

LEVERAGING PUBLIC CLOUD INFRASTRUCTURE FOR COST-EFFECTIVE, AUTO-SCALING SOLUTIONS

Nagarjuna Putta¹, Shyamakrishna Siddharth Chamarthy², Krishna Kishor Tirupati³, Prof. (Dr) Sandeep Kumar⁴, Prof. (Dr) MSR Prasad⁵ & Prof. (Dr) Sangeet Vashishtha⁶

¹SV University, Tirupathi, Andhra Pradesh, India

²Scholar, Columbia University, Sakthinagar 2nd Ave, Nolambur, Chennai, Tamil Nadu, India

³International Institute of Information Technology Bangalore, India

⁴Department of Computer Science and Engineering Koneru Lakshmaiah Education Foundation Vadeshawaram,
A.P., India

⁵Department of Computer Science and Engineering Koneru Lakshmaiah Education Foundation Vadeshawaram,
A.P., India

⁶IIMT University, Meerut, India

ABSTRACT

Public cloud infrastructure has emerged as a transformative solution for businesses seeking cost-effective and scalable IT operations. This paper explores how organizations can leverage public cloud platforms to achieve automatic scalability, ensuring optimal resource utilization in dynamic environments. The ability of public cloud providers, such as AWS, Microsoft Azure, and Google Cloud, to offer on-demand resources significantly reduces upfront capital expenditures and minimizes operational costs. Auto-scaling, a core feature of public clouds, enables applications to automatically adjust resource capacity based on real-time demand, ensuring high availability and performance without manual intervention.

This study highlights the advantages of adopting public cloud solutions, including reduced infrastructure maintenance, pay-as-you-go pricing models, and access to innovative technologies like machine learning and IoT platforms. Furthermore, it discusses how auto-scaling eliminates over-provisioning issues by aligning resource usage with fluctuating workloads, making it ideal for industries with seasonal or unpredictable demand patterns. Challenges related to data security, compliance, and vendor lock-in are also addressed, providing insights into strategies for mitigating these risks.

In conclusion, the use of public cloud infrastructure presents businesses with a flexible and cost-efficient way to meet their IT requirements while maintaining performance and scalability. With the increasing availability of advanced tools and cloud-native services, organizations can further optimize their operations, fostering innovation and growth. This paper underscores the importance of integrating auto-scaling capabilities within cloud environments to ensure a sustainable and competitive edge in today's fast-evolving digital landscape.

KEYWORDS: *Public cloud Infrastructure, Auto-Scaling Solutions, Cost-Effective IT Operations, Resource Optimization, Cloud Platforms, On-Demand Scalability, Pay-As-You-Go Models, Cloud-Native Services, Workload Management, Operational Efficiency*

Article History

Received: 10 Dec 2022 | Revised: 22 Dec 2022 | Accepted: 22 Dec 2022
