



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 4, April 2026

Comprehensive Analysis of Cloud Computing and Cost Optimization Architectures

Keyur Girishbhai Solanki, Dr. Sneha Patel

Student, Bhagwan Mahavir College of Computer Application, BMU, Surat, India

Assistant Professor, Bhagwan Mahavir College of Computer Application, BMU, Surat, India

ABSTRACT: Cloud computing has transformed modern information technology by enabling scalable, elastic, and cost-efficient computing services delivered over the Internet. Since its formalization by the National Institute of Standards and Technology (NIST), cloud computing has evolved beyond traditional service models to incorporate containerization, serverless computing, artificial intelligence integration, and FinOps-driven cost governance. Building upon foundational studies including this paper presents a consolidated analysis of cloud service models, deployment architectures, security challenges, cost optimization strategies, and emerging paradigms such as hybrid cloud and Machine learning operations (MLOps). The study synthesizes over twenty-five research works to provide a structured and comprehensive understanding of the cloud ecosystem and its future direction.

KEYWORDS: Cloud Cloud, Computing, Cost Hybrid IaaS, MLOps, Optimization, PaaS, SaaS, Security, Serverless, SNS,

I. INTRODUCTION

Cloud computing has become one of the most important technologies in modern information systems. Over the past decade, organizations have gradually moved away from traditional on-premise infrastructure and started adopting cloud platforms to host applications, manage data, and deliver digital services. The main reason for this shift is the flexibility and scalability that cloud environments provide.

Despite the advantages, cloud environments also introduce several challenges. Organizations must address issues related to security, system reliability, and cost management. Since cloud resources operate in dynamic environments, sudden spikes in system workload or inefficient resource allocation can lead to performance degradation or unnecessary operational expenses. Therefore, continuous monitoring and proper resource management have become essential components of modern cloud infrastructure.

Another important concern in cloud environments is system failure or service disruption. If abnormal system behavior is not detected early, it may lead to application downtime and affect the overall availability of services. For this reason, monitoring tools and automated alert systems are commonly used to track system performance metrics and notify administrators whenever unusual activity occurs.

In order to understand how monitoring systems work in real cloud environments, this study also includes a practical implementation using Amazon Web Services (AWS). A monitoring and alerting system was implemented using services such as Amazon EC2, Amazon CloudWatch, and Amazon SNS. In this setup, CloudWatch continuously monitors the CPU utilization of an EC2 instance and triggers an alert whenever the CPU usage exceeds a predefined threshold. Once the threshold is crossed, the system automatically sends an email notification to the administrator using Amazon SNS.

This practical implementation demonstrates how cloud monitoring systems can help detect abnormal system behavior early and reduce the risk of service interruptions.

The purpose of this study is to provide a comprehensive analysis of cloud computing services, security challenges, cost optimization strategies, and emerging architectures. The paper also presents a practical case study that demonstrates how cloud monitoring and alerting mechanisms can be implemented using modern cloud platforms.

II. CLOUD FUNDAMENTALS

2.1 Cloud Service Models

Cloud services are typically categorized into three primary models:

2.1.1 Infrastructure as a Service (IaaS)

IaaS provides virtualized computing resources such as servers, storage, and networking. Prodan and Ostermann [14] conducted one of the earliest taxonomies of IaaS providers, highlighting elasticity and pay-per-use models. Virtualization plays a central role in enabling isolation and scalability [19].

2.2 Platform as a Service (PaaS)

PaaS allows developers to build and deploy applications without managing infrastructure. According to Goncalves and Ballon [18], PaaS enhances software lifecycle control and supports collaborative development environments.

2.3 Software as a Service (SaaS)

SaaS delivers complete applications via web interfaces. Youssef [1] highlights its configurability and multi-tenant efficiency. However, multi-tenancy introduces security concerns that require robust isolation mechanisms [6]. Subashini and Kavitha [6] emphasize vulnerabilities across service delivery models. Zissis and Lekkas [7] identify data confidentiality and trust management as major concerns. ENISA’s risk assessment report [10] further classifies operational and compliance risks.

III. CLOUD ECONOMICS AND COST OPTIMIZATION

One of the defining advantages of cloud computing is economic efficiency. Buyya et al. [12] describe the cloud as a utility computing model similar to electricity services.

