infrastructure). Storage virtualization is a virtualization technique based on connecting the storage system to the application server over the network. It is a solution to cloud users to store their data onto unified cloud platforms. Among the above mentioned techniques, we added the following techniques:

Table 2. Comparative Analysis on Different Resource Management Techniques

References	Techniques	Metrics					
		QoS	Utilization rate	Cost	Energy Efficiency	Scalability	Security
Mazumdar et al., (2016) [36]	Resource consolidation	x	✓	✓	✓	x	x
Guang, et al.(2017) [37]		x	✓	✓	✓	x	x
Ligang et al.(2012)[38]		✓	✓	✓	x	x	x
Amit, et al. (2017)[39]	Resource adaptation	x	x	✓	x	✓	x
Abdul, et al. (2016) [40]		x	✓	✓	✓	✓	x
Arani, et al. (2017)[41]		x	✓	✓	✓	✓	✓
Gihun Jung, al., (2011) [42]	Resource allocation	x	✓	x	x	x	x
Sheng, et al. (2013) [33]		✓	✓	✓	x	✓	x
Vakilinia, et al. (2015) [16]		x	✓	✓	✓	x	x
Xiaomin,et al.(2016)[43]	Resource provisioning	x	✓	x	x	✓	x
Shi, et al.(2014)[44]		x	x	✓	x	✓	x
Aarti, et al. (2015)[45]		x	✓	✓	x	x	x
Chung, et al.(2011)	Resource discovery	x	✓	✓	x	✓	x
Asma, et al.(2016)[46]		✓	x	x	x	x	x
Wright, et al. (2012)[47]		x	x	v	x	x	x
Syed, et al. (2016)[15]	Resource scheduling	✓	✓	✓	✓	✓	x
Weiwe, et al.(2016)[23]		x	✓	✓	✓	x	x
Sukhpal, et al., (2016)[48]		✓	✓	✓	✓	✓	✓
Feng, ying (2015)[49]	Resource blokering	x	x	✓	✓	x	x
Chetan, et al.(2014)[50, 51]		✓	x	✓	x	x	x
Shahin, et al. (2015)[16]		Modeling and	x	✓	✓	✓	x
Devi, et al. (2013) [52]	Resource estimation	x	x	✓	✓	x	x
Gusheng, et al.(2016)[53]		✓	x	x	x	x	x
Iliia, et al.(2016)[9]		Resource mapping	✓	✓	✓	✓	x
Yang, et al. (2014)[54]	Resource mapping	x	✓	x	x	x	x
Hao, et al.(2013)[55]		x	✓	✓	x	x	x

3.2. Migration Techniques

Migration technique is the process of moving an operating system instance from one node to another node without interrupting an operation in order to improve the system reliability and availability [56]. It facilitates load balancing, low level system maintenance, fault tolerant, and energy efficiency.

Due to the sensitivity of the migration process, some performance metrics have to be considered such as preparation time, memory pages transferred estimation, total migration time, and application performance degradation.

- a. Live migration
- b. Offline migration
- c. Cold migration
- d. Hybrid migration
- e. Storage Migration

There are many VM storage data migration solutions. In this paper, the following storage data migration solutions should be cited:

- i. LVM mirroring based data migration solution
- ii. LDM Mirroring-based Data Migration Solution
- iii. SmartVirtualization and SmartMigration based data migration solution:

- iv. LUN copy based data migration
- v. ASM based data migration solution
- vi. VIS-based Data Migration Solution
- vii. FastCopy-based Data Migration Solution

There are other storage resource management functions to be considered such as SmartThin, smartTier, SmartQoS, SmartMotion, SmartPartition, etc.

3.3 Other techniques

3.3.1. Server Optimization

The optimization for server migration helps on improving the live migration and optimizing the performance metrics such as total migration time, and downtime while providing uninterrupted services to applications running in VMs. The following techniques should be listed:

- a. Memory Page Compression
- b. Delta page transfer
- c. Data de-duplication
- d. Post-copy technique
- e. Hybrid pre and post copy
- f. Server consolidation
- g. The memory buddies

3.3.2. Load Balancing with Data Migration

The more VMs can be placed on the same host, the greater the workload and the increase of the VMs migration and the degradation of the system performance. Load Balancing with data migration is important in virtual environment. Some load balancing algorithms were proposed such as self acting load balancing algorithm, and many others. The advantage of load balancing algorithms are to enhance the performance in data migration and the resource utilization rate. Considering a large parallel file system, where data are transferred between the memory and the storage devices. There are different challenges that are faced during load balancing such as scalability, and the availability, and the load migration. To overcome those challenges, the hotspot mitigation algorithm [57] came up to help to determine which heavily-loaded VMs to migrate first.

