IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Resource Provisioning Using Agent Based Cloud Computing

[1]Seema Chowhan

[1]Assistant Professor
[1]Department of Computer Science,
[1] Baburaoji Gholap College, Sangvi
Pune,Maharashtra,, 411027, India

Abstract: The emergence of cloud computing has revolutionized the Information and Communication Technology landscape, offering a cost-effective and scalable solution for the on-demand provision of computing resources. Organizations are increasingly drawn to the cloud for outsourcing their computational requirements, benefiting from minimal upfront investment. Central to this paradigm are Service Level Agreements (SLAs), which ensure Quality of Service (QoS) between cloud providers and users. The breach of SLAs can result in contractual penalties and revenue loss, underscoring the criticality of effective SLA management for cloud providers. This involves continuous monitoring and control of various SLA parameters to prevent violations. To address this challenge, there is a growing need for proactive measures to anticipate and mitigate potential breaches by predicting SLA parameter values based on continuous monitoring data. In this research, we propose an agent-based SLA management framework featuring a coordinator agent designed to employ predictive modeling techniques for pre-emptively identifying and addressing SLA violations. Our approach is substantiated by a case study utilizing real-world datasets, demonstrating the efficacy of predictive modeling in enhancing SLA management within cloud environments.

Index Terms - Service Level Agreement (SLA), Cloud computing, Agent based Computing. Multi agent system (MAS).

I. INTRODUCTION

The tremendous growth in communication and information technology along with different characteristics like scalability, reliability and cost-effectiveness have introduced cloud computing as a service over the internet. Physical bonds between user and IT infrastructures have been separated through virtualization technique of cloud computing.

The adoption of cloud computing is a challenge that organization have to face due to the following reasons:

(a) The volume of data generated by software applications, power requirements for analyzing this data, and the performance requirement constraints of real- time processing. (b) The organizational infrastructure and software requirement.

(c) The corporate governance issues regarding the use of
Cloud computing.

(d) Data, applications, and other resources that are outsourced to the cloud may pose security and privacy concerns.

Cloud is a huge pool of computational resources infrastructure that are accessible as a service over internet. This service is provided through data centers and the software that runs them. Service level agreement includes service level agreement parameters, in order to maintain quality of service and specifications, service level parameters are important. For cloud service providers, short term goal is to prevent SLA violations as much as possible to ensure customer satisfaction and avoid penalties. SLA parameters should be continuously measured and monitored so that possible violations can be predicted as soon as they arise. However, accurate prediction of quality of cloud services or SLA violation is extremely challenging because QoS of a cloud service vary drastically at small timescales, due to cloud platform loads, network traffic conditions and other factors. In order to efficiently utilize these resources and ensure that they are available all the times resource provisioning is done based on certain criteria specified in in SLA.

Complex distributed systems and applications can be modeled and implemented using Multi-agent systems (MAS).

Several researchers have proposed Agent based approaches to managing cloud services including SLA management [He et al. 2007]. An agent is autonomous soft- ware that is designed to meet specific goals by co-operating, synchronising and interacting with other agents [Wooldridge 2009]. A simple agent based approach to SLA management at cloud provider level is presented in this paper through interaction and co-ordination between three agents:

Negotiator, Coordinator and Allocator. Coordinator agent goals and functionality with environment is detailed out.

The paper is organized as follows. Next section presents the background and related work. Section three presents proposed Agent based architecture for SLA-management. The paper ends with conclusion and future directions.

II. BACKGROUND AND RELATED WORK

CLOUD COMPUTING

Cloud computing is large scale parallel and distributed computing paradigm. It is driven by economies of scale, in which Virtualization, dynamic scalability, managed computing power, storage, platform are delivered as service from huge resource pool to application over the Internet. In traditional infrastructure model, one has to purchase and deploy physical equipment that takes time to purchase, ship, assemble, configure, install, test and then utilize. The applications that required cloud services, expect quality of service (QoS) guarantees, in accordance with service level agreements. Auto-scaling mechanism in cloud computing is responsible for assuring QoS properties to the applications and making efficient use of resources by keeping operational cost low for service providers.

