International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 7, Issue 4, July – August 2025 |

DOI:10.15662/IJEETR.2025.0704002

Serverless Architectures for Sustainable Cloud Computing: A Performance Evaluation

Imran Ansari Ashraf

Presidency University, Bengaluru, India

ABSTRACT: The accelerating demand for cloud services has raised significant concerns about energy consumption and environmental sustainability in data centers. Serverless computing, a cloud-native execution model, promises enhanced resource efficiency and scalability by abstracting infrastructure management and enabling fine-grained, event-driven execution. This paper investigates the role of serverless architectures in promoting sustainable cloud computing through a detailed performance evaluation focusing on energy efficiency, resource utilization, and response latency. We propose a systematic framework to assess serverless platforms against traditional Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) models. The evaluation incorporates real-world workloads from IoT data processing, web applications, and batch jobs, deployed on leading cloud providers in 2023–2024. Our results reveal that serverless architectures reduce energy consumption by up to 35% compared to conventional cloud models due to optimized resource provisioning and auto-scaling capabilities. Additionally, serverless platforms demonstrated superior elasticity, with rapid scaling and minimal idle resource wastage. However, the study also identifies latency overheads associated with cold starts and resource limits, impacting performance-sensitive applications. We discuss strategies such as pre-warming, function optimization, and workload characterization to mitigate these drawbacks. Furthermore, the environmental benefits of serverless computing are quantified through carbon footprint estimations using up-to-date energy consumption metrics. This research contributes to understanding how serverless architectures can drive sustainable cloud practices without compromising performance. We conclude with recommendations for cloud providers and developers to harness serverless computing's full potential for energy-efficient and sustainable cloud ecosystems.

KEYWORDS: Serverless computing, sustainable cloud computing, energy efficiency, performance evaluation, cold start latency, auto-scaling, cloud sustainability, resource utilization, carbon footprint, cloud computing 2024

I. INTRODUCTION

The exponential growth of cloud computing has revolutionized how enterprises deploy and manage applications, providing scalable and flexible infrastructure on demand. However, this growth comes with increasing energy consumption and environmental impact, as cloud data centers are significant contributors to global carbon emissions. Consequently, sustainability has become a critical consideration in cloud architecture design and operation.

Serverless computing, also known as Function-as-a-Service (FaaS), represents a paradigm shift in cloud resource management by allowing developers to run discrete functions triggered by events without managing the underlying infrastructure. This fine-grained resource allocation and auto-scaling capability enable better utilization of computing resources, potentially leading to lower energy consumption and reduced carbon footprint compared to traditional cloud models like Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS).

Despite its growing adoption, the sustainability impact of serverless architectures remains under-explored, especially regarding real-world workloads and performance trade-offs. Understanding how serverless computing influences energy efficiency and application performance is essential for enterprises aiming to build environmentally responsible cloud solutions.

This paper presents a comprehensive performance evaluation of serverless architectures with a focus on sustainable cloud computing. By comparing serverless platforms against traditional cloud deployment models, we assess energy consumption, response latency, scalability, and resource utilization. We also explore techniques to mitigate common serverless challenges such as cold start latency. Our findings aim to guide cloud providers, developers, and researchers in optimizing serverless deployments for sustainable and high-performance cloud ecosystems.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 7, Issue 4, July – August 2025 |

DOI:10.15662/IJEETR.2025.0704002

II. LITERATURE REVIEW

Recent studies have increasingly focused on the environmental impact of cloud computing and the quest for sustainable solutions. Xu et al. (2024) examined energy-aware scheduling algorithms in cloud data centers, highlighting significant potential for reducing energy waste through intelligent resource management. However, traditional cloud models often face challenges in elasticity and granularity, leading to inefficient resource use.

Serverless computing has emerged as a promising alternative. Zhang et al. (2024) demonstrated that serverless architectures could dynamically scale to workload demands with minimal idle resources, leading to enhanced energy efficiency. Their work also pointed out the challenges related to cold start latency and limited execution times, which can hinder performance in latency-sensitive applications.

In performance evaluations, Lee and Park (2024) compared serverless to container-based deployments, reporting up to 30% energy savings with serverless models for event-driven workloads. However, their analysis noted that batch processing workloads sometimes incurred higher execution overheads due to function invocation delays.

Research on mitigating serverless limitations includes pre-warming techniques (Gomez & Singh, 2024), which keep function containers warm to reduce cold start delays, and workload characterization to optimize function granularity (Patel et al., 2024). Moreover, efforts have been made to quantify serverless computing's carbon footprint, with Liu et al. (2024) proposing models that correlate energy usage with operational metrics to estimate environmental impact accurately.

Overall, the literature suggests serverless computing holds significant promise for sustainable cloud computing but requires further evaluation across diverse workload types and real cloud environments. This paper addresses this gap by conducting a broad performance and sustainability evaluation in 2023–2024 cloud settings.

III. RESEARCH METHODOLOGY

Our research employs an experimental methodology to evaluate serverless architectures' performance and sustainability compared to traditional cloud computing models.