Cost Benefits

Organizations reduce capital expenditure (CapEx) by adopting operational expenditure (OpEx) models [1].

. The World’s Top 3 Cloud Providers: AWS, GCP, and Azure[26]

Provider	Market Share	Strengths	Cost Tools
AWS (Amazon Web Services)	~32%	Largest service portfolio, global reach, mature ecosystem.	AWS Cost Explorer, Budgets, Savings Plans, Reserved Instances, Trusted Advisor, Compute Optimizer.
Google Cloud Platform (GCP)	~11%	Data analytics, machine learning, seamless integration with Google Workspace.	GCP Billing, Cost Management Reports, Committed Use Discounts, Sustained Use Discounts, Active Assist Recommendations.
Microsoft Azure	~23%	Enterprise support, hybrid cloud, integration with Microsoft 365.	Azure Cost Management + Billing, Reservations, Spot VMs, Advisor, Pricing Calculator.

Proposed AWS Monitoring and Alerting System

This paper proposes a simple and effective cloud-based monitoring and alerting model designed to improve system reliability and performance management in cloud environments. The model focuses on continuously monitoring system resources and providing real-time alerts whenever abnormal behavior is detected.

In this project, a basic cloud monitoring and alerting system was implemented using AWS services. The main goal of the system was to monitor the CPU utilization of an EC2 instance and generate an alert whenever the CPU usage crosses a defined threshold.

The monitoring was performed using Amazon CloudWatch, which continuously collects performance metrics from the EC2 instance. A CloudWatch alarm was configured so that whenever CPU utilization exceeds 70%, the alarm is triggered. Once the alarm is triggered, it sends a notification through Amazon SNS to the subscribed email address.

3.1 Why Monitoring is Important in Cloud Systems

In cloud environments, servers and applications run continuously and handle different workloads. Sometimes the system can experience high resource usage due to heavy traffic, misconfiguration, or unexpected workloads. If these issues are not detected early, they can lead to performance degradation or even system downtime. Monitoring systems help track important metrics such as CPU utilization, network activity, and memory usage.

By monitoring these metrics in real time, administrators can quickly detect abnormal behavior and take preventive action before the situation becomes critical.

3.2 Cost Optimization Using Monitoring

In many cloud environments, organizations sometimes allocate more resources than necessary, which leads to unnecessary cloud costs. By analyzing monitoring data from services like CloudWatch, administrators can understand how resources are actually being used. For example, if the CPU usage of a server remains very low most of the time, it indicates that the instance size may be larger than required. In such cases, the instance type can be downgraded to reduce costs.

Similarly, monitoring helps identify sudden spikes in usage. These insights help organizations make better decisions about scaling resources and managing cloud expenses efficiently.

3.3 Failure Prevention and Early Detection

Another important benefit of monitoring is failure prevention. Instead of waiting for a system to fail, monitoring systems help detect warning signs early.

In this project, the CloudWatch alarm is configured to trigger when CPU utilization exceeds 70%. When this threshold is crossed, an alert email is immediately sent to the administrator.

This early warning allows the administrator to investigate the issue quickly. The problem could be caused by high traffic, inefficient processes, or resource limitations. Once the issue is identified, corrective actions such as restarting services, optimizing workloads, or scaling resources can be taken.

Because of this early detection mechanism, the system helps reduce the chances of service interruption.

3.4 Architecture of the Monitoring System

The monitoring system follows a simple but effective architecture.

First, the EC2 instance generates performance metrics such as CPU utilization. These metrics are automatically collected by Amazon CloudWatch.

CloudWatch continuously evaluates these metrics based on the conditions defined in the alarm configuration. If the CPU utilization exceeds the defined threshold, the CloudWatch alarm changes its state to "ALARM".

Once the alarm is triggered, it sends a notification to Amazon SNS. The SNS service then delivers an email notification to the administrator.

This architecture ensures that system health is monitored continuously and alerts are delivered immediately when abnormal conditions occur.

IV. IMPLEMENTATION AND TESTING

To better understand how monitoring works in a cloud environment, a small monitoring and alerting system was implemented using AWS services. The main purpose of this setup was to monitor the CPU utilization of an EC2 instance and generate an alert whenever the CPU usage crosses a predefined threshold.