SLA MANAGEMENT IN CLOUD ENVIRONMENT

Cloud provider has to face problem of unpredicted request load on limited hardware, hence it creates an illusion of availability of unlimited computing resources to the end users. The challenges for cloud computing provider is to allocate

Resources as per Service Level Agreement (SLA) and performance of cloud system should be stable in any dynamic changes of workload as per SLA specified without effecting QoS. QoS involves establishing certain parameters for specific measures of service. Byun et al. [2011], proposed a method in order to minimize the cost from user view and to maximize the resource utilization from cloud providers view. SLAs identify the resources requirement and quality levels required for the execution of job. In such kind of systems availability, reliability, response time and throughput are Quality of service parameters in contractual documents agreed between cloud provider and customer called SLA [Byun et al., 2011], [Wu et al., 2011], [John et al., 2012], [Alkhamees,2021] [Mann et al., 2022] [Qazi et al,,2024] . SLA between Cloud providers and Cloud customer are used to assure QoS that is one of the big issues that resists organization from availing cloud resources. SLA violations may lead to contractual penalties (in turn loss of revenue and business relationship). QoS parameters play an important role in ranking service providers. The uncertainty about the quality of service (QoS) of cloud services reduces user confidence in the technology therefore QoS parameters are continuously monitored and controlled by service providers to avoid SLA violations.

MULTI-AGENT SYSTEM

A multi-agent system constitutes a distributed framework where numerous software autonomous agents collaborate to address a given task or problem. The dynamic nature of cloud computing, necessitating continual monitoring of requests and resources, adaptation to ever-changing requirements, schedules, and prices, as well as the selection of suitable services and plans to fulfill overall cloud objectives, indicates the suitability of employing autonomous agents for cloud management. An agent, as defined by Jennings (N.R), is an intelligent autonomous software entity capable of executing tasks on behalf of users. It exhibits proactive behavior in response to environmental changes and engages in interactions with other agents, steadfastly pursuing its objectives. Within a multi-agent system, a collection of agents collaborates to tackle complex problems utilizing the resources and knowledge base of each agent. Agents must possess various interaction abilities, including negotiation, coordination, cooperation, and teamwork, to effectively engage with other agents. Contemporary agent-based applications frequently entail multiple agents to maintain complex distributed applications. Intelligent agents assume roles on behalf of users, aiding designers and developers in modeling the multi-agent system (MAS) effectively [Cabri, 2004], [Belgacem, 2022].

III. MULTI-AGENT BASED ARCHITECTURE FOR SLA-MANAGEMEN

The appropriateness of utilizing an agent-based approach to address the dynamic nature of the cloud environment is well-recognized and has been explored by numerous researchers [Sim, 2012]; [He, 2007]; [Chen et al., 2014].Saif et al. [2023] proposes an efficient multi-agent autonomic resource provisioning framework for executing business processes in a containerized multi-cloud environment with guaranteed Quality of Service (QoS). Workloads are initially clustered into CPU and I/O intensive categories using an initialization algorithm and K-means clustering.

Mokni et al.[2022], present the implementation of the MAS-GA (Multi-Agent System based Genetic Algorithm) tailored for efficiently schedule a series of interdependent IoT tasks structured as a workflow. A key strength of method lies in its capability to conceptualize IoT workflow planning as a multi-objective optimization challenge, aiming to generate an optimal planning solution that minimize response time, cost considerations, and makespan efficiency.

In this study, we focus on the challenge of Service Level Agreement (SLA) management at the cloud provider level, aiming to prevent SLA violations through continuous data collection. The design and specification of the agent-based architecture are guided by the process and principles outlined in the Prometheus methodology [Padgham and Winiko, 2005]. Individual agents within this architecture demonstrate autonomy and possess critical insights regarding their goals.

A. IDENTIFYING THE SYSTEM GOALS

The main goal of the system is optimum resource utilization and avoiding SLA-violations.