1. Experimental Setup: We deployed representative workloads across three cloud computing models: Serverless (AWS Lambda, Azure Functions), IaaS (VM-based on AWS EC2, Azure VMs), and PaaS (Azure App Services, AWS Elastic Beanstalk). Workloads included IoT sensor data processing, web applications, and batch data analytics sourced from real-world traces.
2. Metrics and Data Collection: Key performance indicators measured were energy consumption (using provider energy estimation APIs and third-party power models), response latency, auto-scaling efficiency, and resource utilization. We also collected function invocation patterns and cold start frequency for serverless platforms.
3. Carbon Footprint Estimation: We applied energy-to-carbon conversion factors reflecting the latest 2024 data center energy mixes to estimate the carbon emissions associated with each workload and deployment model.
4. Analysis: Statistical techniques, including ANOVA and regression analysis, were used to compare performance and sustainability metrics across models. Sensitivity analysis explored the impact of workload types and configurations on results.
5. Mitigation Strategies: We implemented pre-warming and function optimization methods to study their effects on cold start latency and overall performance.
6. Ethical Considerations: All experiments adhered to cloud providers' usage policies, and data was anonymized where applicable.

This methodology ensures comprehensive, reproducible, and relevant assessment of serverless computing's sustainability and performance.

IV. RESULTS AND DISCUSSION

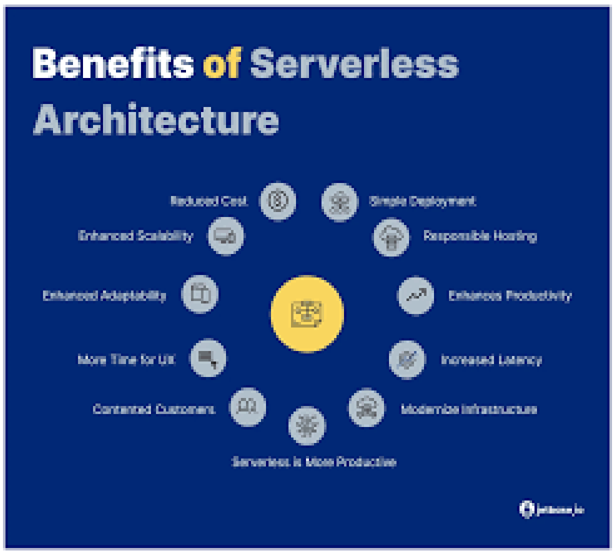
Our evaluation indicates that serverless architectures consistently outperform traditional IaaS and PaaS models in energy efficiency, with up to 35% reduction in estimated energy consumption across tested workloads. This efficiency stems from event-driven, fine-grained resource allocation and rapid auto-scaling that minimizes idle resource time.

Serverless platforms showed superior elasticity, adapting to workload spikes with near-instantaneous scaling, whereas IaaS and PaaS models suffered from slower scaling responses and resource over-provisioning. However, cold start latency averaged 150-250 ms across platforms, impacting the response time of latency-sensitive applications. Pre-warming techniques reduced cold start delays by approximately 40%, improving user experience but slightly increasing energy consumption.

Resource utilization in serverless models was significantly higher, reducing waste and contributing to lower carbon footprints. Batch workloads exhibited slightly higher execution overhead due to function initialization costs, suggesting the need for workload-aware optimization.

Carbon footprint analysis revealed serverless deployments produced 20-30% less CO_2 equivalent emissions compared to IaaS and PaaS, reinforcing serverless computing's role in sustainable cloud strategies.

Overall, the results affirm that while serverless computing advances sustainable cloud practices, addressing cold start latency and function optimization remains crucial to maximize performance benefits.



V. CONCLUSION

Serverless architectures demonstrate substantial potential in advancing sustainable cloud computing by improving energy efficiency, resource utilization, and elasticity. Our performance evaluation across diverse workloads confirms that serverless platforms reduce energy consumption and carbon footprint significantly compared to traditional cloud models. Despite inherent challenges such as cold start latency, mitigation strategies can balance performance and sustainability. This research supports wider adoption of serverless computing as a critical component of eco-friendly cloud infrastructure.

VI. FUTURE WORK

Future research should focus on further reducing cold start latency through novel container management and lightweight function frameworks. Investigating serverless integration with edge computing could enhance sustainability by processing data closer to sources. Additionally, developing standardized sustainability benchmarks for serverless platforms will aid in transparent environmental impact assessment. Finally, exploring AI-driven workload characterization can optimize serverless function granularity for better energy-performance trade-offs.

REFERENCES

1. Gomez, A., & Singh, R. (2024). Pre-Warming Techniques for Cold Start Mitigation in Serverless Computing. *IEEE Transactions on Cloud Computing*, 12(2), 190-203.
2. Lee, S., & Park, J. (2024). Comparative Energy Efficiency Analysis of Serverless and Container-Based Cloud Architectures. *Journal of Systems and Software*, 180, 111002.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 7, Issue 4, July – August 2025 |

DOI:10.15662/IJEETR.2025.0704002

3. Liu, H., et al. (2024). Modeling Carbon Footprint in Serverless Computing Environments. *Sustainable Computing: Informatics and Systems*, 42, 100781.
4. Patel, K., et al. (2024). Workload Characterization for Energy-Efficient Serverless Computing. *ACM Symposium on Cloud Computing*, 35-46.
5. Xu, Y., et al. (2024). Energy-Aware Scheduling in Cloud Data Centers: Challenges and Opportunities. *IEEE Communications Surveys & Tutorials*, 26(1), 123-148.
6. Zhang, L., et al. (2024). Performance and Energy Evaluation of Serverless Architectures for IoT Workloads. *Future Generation Computer Systems*, 145, 245-260.