iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IX **Month of publication:** September 2026

DOI: https://doi.org/10.22214/ijraset.2026.84799

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IX Sep 2026- Available at www.ijraset.com

Sustainable Cloud Computing in the AI Era: Energy-Aware Infrastructure, Edge Integration and Security Challenges

Fayiza Rinaz K M[1], Ayesha Thahleela K K[2], Aaslim Sanofar A[3], Santhiya A[4], Vinshi P[5], G. Abirami[6]
[1, 2, 3, 4]*U.G. Student, Department of Bachelor of Computer Applications, Annai Womens College, Karur, TamilNadu, India,*
[5, 6]*Assistant Professor, Department of BCA IT, Annai Womens College, Karur ,India*

Abstract: ***Cloud computing has become a part of modern technology. It is used in intelligence, online education, banking, healthcare, entertainment, business applications and many other areas. The growth of intelligence is increasing the demand for computing power, storage, networking and data-centre resources. Because of this growth, energy consumption, security and efficient resource management have become issues in cloud computing. This paper discusses cloud computing in the AI era with a focus on energy-efficient infrastructure, resource management, edge computing and cloud security. It also discusses how artificial intelligence can be used to improve cloud resource management and reduce resource usage. A simple conceptual approach based on workload monitoring, resource allocation, edge processing and security monitoring is presented. The study concludes that future cloud systems need to provide performance while also reducing unnecessary energy consumption and maintaining security and reliability.***

Keywords: Cloud Computing, Sustainable Computing, Artificial Intelligence, Energy Efficiency, Edge Computing, Cloud Security, Data Centres

I. INTRODUCTION

Cloud computing has changed the way computing resources are used. In the traditional approach, organizations need to purchase and maintain their own servers, storage devices and networking equipment. Cloud computing provides another option by allowing users to access computing resources through the Internet according to their requirements.

Cloud computing is now widely used for storing data, running applications, developing software, online learning, business activities, artificial intelligence and many other services.

The growth of artificial intelligence has increased the demand for cloud computing. AI applications require large amounts of computing power and data storage. Modern data centres therefore need processors, specialized hardware, storage systems and cooling facilities.

The International Energy Agency has reported growth in electricity consumption by data centres and expects this demand to continue increasing as AI and other digital services expand.

This situation creates a challenge for the cloud-computing industry. More computing power is required, but unnecessary energy consumption should also be avoided. Cloud providers and organizations therefore need to consider performance, cost, security and energy efficiency.

This paper discusses cloud computing in the AI era. It focuses on energy-aware infrastructure, cloud resource management, edge computing, security and the role of artificial intelligence in improving cloud efficiency.

II. EVOLUTION OF CLOUD COMPUTING

Cloud computing developed from technologies such as distributed computing, virtualization, grid computing and utility computing. Virtualization is one of the technologies behind cloud computing. It allows one physical server to run virtual machines. This helps organizations use hardware resources effectively.

Cloud computing is generally divided into three service models:

- Infrastructure as a Service (IaaS): Provides resources such as virtual machines, storage and networking.
- Platform as a Service (PaaS): Provides an environment for developing and deploying applications.
- Software as a Service (SaaS): Provides ready-to-use software applications through the Internet.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IX Sep 2026- Available at www.ijraset.com

III. CLOUD COMPUTING IN THE AI ERA

Artificial intelligence is changing the requirements of cloud computing. AI applications need large amounts of processing power for activities such as model training, data analysis and inference.

AI workloads can sometimes be more demanding than traditional applications. They may require hardware such as GPUs and other computing accelerators. The growth of AI also creates opportunities for cloud computing. AI can be used to predict workloads, identify resource usage, improve resource allocation and support automated cloud management.

Therefore, there is a relationship between AI and cloud computing. AI increases the demand for resources while AI can also help cloud platforms use those resources more efficiently.

IV. ENERGY CHALLENGE IN MODERN DATA CENTRES

Energy consumption is becoming an important issue in modern cloud computing. Data centres require electricity not only for servers but also for storage, networking, cooling and other supporting systems.

As AI applications become more popular, the demand for computing infrastructure is increasing. The IEA estimates that data-centre electricity consumption will continue to grow through the end of this decade.