These goals can be divided into several sub goals as given below

(1) Finding appropriate SLA template as per consumer needs.
(2) Initiating the resource provisioning based on SLA template.
(3) Resource allocation.
(4) Record resource utilization parameters.
(5) Forecasting SLA-parameters using appropriate prediction model and resource utilization parameters.
(6) Predicting possibility of SLA violation.
(7) Creating prediction models.
(8) Choosing model depending on SLA parameter and the task.
(9) Changing or refining SLA.
(10) Requesting changes in Resource allocation.
(11) Add, delete, Modify SLA templates.

B. IDENTIFYING DIFFERENT AGENT TYPES

In the intricate system of resource provisioning, numerous sub-goals must be allocated among different types of agents. Three distinct types of agents—SLA Negotiator Agent, SLA Coordinator Agent, and Resource Allocator Agent—are employed to categorize these sub-goals. The SLA parameters encompass both the provider's predefined parameters and the customer-specified Quality of Service (QoS) parameters. A knowledge base repository stores resource information, prediction parameters, prediction models, request and resource mappings, and resource utilization records for various SLA parameters. Figure 1 illustrates the high-level architecture supporting SLA-oriented resource provisioning and allocation in cloud computing. Figure

2 depicts the multi-agent system with its three agents interacting with each other and the cloud infrastructure environment.