In this implementation, Amazon CloudWatch was used to monitor the CPU utilization of the EC2 instance. A CloudWatch alarm was configured to trigger when CPU usage goes above 70%. Once the alarm is triggered, a notification is sent using Amazon SNS to the administrator’s email. The architecture of the monitoring system used in this project is shown in the figure below.

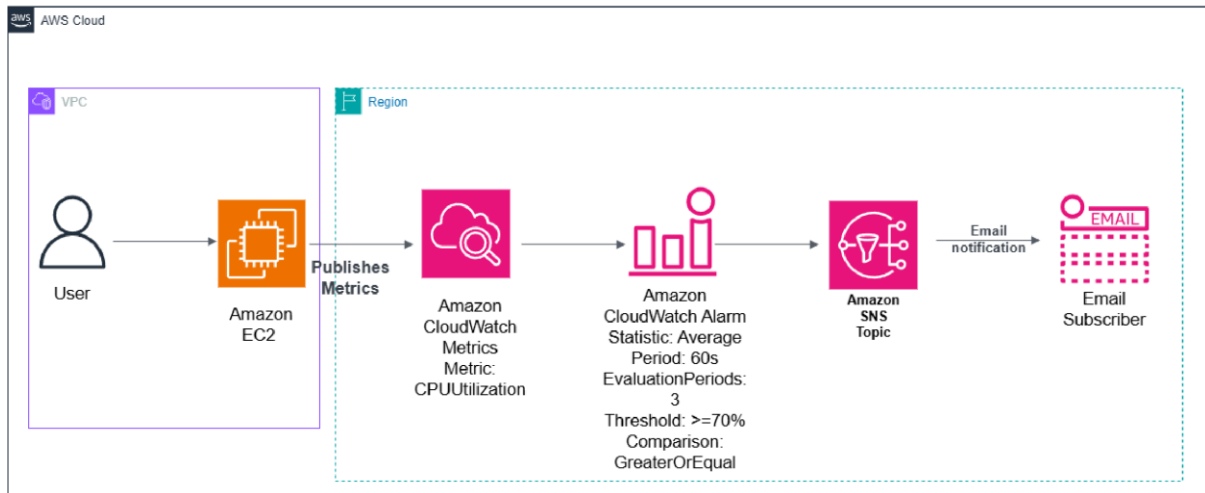


Figure 1: Architecture of the AWS monitoring and alerting system

As shown in the architecture diagram, the EC2 instance sends performance metrics to Amazon CloudWatch. CloudWatch continuously analyzes these metrics and checks whether the defined threshold conditions are met. If the CPU utilization exceeds the configured limit, the CloudWatch alarm is triggered. This alarm then sends a notification to Amazon SNS, which delivers an email alert to the administrator.

After setting up the monitoring architecture, the system was tested to verify whether the alert mechanism works properly. First, an EC2 instance was launched and configured to send performance metrics to CloudWatch.

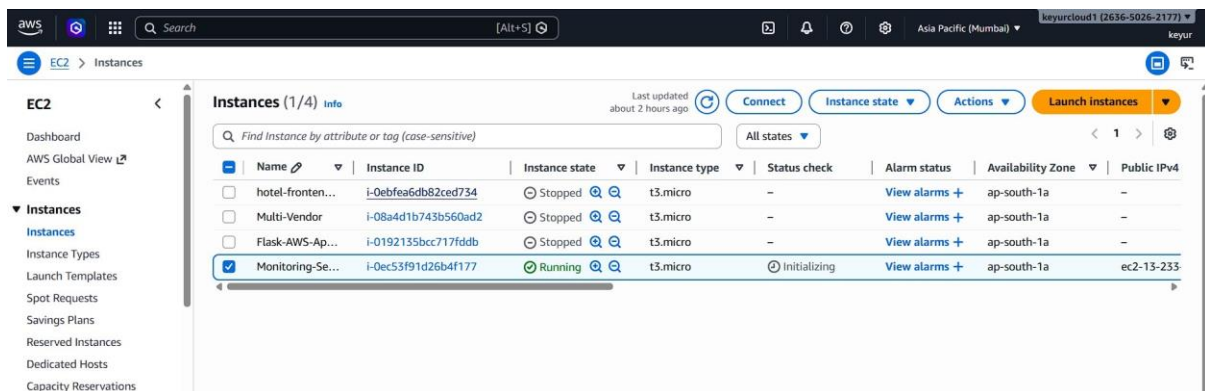


Figure 2: EC2 instance running in AWS environment

Once the instance was running, a CloudWatch alarm was created to monitor CPU utilization. The alarm threshold was set to 70%, which means the alarm would trigger whenever the CPU usage crosses this limit.

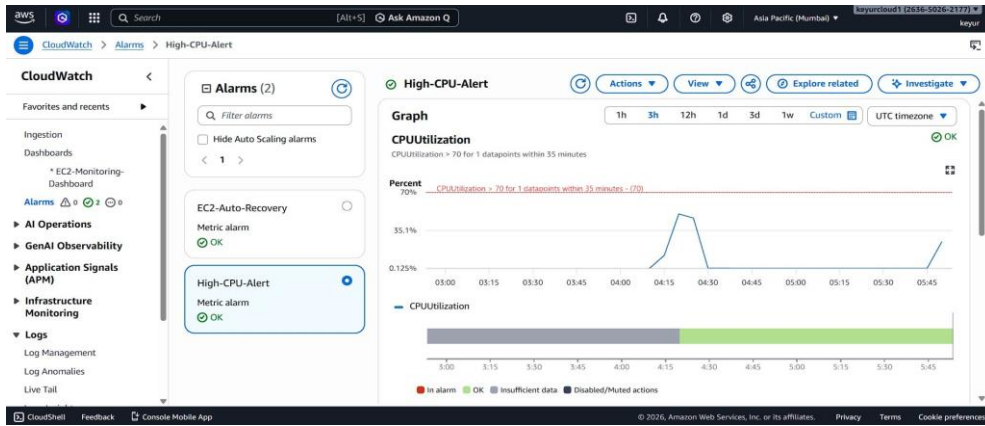


Figure 3: CloudWatch alarm configured to monitor CPU utilization

To test the monitoring system, a CPU intensive process was executed on the EC2 instance. This was done intentionally to increase CPU utilization and trigger the alarm condition. After some time, the CPU utilization increased and crossed the defined threshold. And the CloudWatch alarm changed its state from OK to ALARM.

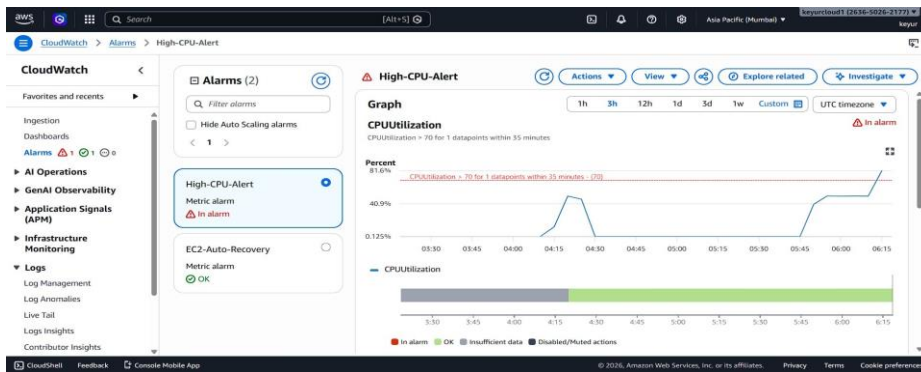


Figure 4: CloudWatch alarm triggered after CPU utilization exceeded the threshold

Once the alarm was triggered, Amazon SNS automatically sent an email notification to the subscribed email address. This confirmed that the monitoring and alerting system was working correctly.