Unnecessary energy consumption can occur because of reasons including:

1) Underused servers.
2) Poor workload scheduling.
3) Excessive data movement.
4) Inefficient cooling.
5) Oversized virtual machines.
6) Unnecessary data storage.
7) Inefficient deployment of AI applications.

Improving the use of resources can help reduce energy consumption while maintaining good performance.

V. PROPOSED ENERGY-AWARE CLOUD FRAMEWORK

This paper proposes a conceptual framework for improving energy efficiency in cloud computing. The framework contains five steps.

A. Workload Monitoring

The system continuously observes the workload of resources. Information such as CPU usage, memory usage, network traffic and accelerator usage can be monitored.

B. Workload Prediction

Previous workload information can be used to estimate demand. This can help the cloud system prepare resources when demand is expected to increase.

C. Dynamic Resource Allocation

Resources can be increased or decreased according to the actual requirements of an application. This can help prevent resource over-allocation.

D. Energy-Aware Scheduling

When several computing resources are available, the system can select resources by considering both performance and energy consumption.

E. Continuous Monitoring

The system should continuously monitor performance and resource usage. The collected information can be used to improve resource-allocation decisions.

The basic process can be represented as:

Workload → Monitoring → Prediction → Resource Allocation → Energy-Aware Scheduling → Monitoring → Optimization

This creates a process for improving cloud resource usage.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IX Sep 2026- Available at www.ijraset.com

VI. CLOUD AND EDGE COMPUTING

Cloud computing and edge computing are both important technologies, but they work in different ways. In cloud computing, most data is sent to a central data centre for processing and storage. It is useful for applications that need large storage space and high computing power. Edge computing processes some data closer to where it is produced. This reduces the time needed to send data to a cloud server and receive a response. It is useful for applications that need quick responses, such as smart devices, cameras, vehicles and industrial systems. Cloud computing may require data to travel through the network, while edge computing can process some information locally. This can reduce network traffic and improve response time.

However, managing edge devices can be more difficult because they may be located in different places. Cloud systems are generally easier to manage from a central location.

Both technologies have their advantages. Cloud computing is suitable for large-scale storage, data analysis and AI model training, while edge computing is useful for real-time processing. Using cloud and edge computing together can provide better performance, flexibility and efficient use of resources.

VII. SECURITY CHALLENGES

Security is one of the major concerns in cloud computing. Since data and applications are often stored on infrastructure operated by cloud providers, appropriate security measures are required.

A. Identity and Access Management

Only authorized users should be able to access resources. Users should receive the permissions required for their work.

B. Data Protection

Important data should be protected during storage and transmission. Encryption can be used to reduce the risk of unauthorized access.

C. Cloud Misconfiguration

Incorrect cloud settings can expose applications and data. Regular checking of cloud configurations can help identify problems.

D. Multi-Tenant Security

Cloud providers may serve customers using shared infrastructure. Proper isolation is required to prevent access between different users.

E. AI Security

AI-based cloud systems also require protection for training data, models and AI services. Unauthorized access or modification of AI models can create risks.

Security should therefore be included from the beginning of cloud-system design.

VIII. SUSTAINABILITY AND PERFORMANCE

Energy efficiency should not be achieved by reducing computing performance. A good cloud system should maintain a balance between several factors.

These factors include:

1) Performance
2) Energy consumption
3) Cost
4) Security
5) Availability
6) Scalability

For example, placing many workloads on a small number of servers may reduce energy usage but can create performance problems. Similarly, moving all processing to edge devices may reduce network delay but can make system management more difficult.

Therefore, cloud systems should be designed to achieve a balance between performance and sustainability.

IX. ROLE OF ARTIFICIAL INTELLIGENCE IN CLOUD OPTIMIZATION

Artificial intelligence can be used to improve cloud resource management.

AI-based systems can study workload patterns and predict future demand. This information can help cloud platforms allocate resources before demand increases.

AI can also help identify changes in resource usage. For example, a sudden increase in network traffic or CPU usage could indicate a change in workload or a possible security problem.

