

OPTIMIZING CLOUD INFRASTRUCTURE FOR SCALABLE DATA PROCESSING SOLUTIONS

Rajas Paresh Kshirsagar¹, Phanindra Kumar Kankanampati², Ravi Kiran Pagidi³, Aayush Jain⁴, Dr. Shakeb Khan⁵ & Prof.(Dr.) Arpit Jain⁶

¹Independent Researcher, N.Y. University, Malad (W), Mumbai, Maharashtra, India,

²Independent Researcher, Binghamton University, Miyrapur, Hyderabad, India

³Independent Researcher, Technological University, Hyderabad, India

⁴Independent Researcher, Vivekananda Institute of Professional Studies, Pitampura, Delhi

⁵Research Supervisor, Maharaja Agrasen Himalayan Garhwal University, Uttarakhand

⁶Independent Researcher, KL University, Vijaywada, Andhra Pradesh, India

ABSTRACT

The rapid growth of data in modern enterprises necessitates the development of scalable and efficient data processing solutions. Cloud infrastructure has emerged as a powerful tool for addressing these needs, offering elasticity, flexibility, and cost-efficiency. This paper explores various strategies for optimizing cloud infrastructure to support scalable data processing, focusing on key elements such as dynamic resource management, distributed computing, and performance optimization.

The research highlights the significance of auto-scaling mechanisms, which automatically adjust computing resources based on real-time demand, ensuring that data workloads are processed efficiently. It also delves into the role of containerization, particularly through platforms like Kubernetes, to streamline the deployment and management of scalable applications. In addition, the paper examines serverless computing as a cost-effective approach for handling sporadic workloads, reducing the need for constant server management while maintaining processing power on demand.

Moreover, the study discusses techniques for optimizing storage and network bandwidth to avoid bottlenecks in data processing pipelines. It also addresses the integration of cloud-native tools, such as managed databases and data analytics services, which simplify infrastructure management and improve overall system efficiency.

By adopting these strategies, organizations can design robust cloud architectures that support the scalable and efficient processing of large datasets, enabling them to meet growing business demands and enhance operational agility. This research provides a comprehensive guide for leveraging cloud technologies to optimize data processing, contributing to long-term organizational success in a data-driven world.

KEYWORDS: *Cloud Infrastructure, Scalable Data Processing, Dynamic Resource Management, Auto-Scaling, Containerization, Kubernetes, Serverless Computing, Storage Optimization, Distributed Computing, Cloud-Native Tools, Data Pipelines*

Article History

Received: 10 May 2024 | Revised: 14 May 2024 | Accepted: 30 Jun 2024