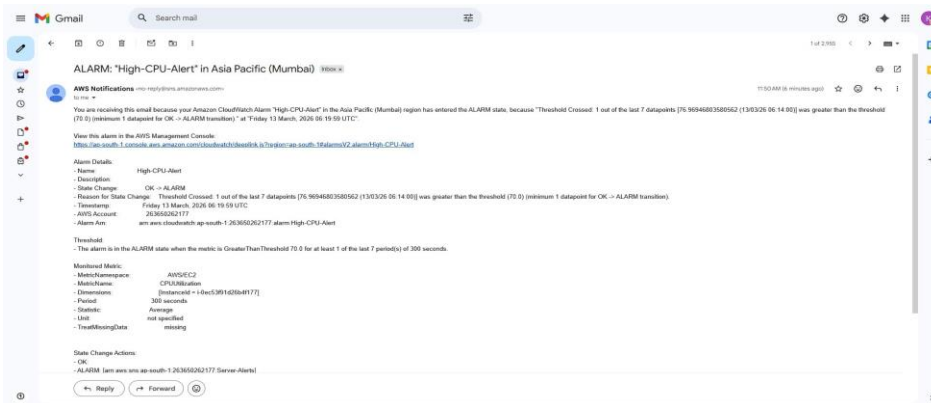


Figure 5: Email notification received through Amazon SNS

In addition to the alert system, a CloudWatch dashboard was also created to visualize the monitoring data in real time. The dashboard provides a graphical view of CPU utilization and helps administrators quickly understand system performance.

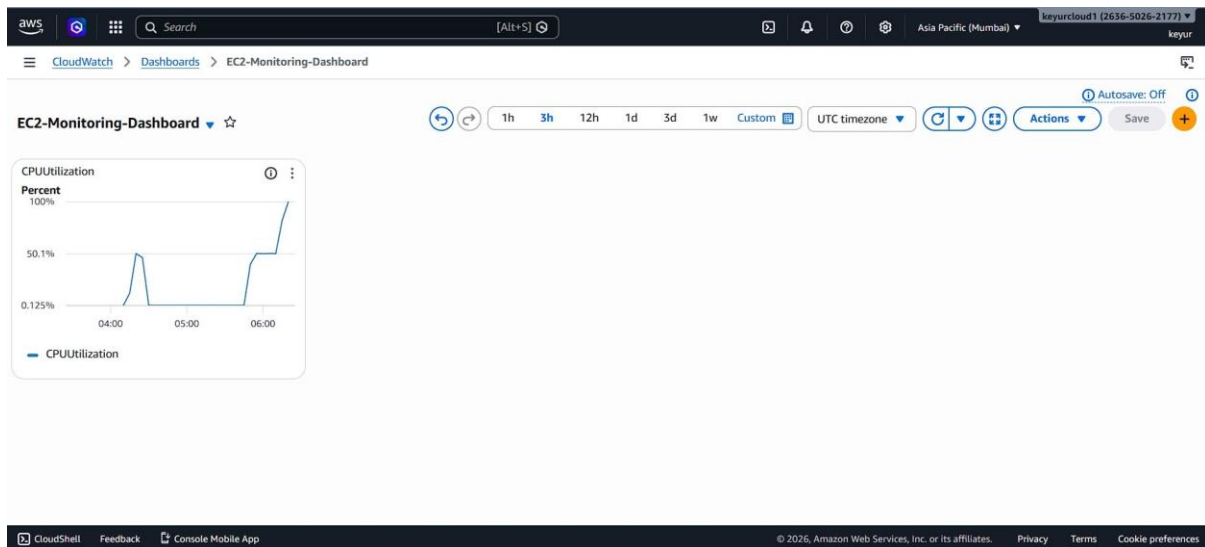


Figure 6: CloudWatch dashboard displaying CPU utilization metrics

V. CONCLUSION

The implementation demonstrated how cloud monitoring tools can help administrators detect abnormal resource usage early. By continuously tracking system metrics and generating alerts when certain conditions are met, monitoring systems help reduce the risk of service interruption.

Another advantage of monitoring systems is that they provide visibility into resource usage patterns. This information can be useful for optimizing resource allocation and managing cloud costs more efficiently.

Overall, this implementation shows that even a simple monitoring setup can significantly improve the reliability and management of cloud infrastructure.

Cloud computing has evolved from a service-oriented abstraction model into a sophisticated digital ecosystem supporting enterprise transformation, artificial intelligence, and global connectivity. While security and governance challenges persist, technological advancements and regulatory frameworks are strengthening trust in cloud platforms.

This research paper examined various aspects of cloud computing, including service models, deployment architectures, security challenges, cost optimization strategies, and emerging technological trends. The study highlights how cloud environments provide organizations with improved scalability and operational flexibility while also introducing new challenges related to security, governance, and resource management.

To demonstrate the practical importance of monitoring in cloud infrastructure, a monitoring and alerting system was implemented using Amazon Web Services. The system utilized services such as Amazon EC2, Amazon CloudWatch, and Amazon SNS to monitor CPU utilization and generate alerts when system thresholds were exceeded. Through this implementation, it was observed that monitoring systems can effectively detect abnormal resource utilization and notify administrators in real time, allowing them to respond quickly to potential issues.