Another possible application is data-centre cooling. AI-based systems can analyze temperature and workload information to help improve cooling management.

AI can therefore support cloud providers in improving resource utilization, reducing energy consumption and maintaining reliable services.

X. SUSTAINABLE CLOUD STRATEGY

A sustainable cloud strategy should include practical measures.

1) Resource Right-Sizing: Cloud resources should match the requirements of applications. Using larger resources than necessary can increase cost and energy consumption.
2) Automatic Scaling: Resources can be automatically increased when demand increases and reduced when demand decreases.
3) Workload Consolidation: Suitable workloads can be combined on resources to improve utilization while maintaining performance and security.
4) Efficient Data Management: Unnecessary data copies and excessive data movement should be reduced.
5) Renewable Energy: Cloud providers can increase the use of renewable energy sources for data-centre operations where suitable.
6) Edge Processing: Some data can be processed closer to its source instead of sending everything to a central cloud.
7) Continuous Monitoring: Energy consumption, resource utilization, performance and cost should be monitored regularly.

XI. DISCUSSION

The development of cloud computing shows that scalability and performance are no longer the only important factors. Energy efficiency and sustainability are also becoming important.

The fast rise of AI is creating greater demand for computing power. At the same time, better hardware and software can make individual AI tasks more efficient. Even with these improvements, the overall use of AI can still consume more energy. This means improving each system is not enough. Cloud companies and users also need to manage workloads efficiently and avoid using more resources than necessary.

The framework discussed in this paper includes workload monitoring, prediction, dynamic resource allocation, energy-aware scheduling, edge processing and continuous security monitoring. Future versions of this system can evaluate its effectiveness using measures such as energy consumption, response time, resource utilization, operating cost and availability.

XII. FUTURE SCOPE

Cloud computing will keep growing as AI and edge computing become more common. Future work can focus on developing AI algorithms for predicting cloud workloads.

Cloud scheduling systems can also be designed to consider energy availability, workload importance, cost and performance. Edge-cloud systems can be studied further for applications that require fast results.

Another important area is AI security. Future systems will need methods to protect AI models, training data and cloud-based AI services.

Research on energy-efficient data centres and the use of renewable energy to power cloud systems can also help make future computing more environmentally sustainable.

XIII. CONCLUSION

Cloud computing has become a part of modern digital systems. The growth of artificial intelligence is increasing the need for computing power and data-centre resources, making energy efficiency and sustainability more important.