3.3.3. Hotspot & Coldspot Migration

Hotspots and coldspots are based on the thresholds set by the cloud owner or based on the SLA specified by the clients. Hotspot uses a memory-to-memory stop-and-copy migration technique that copies the container's memory state from the memory of the destination VM without saving it to stable storage.

The advantage of hotspot migration is to minimize cost and the disadvantage is that it does not consider the application performance when deciding where and when to migrate. In order to migrate data, hotspot estimates the resource utilization by considering: Low utilization and high utilization threshold [58].

4. Solution to Resource Management Issues

The top resource management issues are as the follows:

- a. Dynamic Fault-tolerant: Dynamic fault-tolerant scheduling algorithm takes in consideration to use full-idle resources and VM migration [43]
- b. Failure-Aware resource provisioning: Failure-aware resource provisioning algorithm considers the workload model to redirect users' requests to the appropriate cloud providers [59].
- c. Energy efficiency: Towards efficient resource provisioning allows users to identify the best trade-off point between performance and energy efficiency [60].
- d. Resource availability: Considers these metrics: mapping, finishing time, cost and resources to meet the optimization objectives [61].
- e. Bandwidth minimization and low latency: Author proposed a bandwidth minimization and low latency based on a polynomial-time routing algorithm, integer linear programming (ILP) and heuristic algorithm [62].

- f. Hybrid approach: Hybrid approach combined the concept of the autonomic computing and Reinforcement Learning to predict the future demands and to deal with fluctuating demands of the cloud services [63].
- g. Cost minimization: Authors used Linear programming approaches to optimize the cost and latency at the same time.

5. Open Problems in cloud resource management

In this work, we analyzed the resource management technique issues and discussed their merits such as throughput, availability, high scalability, quality of Service, optimal utility, reduced latency, reduced overheads, communication cost, computational time, and energy efficiency. The resource management technique demerits are: less security, Low priority, system performance degradation due to migration effect, network delay, and more time is used for balancing the workload. Also we are mentioning some of the unresolved issues such as:

5.1. Performance issues in cloud resources

The performance guarantee is the important issue between the cloud service providers and cloud users. The possible ways to improve the performance of resources and applications running in the cloud are suggested here.

- a. Identify the performance parameters and evaluation metrics: The following metrics are used to evaluate the performance of cloud resources: IOPS including read IOPS, write IOPS, total response time, read response time or write response time, total bandwidth, write bandwidth, and so on.
- b. Apply performance metrics: Use the mentioned metrics to measure and capture the QoS level to gain user satisfaction. The QoS levels include the number of SLA violations, response time, delay, etc.
- c. Apply the energy efficiency metrics: Some energy improvements metrics are considered to measure the overall power and energy consumption. The power saving methods were suggested such as DVFS, turning on/off, DVS, VM migration [64], VM placement [65], Hibernation [66].
- d. Apply monetary metrics: the monetary metrics is based on the evaluation of service provider's cost, and user cost.

5.2. Security issues in cloud resources

The cloud providers may face with the security and privacy issues in PaaS and IaaS. Despite that the cloud infrastructure and other related services provide basic security firewall, the problem of security and privacy are still the major concerns in cloud computing since the users have no control over their data. According to authors in [67], the cloud providers may convince the third parties the safety of their data and their infrastructure against both exterior and domestic attacks. That hope is however rely on organizational procedures but it cannot be verified or remotely validated by the third parties [68].

Security issues in cloud resources can be classified into security issues in SaaS, in PaaS, and in IaaS. They can be grouped into key security elements such as Data security; Network security; Data locality; Data integrity; Data segregation; Data access; and Authentication and authorization. Cloud security alliance reports [69, 70] and studies on cloud security issues [71] identify several security threats such as data breaches, data leakage, data loss, denial of service, accounts, services and traffic hijacking, insecure application programming interfaces, and many more others which can occur in cloud system information management. Some security approaches were suggested such as confidentiality, availability, reliability, authentication and authorization, integrity, etc. Other approaches on authorization and authentication to enforce identity and access management policies are Public key infrastructure (PKI) X509 certificates, IBE, IBS, CP-ABE [72-77].

5.3. Scheduling in Cloud Resources

Cloud computing is the innovation gives away to use shared pool of resources globally. Those resources can be dynamically upscale or downscale accordingly. In order to adapt to the effective cloud resource management and deal with resource scheduling, we adopt two types of

resource scheduling: Processor resource scheduling and server resource scheduling. The efficient processor resource scheduling leads to many advantages. We deal with processor resource scheduling based on the preemptive strategy. In cloud computing, the processor resources are encapsulated into jobs. The jobs with highest priority should be scheduled prior to jobs with lowest priority. Common metrics can be considered: makespan, load balancing, average waiting time, throughput, turnaround time, energy efficient, etc. Lack of effective process resource scheduling methods may lead to overheating of cloud resources (Compute and storage). Single and multiprocessor resource scheduling should be considered to reduce energy consumption [78-80]

Server resource scheduling: Virtualization is a part of cloud computing and is very worthwhile concept. Virtual machines are managed and hosted by physical servers and share the physical resources such as memory, storage, processor. The purpose for server resource scheduling is to find the idle machines to be turn off when no services running on them to save the energy. Another purpose of server resource scheduling is to balance the workload between servers to improve the system performance, maintain the system stability, and build the system fault tolerance.

6. Conclusion

In our paper, we have surveyed various cloud resource management techniques for cloud computing. Among those techniques we cite resource provisioning, resource discovery, resource monitoring, resource mapping, resource allocation, and resource consolidation. The main purpose of cloud resource management techniques is to satisfy the customer requirements. As described above, though there are huge advantages in applying resource management techniques, still, several outstanding issues exist which can be solved to make the cloud system more manageable and efficient. Those open challenges are resource performance issues, security and privacy issue, and server, processor, workload resources scheduling issues. The following metrics were mentioned: CPU utilization, throughput, reduced number of SLA violations, response time, reduced delay, energy efficiency, total bandwidth, reduced latency, reduced overheads, specialized environment, communication cost, etc. In our next paper we will focus on cloud resource scheduling based on meta-heuristic techniques.

References

- [1] Jennings, B, R Stadler. Resource management in clouds: Survey and research challenges. *Journal of Network and Systems Management*, 2015; 23(3): p. 567-619.
- [2] Mell, P, T Grance. The NIST definition of cloud computing. 2011.
- [3] Arianyan, E, H Taheri, S Sharifian. Novel energy and SLA efficient resource management heuristics for consolidation of virtual machines in cloud data centers. *Computers & Electrical Engineering*, 2015; 47: p. 222-240.
- [4] Zhang, J, H Huang, X Wang. Resource provision algorithms in cloud computing: A survey. *Journal of Network and Computer Applications*, 2016; 64: p. 23-42.
- [5] Shi, W, et al. An online auction framework for dynamic resource provisioning in cloud computing. *IEEE/ACM Transactions on Networking*, 2016; 24(4): p. 2060-2073.
- [6] Wang, X, et al. An intelligent economic approach for dynamic resource allocation in cloud services. *IEEE Transactions on Cloud Computing*, 2015; 3(3): p. 275-289.
- [7] Jintao, J, Y Wensen, G Lei. Research on Batch Scheduling in Cloud Computing. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 2016; 14(4): p. 1454-1461.
- [8] Kaur, M, SS Kadam. Discovery of resources using MADM approaches for parallel and distributed computing. *Engineering Science and Technology, an International Journal*, 2017.
- [9] Pietri, I, R Sakellariou. Mapping virtual machines onto physical machines in cloud computing: A survey. *ACM Computing Surveys (CSUR)*, 2016; 49(3): p. 49.
- [10] Endo, PT, et al. *Self-organizing strategies for resource management in Cloud Computing: State-of-the-art and challenges*. in Cloud Computing and Communications (LatinCloud), 2nd IEEE Latin American Conference on. 2013. IEEE.
- [11] Singh, A, K Chatterjee. Cloud security issues and challenges: A survey. *Journal of Network and Computer Applications*, 2017; 79: p. 88-115.
- [12] Nguyen Van, H, F Dang Tran, J-M. Menaud. *Autonomic virtual resource management for service hosting platforms*. in Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing. 2009. IEEE Computer Society.

- [13] Manvi, SS, GK Shyam. Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey. *Journal of Network and Computer Applications*, 2014; 41: p. 424-440.
- [14] Parikh, SM. *A survey on cloud computing resource allocation techniques*. in Engineering (NUICONE), 2013 Nirma University International Conference on. 2013. IEEE.
- [15] Madni, SHH, et al. Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities. *Journal of Network and Computer Applications*, 2016; 68: p. 173-200.
- [16] Vakiliinia, S, MM Ali, D Qiu. Modeling of the resource allocation in cloud computing centers. *Computer Networks*, 2015; 91: p. 453-470.
- [17] Han, D, W Sun, X Fan. Dynamic energy management in smart grid: A fast randomized first-order optimization algorithm. *International Journal of Electrical Power & Energy Systems*, 2018; 94: p. 179-187.
- [18] Lee, L-T, et al., A dynamic resource management with energy saving mechanism for supporting cloud computing. *International Journal of Grid and Distributed Computing*, 2013; 6(1): p. 67-76.
- [19] Kim, N, J Cho, E Seo. Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems. *Future Generation Computer Systems*, 2014. 32: p. 128-137.
- [20] Ding, Y, et al. Energy efficient scheduling of virtual machines in cloud with deadline constraint. *Future Generation Computer Systems*, 2015; 50: p. 62-74.
- [21] Wang, W, Y Jiang, W Wu. Multiagent-Based Resource Allocation for Energy Minimization in Cloud Computing Systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2016.
- [22] Younge, AJ, et al. *Efficient resource management for cloud computing environments*. in Green Computing Conference, 2010 International. 2010. IEEE.
- [23] Kong, W, Y Lei, J Ma. Virtual machine resource scheduling algorithm for cloud computing based on auction mechanism. *Optik-International Journal for Light and Electron Optics*, 2016; 127(12): p. 5099-5104.
- [24] Bacalhau, ET, FL Usberti, C Lyra. *A dynamic programming approach for optimal allocation of maintenance resources on power distribution networks*. in Power and Energy Society General Meeting (PES), 2013 IEEE. 2013. IEEE.
- [25] Lee, G, RH Katz. *Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud*. in HotCloud. 2011.
- [26] Ngenzi, A, SR Nair. *Dynamic resource management in Cloud datacenters for Server consolidation*. arXiv preprint arXiv:1505.00577, 2015.
- [27] Coffman, EG, MR. Garey, DS Johnson, Dynamic Bin Packing. *SIAM Journal on Computing*, 2006. 12(2): p. 227-258.
- [28] Koneru, S, VR Uddandi, S Kavuri. Resource Allocation Method using Scheduling methods for Parallel Data Processing in Cloud. *International Journal of Computer Science and Information Technologies [IJCSIT]*, 2012; 3(4): p. 4625-4628.
- [29] Jiang, W. et al. A Resource scheduling strategy in cloud computing based on multi-agent genetic algorithm. *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, 2013; 11(11): p. 6563-6569.
- [30] Maguluri, ST, R Srikant, L Ying. *Stochastic models of load balancing and scheduling in cloud computing clusters*. in international conference on computer communications. 2012.
- [31] Feng, T, J Bi, K Wang. *Joint allocation and scheduling of network resource for multiple control applications in SDN*. in Network Operations and Management Symposium (NOMS), 2014 IEEE. 2014. IEEE.
- [32] Moens, H, et al. *Hierarchical network-aware placement of service oriented applications in clouds*. in Network Operations and Management Symposium (NOMS), 2014 IEEE. 2014. IEEE.
- [33] Di, S, C-L. Wang, Dynamic optimization of multiattribute resource allocation in self-organizing clouds. *IEEE Transactions on parallel and distributed systems*, 2013; 24(3): p. 464-478.
- [34] Leelipushpam, PGJ, J. Sharmila. *Live VM migration techniques in cloud environment-a survey*. in Information & Communication Technologies (ICT), 2013 IEEE Conference on. 2013. IEEE.
- [35] Rodríguez-Haro, F, et al. A summary of virtualization techniques. *Procedia Technology*, 2012; 3: 267-272.
- [36] Mazumdar, S, M Pranzo. Power efficient server consolidation for Cloud data center. *Future Generation Computer Systems*, 2017; 70: 4-16.
- [37] Han, G, et al. Resource-utilization-aware energy efficient server consolidation algorithm for green computing in IIOT. *Journal of Network and Computer Applications*, 2017.
- [38] He, L, et al. Developing resource consolidation frameworks for moldable virtual machines in clouds. *Future Generation Computer Systems*, 2014; 32: 69-81.
- [39] Mitra, A, N O'Regan, D Sarpong. Cloud resource adaptation: A resource based perspective on value creation for corporate growth. *Technological Forecasting and Social Change*, 2017.
- [40] Hummaida, AR, NW Paton, R Sakellariou. Adaptation in cloud resource configuration: a survey. *Journal of Cloud Computing*, 2016; 5(1): 1-16.
- [41] Bhattacharya, A, P De. A survey of adaptation techniques in computation offloading. *Journal of Network and Computer Applications*, 2017; 78: 97-115.

- [42] Jung, G, KM Sim. *Agent-based adaptive resource allocation on the cloud computing environment*. in Parallel Processing Workshops (ICPPW), 2011 40th International Conference on. 2011. IEEE.
- [43] Zhu, X, et al. Fault-tolerant scheduling for real-time scientific workflows with elastic resource provisioning in virtualized clouds. *IEEE Transactions on Parallel and Distributed Systems*, 2016; 27(12): 3501-3517.
- [44] Shi, W., et al., An online auction framework for dynamic resource provisioning in cloud computing. *ACM SIGMETRICS Performance Evaluation Review*, 2014. 42(1): 71-83.
- [45] Singh, A, D Juneja, M Malhotra, A novel agent based autonomous and service composition framework for cost optimization of resource provisioning in cloud computing. *Journal of King Saud University-Computer and Information Sciences*, 2015.
- [46] Alkalbani, AM, FK Hussain. *A comparative study and future research directions in cloud service discovery*. in Industrial Electronics and Applications (ICIEA), 2016 IEEE 11th Conference on. 2016. IEEE.
- [47] Wright, P, et al. A constraints-based resource discovery model for multi-provider cloud environments. *Journal of cloud computing: advances, systems and applications*, 2012; 1(1): 6.
- [48] Singh, S, I Chana, A survey on resource scheduling in cloud computing: Issues and challenges. *Journal of grid computing*, 2016; 14(2): p. 217-264.
- [49] Ma, F, Y, Yang. *An energy sentient service brokering strategy in cloud computing*. in Control And Decision Conference (CCDC), 2017 29th Chinese. 2017. IEEE.
- [50] Awasthi, C, P, Kanungo. *Client Requirement Modeling Using Resource Broker Architecture in Cloud Computing Environment*. in Cloud Computing in Emerging Markets (CCEM), 2014 IEEE International Conference on. 2014. IEEE.
- [51] Roy, DG, et al. *Multi-cloud scenario based QoS enhancing virtual resource brokering*. in Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on. 2016. IEEE.
- [52] Devi, KV, K Mehata. *Modelling resource estimation and scheduling for processing in cloud computing*. in Advanced Computing (ICoAC), 2013 Fifth International Conference on. 2013. IEEE.
- [53] Fan, G, H Yu, L Chen. A formal aspect-oriented method for modeling and analyzing adaptive resource scheduling in cloud computing. *IEEE Transactions on Network and Service Management*, 2016. 13(2): 281-294.
- [54] Song, Y, R Routray, R Jain. *Virtual-to-physical mapping inference in virtualized cloud environments*. in Cloud Engineering (IC2E), 2014 IEEE International Conference on. 2014. IEEE.
- [55] Di, H, et al. *Reliable virtual infrastructure mapping with efficient resource sharing*. in Communications, Circuits and Systems (ICCCAS), 2013 International Conference on. 2013. IEEE.
- [56] Kale, RMCO, *Virtual Machine Migration Techniques in Cloud Environment: A Survey*.
- [57] Wood, T, et al. Black-box and gray-box strategies for virtual machine migration. in *networked systems design and implementation*. 2007.
- [58] Xu, Y, Y Sekiya. *Virtual machine migration strategy in federated cloud*. in Internet Conference. 2010.
- [59] Javadi, B, J Abawajy, R Buyya. Failure-aware resource provisioning for hybrid Cloud infrastructure. *Journal of Parallel and Distributed Computing*, 2012; 72(10): 1318-1331.
- [60] Nghiem, PP, S Figueira. Towards efficient resource provisioning in MapReduce. *Journal of Parallel and Distributed Computing*, 2016; 95: 29-41.
- [61] Kaur, P, S Mehta. Resource provisioning and work flow scheduling in clouds using augmented Shuffled Frog Leaping Algorithm. *Journal of Parallel and Distributed Computing*, 2017; 101: 41-50.
- [62] Yang, S, et al., Energy-Aware Provisioning in Optical Cloud Networks. *Computer Networks*, 2017; 118: p. 78-95.
- [63] Ghobaei-Arani, M, S Jabbehdari, MA Pourmina. An autonomic resource provisioning approach for service-based cloud applications: A hybrid approach. *Future Generation Computer Systems*, 2017.
- [64] Tesfatsion, SK, E Wadbro, J Tordsson. *A combined frequency scaling and application elasticity approach for energy-efficient cloud computing*. Sustainable Computing: Informatics and Systems, 2014; 4(4): p. 205-214.
- [65] Khosravi, A, SK Garg, R Buyya. *Energy and carbon-efficient placement of virtual machines in distributed cloud data centers*. in European Conference on Parallel Processing. 2013. Springer.
- [66] Subrata, R, AY Zomaya, B Landfeldt. Cooperative power-aware scheduling in grid computing environments. *Journal of Parallel and Distributed Computing*, 2010; 70(2): p. 84-91.
- [67] Sen, J. *Security and privacy issues in cloud computing*. Architectures and Protocols for Secure Information Technology Infrastructures, 2013: p. 1-45.
- [68] Nagesh, OS, T Kumar, VR Vedula, *A Survey on Security Aspects of Server Virtualization in Cloud Computing*. *International Journal of Electrical and Computer Engineering (IJECE)*, 2017; 7(3): p. 1326-1336.
- [69] Islam, T, D Manivannan, S Zeadally. A classification and characterization of security threats in cloud computing. *Int. J. Next-Gener. Comput*, 2016; 7(1).
- [70] Sutikno, T, D Stiawan, IMI Subroto. Fortifying big data infrastructures to face security and privacy issues. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 2014; 12(4): p. 751-752.

-
- [71] Gholami, A, E. Laure, *Security and privacy of sensitive data in cloud computing: a survey of recent developments*. arXiv preprint arXiv:1601.01498, 2016.
- [72] Kim, H, SC. Timm. X. *509 authentication and authorization in fermi cloud*. in Proceedings of the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing. 2014. IEEE Computer Society.
- [73] Nzanywayingoma, F, Q Huang. *Securable personal health records using ciphertext policy attribute based encryption*. in E-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference on. 2012. IEEE.
- [74] Popa, RA, et al. *Enabling Security in Cloud Storage SLAs with CloudProof*. in USENIX Annual Technical Conference. 2011.
- [75] Garfinkel, T, et al. Terra: A virtual machine-based platform for trusted computing. in *ACM SIGOPS Operating Systems Review*. 2003. ACM.
- [76] Santos, N, KP Gummadi, R Rodrigues. *Towards Trusted Cloud Computing*. HotCloud, 2009; 9(9): p. 3.
- [77] Zhu, S, G Gong. Fuzzy authorization for cloud storage. *IEEE Transactions on Cloud Computing*, 2014. 2(4): p. 422-435.
- [78] Kim, KH, A Beloglazov, R Buyya. *Power-aware provisioning of cloud resources for real-time services*. in Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and e-Science. 2009. ACM.
- [79] Pillai, P, KG Shin. Real-time dynamic voltage scaling for low-power embedded operating systems. in *ACM SIGOPS Operating Systems Review*. 2001. ACM.
- [80] Martin, SM, et al. *Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads*. in Proceedings of the 2002 IEEE/ACM international conference on Computer-aided design. 2002. ACM.