REFERENCES

- [1] A. E. Youssef, "Exploring Cloud Computing Services and Applications," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 3, no. 6, pp. 838–847, 2012.
- [2] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," NIST Special Publication 800-145, 2011. <https://doi.org/10.6028/NIST.SP.800-145>
- [3] M. Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," *UC Berkeley Technical Report*, 2009.
- [4] M. Armbrust et al., "A View of Cloud Computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010. <https://doi.org/10.1145/1721654.1721672>
- [5] T. Dillon, C. Wu, and E. Chang, "Cloud Computing: Issues and Challenges," *24th IEEE International Conference on Advanced Information Networking and Applications*, 2010. <https://doi.org/10.1109/AINA.2010.187>
- [6] S. Subashini and V. Kavitha, "A Survey on Security Issues in Service Delivery Models of Cloud Computing," *Journal of Network and Computer Applications*, vol. 34, pp. 1–11, 2011. <https://doi.org/10.1016/j.jnca.2010.07.006>
- [7] D. Zisis and D. Lekkas, "Addressing Cloud Computing Security Issues," *Future Generation Computer Systems*, vol. 28, pp. 583–592, 2012. <https://doi.org/10.1016/j.future.2010.12.006>
- [8] K. Popovic and Z. Hocenski, "Cloud Computing Security Issues and Challenges," *MIPRO Conference*, 2010.
- [9] Cloud Security Alliance, "Top Threats to Cloud Computing," CSA Report, 2010.
- [10] ENISA, "Cloud Computing Risk Assessment," European Union Agency for Cybersecurity, 2009. <https://doi.org/10.2824/93560>
- [11] W. Jansen and T. Grance, "Guidelines on Security and Privacy in Public Cloud Computing," NIST SP800-144, 2011.
- [12] R. Buyya, C. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality," *Future Generation Computer Systems*, vol. 25, pp. 599–616, 2009. <https://doi.org/10.1016/j.future.2008.12.001>
- [13] L. M. Vaquero et al., "A Break in the Clouds: Towards a Cloud Definition," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 1, pp. 50–55, 2009.
- [14] R. Prodan and S. Ostermann, "A Survey and Taxonomy of Infrastructure as a Service and Web Hosting Cloud Providers," *IEEE Grid Computing Conference*, 2009.
- [15] Y. Chen, X. Li, and F. Chen, "Overview and Analysis of Cloud Computing Research and Application," *International Conference on E-Business and E-Government*, 2011.
- [16] A. Fox et al., "Cloud Computing: The State of the Art," *Science*, vol. 328, no. 5978, pp. 592–594, 2010.
- [17] N. Robinson et al., "The Cloud: Understanding the Security, Privacy and Trust Challenges," RAND Corporation, 2010.
- [18] A. Joint and E. Baker, "Knowing the Past to Understand the Present: Issues in Contracting for Cloud-Based Services," *Computer Law & Security Review*, vol. 27, pp. 407–415, 2011.
- [19] V. Sarathy, P. Narayan, and R. Mikkilineni, "Next Generation Cloud Computing Architecture," *IEEE Workshop on Collaboration and Cloud Computing*, 2010.
- [20] R. Choubey, R. Dubey, and J. Bhattacharjee, "A Survey on Cloud Computing Security, Challenges and Threats," *International Journal on Computer Science and Engineering*, vol. 3, no. 3, 2011.
- [21] M. Zhang et al., "Cloud Computing: State-of-the-Art and Research Challenges," *Journal of Internet Services and Applications*, 2010.
- [22] J. Bordkin, "Gartner: Seven Cloud-Computing Security Risks," Gartner Report, 2008.
- [23] A. Alshwaiir, A. Youssef, and A. Emam, "A New Trend for E-Learning in KSA Using Educational Cloud," *Advanced Computing: An International Journal*, 2012.
- [24] R. L. Krutz and R. D. Vines, *Cloud Security: A Comprehensive Guide to Secure Cloud Computing*, Wiley Publishing, 2010.
- [25] Amazon Web Services, "Overview of Amazon Web Services," AWS Whitepaper, 2020.
- [26] <https://www.visualcapitalist.com/the-worlds-largest-cloud-providers-ranked-by-market-share/>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details