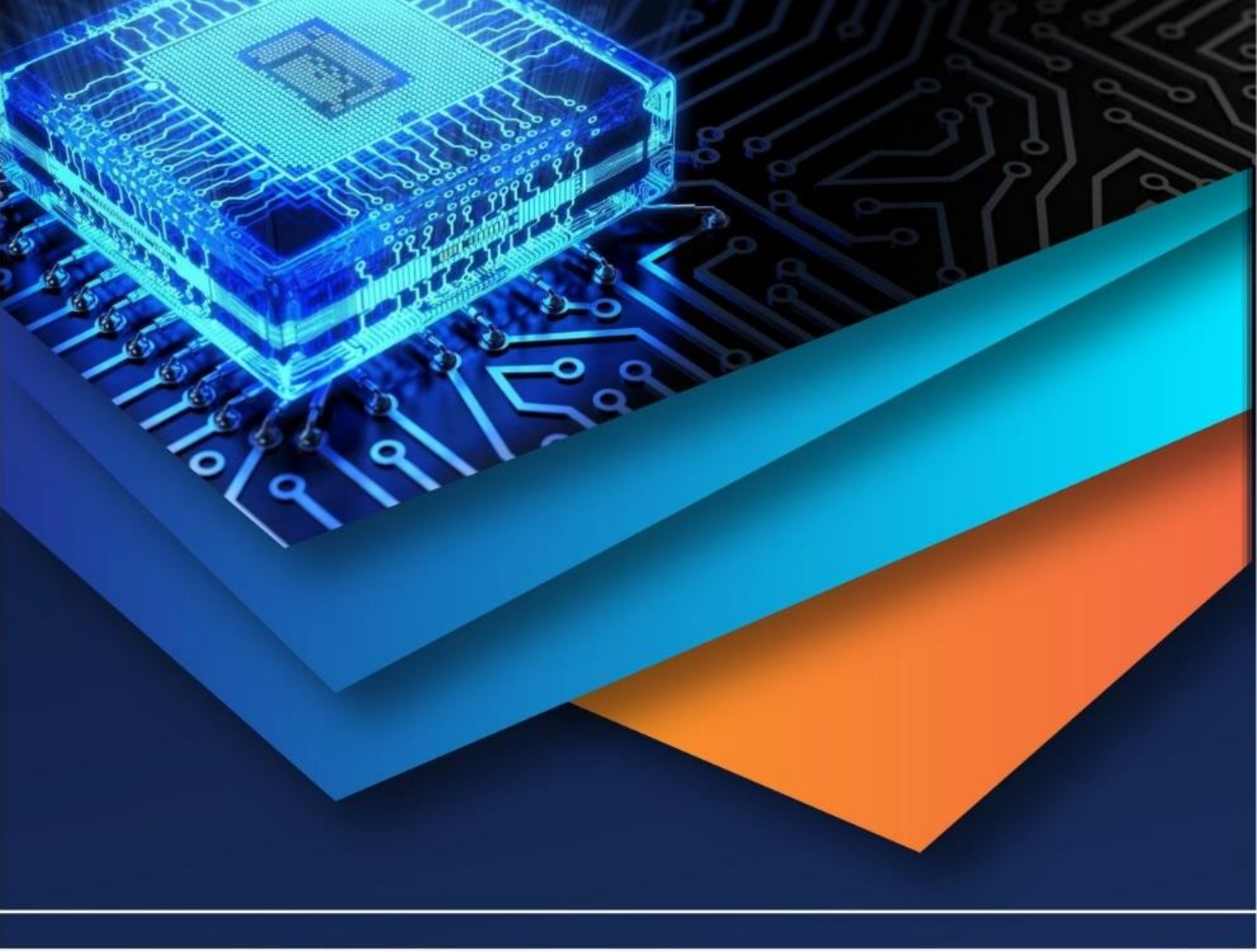
This paper discussed cloud computing with a focus on using energy wisely, edge computing, cloud security and using AI to improve cloud management. The proposed framework uses workload monitoring, workload prediction, dynamic resource allocation, energy-aware scheduling and continuous monitoring to improve resource utilization.

Cloud computing and edge computing can work together. Cloud computing is suitable for storing and processing large amounts of data, while edge computing can handle tasks that need to be completed quickly near users and devices.

Overall, future cloud systems need to be fast while also using resources wisely, maintaining strong security and avoiding unnecessary energy consumption. Sustainable cloud computing can therefore help support the growth of AI and other digital technologies.

REFERENCES

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," NIST Special Publication 800-145, National Institute of Standards and Technology, 2011.
[2] M. Armbrust et al., "A View of Cloud Computing," Communications of the ACM, vol. 53, no. 4, pp. 50–58, 2010.
[3] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility," Future Generation Computer Systems, vol. 25, no. 6, pp. 599–616, 2009.
[4] M. Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," UC Berkeley Reliable Adaptive Distributed Systems Laboratory, 2009.
[5] K. Hashizume, D. G. Rosado, E. Fernández-Medina and E. B. Fernandez, "An Analysis of Security Issues for Cloud Computing," Journal of Internet Services and Applications, vol. 4, 2013.
[6] M. Jensen, J. Schwenk, N. Gruschka and L. L. Iacono, "On Technical Security Issues in Cloud Computing," in IEEE International Conference on Cloud Computing, 2009.
[7] International Energy Agency, Energy and AI, IEA, Paris, 2025.
[8] International Energy Agency, Key Questions on Energy and AI, IEA, Paris, 2026.
[9] B. P. Rimal, E. Choi and I. Lumb, "A Taxonomy and Survey of Cloud Computing Systems," in Proceedings of the Fifth International Joint Conference on INC, IMS and IDC, 2009.
[10] M. Satyanarayanan, "The Emergence of Edge Computing," Computer, vol. 50, no. 1, pp. 30–39, 2017.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY