Green Computing Optimization for Multi-Region Streaming Platforms

Sravya Sambaturu

Abstract: ***Global streaming services are always under pressure to control their resource usage effectively while still offering millions of users worldwide flawless experiences. As user demands for continuous, high-quality content increase, it is becoming increasingly important to strike the right balance between low latency, high availability, and resource efficiency. However, the infrastructure required to meet these demands often results in significant energy consumption and operational costs, presenting a major challenge for sustainability. The application of green computing principles to multi-region cloud infrastructure optimization is examined in this article, with an emphasis on tactics that minimize energy consumption and operational costs without compromising the performance that users rely on. Platforms can preserve their competitive advantage while making significant progress toward environmental responsibility by implementing more intelligent, energy-efficient procedures.***

Keywords: ***Cloud Infrastructure, Energy Efficiency, Green Computing, Resource Optimization***

I. INTRODUCTION

The streaming industry relies on massive infrastructure, with many servers spread across several areas to provide uninterrupted services. However, this resource-intensive configuration dramatically increases carbon emissions and operating inefficiencies. Green computing methods can solve these concerns by optimizing resource utilization while preserving quality [1].

II. CHALLENGES IN TRADITIONAL STREAMING INFRASTRUCTURE

Before delving into green computing principles, it is critical to understand the primary issues that standard streaming infrastructures confront. These platforms frequently rely on a static, over-provisioned architecture, in which resource allocation is based on peak traffic demands rather than real usage patterns. As a result, during off-peak hours, significant elements of the infrastructure remain idle, wasting power and contributing to excessive energy expenses and carbon emissions [2].

Furthermore, the nature of streaming services necessitates rapid responsiveness and low latency [3], which has traditionally required platforms to over-allocate resources to avoid slowdowns during peak usage periods; however [4], this "always on" model results in significant inefficiencies, particularly when traffic is light [5]. This becomes a huge issue for platforms that want to minimize their environmental effect while yet providing a flawless user experience [6]. Addressing these inefficiencies with green computing has the potential to tackle many of these long-standing issues [7].

III. BACKGROUND AND PRINCIPLES

To accommodate users across the globe, streaming platforms typically rely on multi-region architectures. These architectures involve replicating services in different regions to handle varying loads and ensure reliability. However, a common issue arises from over-provisioning: servers and resources are often allocated based on peak demand but remain underutilized during off-peak hours, resulting in wasted energy and inefficient operations.

Green computing—an approach aimed at minimizing energy consumption, optimizing resource usage, and increasing system adaptability—can be employed to fine-tune these infrastructures. By aligning these principles with operational practices, streaming platforms can reduce operational costs and carbon emissions while still delivering. High-performance services to their global user base.

Key Principles of Green Computing:

1. **Energy Efficiency**: Prioritizing energy-efficient hardware and software solutions helps reduce the overall power consumption of the infrastructure. Leveraging lower-power servers and optimizing energy usage through software-level management can greatly minimize unnecessary energy waste.
2. **Resource Optimization**: This principle focuses on better utilization of available resources by matching the resources to actual demand, ensuring that servers and instances are only active when needed. Techniques like virtual machine consolidation, right-sizing instances, and server resource sharing can improve resource efficiency.
3. **Dynamic Scalability**: Instead of relying on static, fixed resource allocation, dynamic scalability allows resources to be adjusted in real time based on demand fluctuations. Elastic scaling technologies enable servers to scale up or down depending on traffic, reducing idle resources during off-peak hours.
4. **Sustainability and Carbon Footprint Reduction**: Integrating renewable energy sources into the infrastructure, such as sourcing energy from wind, solar, or other green sources, significantly reduces the carbon footprint. Additionally, managing workloads efficiently across regions can help lower the carbon intensity of the energy used.

By applying these green

Manuscript received on 17 December 2024 | First Revised Manuscript received on 25 December 2024 | Second Revised Manuscript received on 03 January 2025 | Manuscript Accepted on 15 January 2025 | Manuscript published on 30 January 2025.
*****Correspondence Author(s)

Sravya Sambaturu*, DevOps, IT Industry, Hyderabad (Telangana), India. Email ID: sravya.sambaturu75@gmail.com, ORCID ID: 0000-0001-9061-6770

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

computing principles, streaming platforms can not only reduce their environmental impact but also enhance their operational efficiency and cost-effectiveness, all while maintaining the high-quality user experience that their global audience expects.

IV. SCENARIO: OVER-PROVISIONED MULTI-REGION SETUP

A. Initial Setup

A streaming platform operates across three regions (Americas, Europe, and Asia-Pacific), with the following details:

- **Total user base:** 20 million active users.
- **Peak hours:** Vary by region (Americas: 6 PM–12 AM, Europe: 6 PM–12 AM, Asia-Pacific: 8 PM–2 AM).
- **Instances per region:** 500 instances (Standard, 8 vCPUs, 32GB RAM, $0.40/hour each).
- Average CPU utilization during peak: 60%.
- Average CPU utilization during off-peak: 10%.

B. Problem Identification

- **Underutilization during off-peak hours:** Each region maintains a full complement of instances, despite low activity.
- **High energy consumption:** Idle instances continue to draw power, contributing to unnecessary CO2 emissions.
- **Static allocation:** Resources are not dynamically scaled based on demand fluctuations.

C. Proposed Approach

i. Dynamic Workload Allocation

1. **Global Load Balancing:**
   - Use traffic steering to redirect users to regions with available capacity during off-peak hours.
   - Example: Users in Europe during their off-peak hours can be redirected to servers in the Americas operating at higher utilization.
2. **Elastic Scaling:**
   - Configure auto-scaling groups to dynamically adjust the number of instances based on real-time demand.
   - Scale down to a minimum baseline during off-peak hours.
3. **Instance Right-Sizing:**
   - Replace underutilized standard instances with smaller instance types (e.g., 4 vCPUs, 16GB RAM) for off-peak workloads.
4. **Green Energy Integration:**
   - Deploy workloads in regions with renewable energy sources during non-peak hours to reduce carbon footprint.

V. ASSUMPTIONS

- Each idle instance consumes 0.2 kWh per hour.
- Carbon intensity of energy: 0.5 kg CO2 per kWh.
- Smaller instances (4 vCPUs) cost $0.20/hour and consume 0.1 kWh per hour.

VI. RESULT AND DISCUSSION

We analyze the impact of implementing green computing strategies on resource consumption, operational costs, and environmental sustainability. The following tables illustrate the comparison between the initial setup and the optimized approach, highlighting the significant improvements in cost reduction, energy efficiency, and carbon footprint reduction.

Table 1: Initial Resource usage and Costs

Metric	Peak Hours(6hrs)	Off-Peak Hours(18hrs)	Total Daily
Instance per region	500	500	-
CPU utilization	60%	10%	-
Energy consumption (kWh)	300,000	900,000	1,200,000
Daily cost ($)	24,000	72,000	96,000
CO2 emissions (kg)	600,000	1,800,000	2,400,000

Table 2: Optimized Resource usage and Costs

Metric	Peak Hours (6 hrs.)	Off-Peak Hours (18 hrs.)	Total Daily
Instances per region	500 (full-size)	250 (small-size)	-
CPU utilization	80%	50%	-
Energy consumption (kWh)	300,000	225,000	525,000
Daily cost ($)	24,000	27,000	51,000
CO2 emissions (kg)	600,000	112,500	712,500

Table 3: Summary of Benefits

Metric	Initial State	Optimized State	Savings
Daily Cost ($)	96,000	51,000	45,000 (47%)
Annual Cost ($)	35,040,000	18,615,000	16,425,000
Daily Energy (kWh)	1,200,000	525,000	675,000(56%)
AnnualCO2 (kg)	876,000,000	260,625,000	615,375,000

VII. CONCLUSION

This case study highlights the effectiveness of green computing principles in optimizing multi-region cloud infrastructures. By adopting strategies such as dynamic workload allocation, elastic scaling, and right-sizing, streaming platforms can significantly reduce energy consumption and operational costs, while also cutting down on their carbon emissions. These optimizations not only benefit the environment but also improve the long-term sustainability and profitability of streaming services. The integration of green energy and resource-efficient practices allows organizations to meet the increasing demand for streaming content without compromising on performance or environmental responsibility.

DECLARATION STATEMENT

I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been sponsored or funded by any organization or agency. The independence of this research is a crucial factor in affirming its impartiality, as it has been conducted without any external sway.
- **Ethical Approval and Consent to Participate:** The data provided in this article is exempt
from the requirement for ethical approval or participant consent.
- **Data Access Statement and Material Availability:** The

adequate resources of this article are publicly accessible.

- **Authors Contributions:** The authorship of this article is contributed solely.

REFERENCES

1. Reddy, S. P., & Chandan, H. K. S. (2014, February). Energy-aware scheduling of real-time and non-real-time tasks on cloud processors (Green Cloud Computing). International Conference on Information Communication and Embedded Systems (ICICES2014). DOI: https://doi.org/10.1109/ICICES.2014.7033827
2. Shailesh S. Deore, Ashok N. Patil, and Ruchira Bhargava. 2012. Energy-efficient scheduling scheme for virtual machines in cloud computing. International Journal of Computer Applications 56, 10 (Oct. 2012). DOI: https://doi.org/10.5120/8926-2999
3. J Kavitha, Rokesh Kumar Yarava, Energy Efficiency Analysis Between Green Computing and Cloud Computing. (2019). In International Journal of Recent Technology and Engineering (Vol. 8, Issue 3S2, pp. 254–258). DOI: https://doi.org/10.35940/ijrte.C1047.1083S219
4. Jain, J. (2019). Modern with Advanced Direction in Green Cloud. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 2, pp. 3090–3095). DOI: https://doi.org/10.35940/ijeat.F9184.129219
5. Vatsal, S., & Verma, Dr. S. B. (2023). Data Centre Efficiency Enhancement by Metrics Oriented Approach to Revamp Green Cloud Computing Concept. In International Journal of Innovative Technology and Exploring Engineering (Vol. 12, Issue 8, pp. 1–14). DOI: https://doi.org/10.35940/ijitee.F9532.0712823
6. Naik, O., Jaiswal, S., Bhuyar, Dr. K., & Kodape, Dr. S. (2023). Green Hydrogen Production as a Sustainable Initiative for Alternative Energy Source: A Review. In International Journal of Basic Sciences and Applied Computing (Vol. 9, Issue 10, pp. 1–5). DOI: https://doi.org/10.35940/ijbsac.I0503.0691023
7. Paul, D., Pal, O. K., Islam, Md. M., Mohammad, M., & Babu, R. M. (2023). Design and Implementation of an Efficient Smart Digital Energy Meter. In International Journal of Soft Computing and Engineering (Vol. 13, Issue 1, pp. 25–30). DOI: https://doi.org/10.35940/ijsce.A3600.0313123

AUTHOR'S PROFILE

Sravya Sambaturu is currently working as a DevOps Engineer with over five years of experience in the IT industry. A graduate of JNTUH, she focuses on cloud computing using AWS and GCP, along with implementing DevOps tools and practices. Her work involves managing GitLab CI/CD pipelines, container orchestration with Kubernetes, and automating workflows to improve efficiency. She enjoys exploring new tools and techniques in cloud technologies and DevOps, aiming to build practical, scalable solutions.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)/ journal and/or the editor(s). The Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.