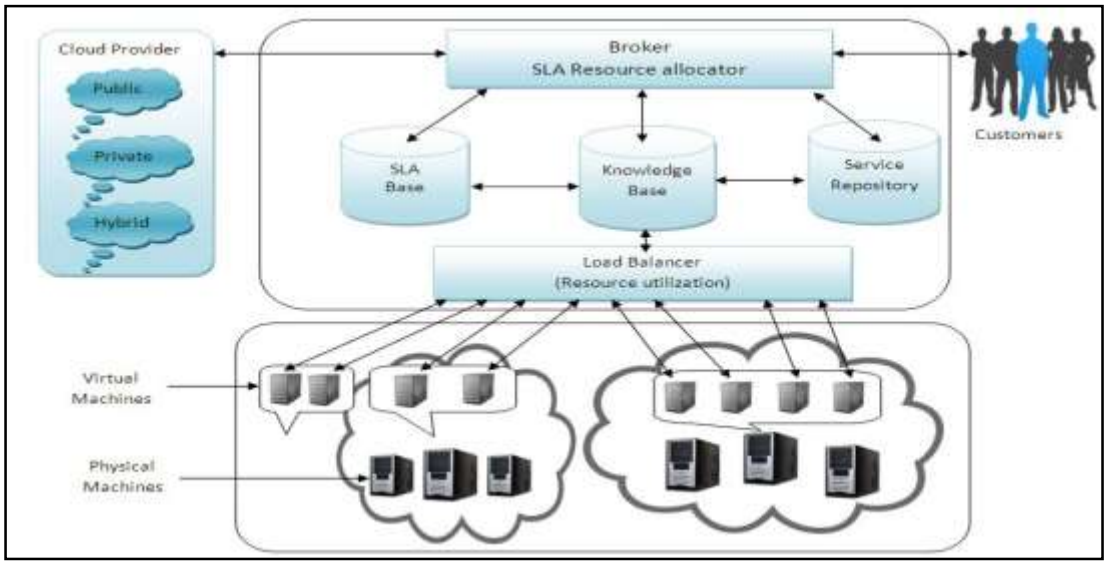

Fig.1 System Architectural frame work for resource Provisioning

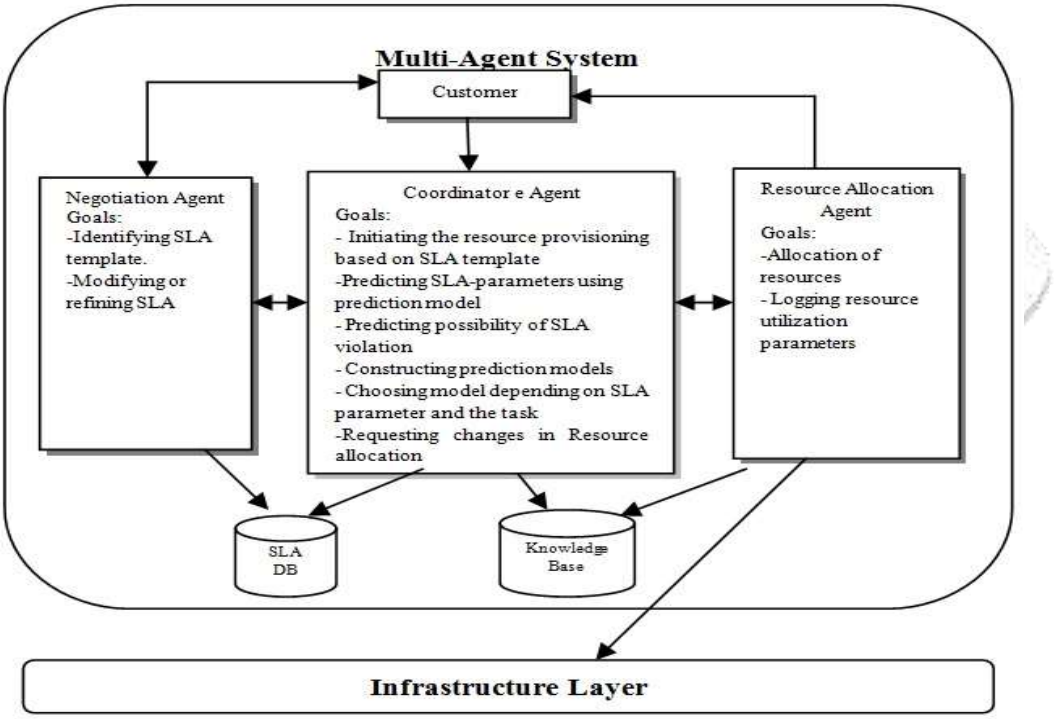

Fig. 2 Multi-agent Architecture for cloud

- **SLA Negotiator Agent:**

SLA database have predefined SLA template, agent selects a predefined SLA template from SLA database as per its requirement. If customer may have specific QoS requirement then negotiator agent will go through a negotiation process with the customer to achieve a mutually agreed SLA (Negotiated SLA). In order to ensure the agreed SLA, the negotiation agent interacts with Coordinator agent to get resource information, collects information about SLA template from SLA database and QoS terms, negotiation strategy and service information from knowledge data base.

- **SLA Coordinator Agent:**

Initiate SLA after receiving a request from the negotiation agent. Collects data about resources and allocates resources by interacting with resource allocator agents. System efficiency is improved by the coordinator agent by mapping user QoS parameters to low-level system requirements. It also uses predictive information collected from the predictive modeling approach to avoid SLA violations by either interacting with the Resource Allocator agent for adjust resources or with the Negotiator agent for adjusting SLA. SLA template

can be added. The Coordinator agent can add, modify, and delete by the coordinator agent thus restricting the range of available templates depending on the resource availability. The coordinator agent achieves the system goals of optimal utilization of system resources. Thus plays a very important role in decision making.

- **Resource Allocator Agent:**

Resource allocator agent does very important task of resource provisioning and allocation. It does resource allocation as per requested by Coordinator agent. It carries out scheduling of resources by collecting data about resources from infrastructure layer and continuously monitors resources to find their availability.

- **SLA Base:**

  Negotiated service level agreements and SLA templates are stored in SLA database.

- **Knowledge Base:**

The providers predefined parameters and the customer specified QoS Parameters are considered as SLA parameters. Cloud resource information, prediction parameters, prediction Model, request and resource mapping, resource utilization records for different SLA parameters are stored in knowledge base repository.

CONCLUSION

Cloud data centers serve as vital hubs for resource provisioning, striving to optimize existing resources before expanding to enhance performance, reliability, and availability. This paper introduces an agent-based approach to construct a multi-agent model aimed at efficiently managing cloud resources. The proposed multi-agent SLA negotiation model, applied at the cloud provider level, seamlessly integrates into broader cloud-level schemes. The coordinating agent's decision-making capability can be bolstered through predictive analytics, tailored to specific SLA parameters. Additionally, the exploration of various critical SLA parameters and effective predictive algorithms is imperative. Future plans involve implementing the model by selecting an appropriate agent-based development platform.

REFERENCES

Byun, E.-K., Kee, Y.-S., Kim, J.-S., and Maeng, S. 2011. Cost-optimized provisioning of elastic resources for application workflows. Future Generation Computer Systems for software as a service provider (saas) in cloud computing environments Proceedings of the 2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. IEEE Computer Society, 195-204.

Chen, J., Han, X., and Jiang, G. 2014. A negotiation model based on the multiagent system under cloud computing.In The Ninth International Multi-Conference on Computing in the Global Information Technology, pp. 157-164.

C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.

Cabri, G., Ferrari, L., Leopardi, L.: Role-based Approaches for Agent Development. In: Proceedings of the 3rd Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), pp. 1504–1505 (2004)

He, Q., Yan, J., Kowalczyk, R., Jin, H., and Yang, Y. 2007. An agent-based framework for service level agreement management. In 2007 11th International Conference on Computer-Supported Cooperative Work in

Jennings, N.R., Wooldridge, M.: Agent Technology: Foundations, Applications, and Markets, pp. 29–40. Springer (2002)

Jowda, F., & Al-Asfoor, M. (2022). A survey on cloud resources allocation using multi-agent system. In *Expert Clouds and Applications: Proceedings of ICOECA 2021* (pp. 139-151). Springer Singapore.

M. Young, the Technical Writers Handbook. Mill Valley, CA: University Science, 1989. A practical guide to cloud service level agreements.

Mann, S., Pathak, N., Sharma, N., Kumar, R., Porwal, R., Sharma, S. K., & Aung, S. M. Y. (2022). Study of Energy-Efficient Optimization Techniques for High-Level Homogeneous Resource Management. Wireless Communications and Mobile Computing, 2022.

Mokni, M., Yassa, S., Hajlaoui, J. E., Chelouah, R., & Omri, M. N. (2022). Cooperative agents-based approach for workflow scheduling on fog-cloud computing. Journal of Ambient Intelligence and Humanized Computing, 13(10), 4719-4738.

Padgham, L. and Winikoff, M. 2005. Developing intelligent agent systems: A practical guide, Volume 13. John Wiley & Sons.

Qazi, F., Kwak, D., Khan, F. G., Ali, F., & Khan, S. U. (2024). Service Level Agreement in cloud computing: Taxonomy, prospects, and challenges. Internet of Things, 101126.

Saif, M. A. N., Niranjan, S. K., Murshed, B. A. H., Al-ariki, H. D. E., & Abdulwahab, H. M. (2023). Multi-agent QoS-aware autonomic resource provisioning framework for elastic BPM in a containerized multi-cloud environment. Journal of Ambient Intelligence and Humanized Computing, 14(9), 12895-12920.

Sim, M. K. 2012. Agent-based cloud computing. IEEE Transactions on Services Computing Vol.5, 4.

Talia D. Cloud computing and software agents: Towards cloud intelligent services. In: Fortino G, Garro A, Palopoli L, Russo W, Spezzano G, editors. Proceedings of the 12th

Wooldridge, M. 2009. An introduction to multiagent systems. John Wiley & Sons. Workshop on Objects and Agents (WOA-2011). Italy: Rende (CS); July 4-6, 2011

Wu, L., Garg, S. K.,and Buyya, R. 2011. Sla-based resource allocation

Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces(Translation Journals style)," IEEE Trans. J. Magn.Jpn., vol. 2, Aug. 1987, pp. 740–741 [Dig. 9th Annu. Conf. Magnetics Japan, 1982, p. 301].

AUTHORS PROFILE

Dr. Seema Chowhan is working as a faculty and head in the subject of computer science at Baburaoji Gholap College Pune, India affiliated to Savitribai Phule Pune University, Pune. She has 24+ years of experience in teaching UG and PG courses. She has completed an M.Phil and Ph.D in computer science. Her research interests include Cloud Computing, Networking, and Machine learning.