World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:12, 2023

Optimizing Data Transfer and Processing in Multi-Cloud Environments for Big Data Workloads

Authors : Gaurav Kumar Sinha

Abstract : In an era defined by the proliferation of data and the utilization of cloud computing environments, the efficient transfer and processing of big data workloads across multi-cloud platforms have emerged as critical challenges. This research paper embarks on a comprehensive exploration of the complexities associated with managing and optimizing big data in a multi-cloud ecosystem.The foundation of this study is rooted in the recognition that modern enterprises increasingly rely on multiple cloud providers to meet diverse business needs, enhance redundancy, and reduce vendor lock-in. As a consequence, managing data across these heterogeneous cloud environments has become intricate, necessitating innovative approaches to ensure data integrity, security, and performance.The primary objective of this research is to investigate strategies and techniques for enhancing the efficiency of data transfer and processing in multi-cloud scenarios. It recognizes that big data workloads are characterized by their sheer volume, variety, velocity, and complexity, making traditional data management solutions insufficient for harnessing the full potential of multi-cloud architectures.The study commences by elucidating the challenges posed by multi-cloud environments in the context of big data. These challenges encompass data fragmentation, latency, security concerns, and cost optimization. To address these challenges, the research explores a range of methodologies and solutions. One of the key areas of focus is data transfer optimization. The paper delves into techniques for minimizing data movement latency, optimizing bandwidth utilization, and ensuring secure data transmission between different cloud providers. It evaluates the applicability of dedicated data transfer protocols, intelligent data routing algorithms, and edge computing approaches in reducing transfer times.Furthermore, the study examines strategies for efficient data processing across multi-cloud environments. It acknowledges that big data processing requires distributed and parallel computing capabilities that span across cloud boundaries. The research investigates containerization and orchestration technologies, serverless computing models, and interoperability standards that facilitate seamless data processing workflows.Security and data governance are paramount concerns in multi-cloud environments. The paper explores methods for ensuring data security, access control, and compliance with regulatory frameworks. It considers encryption techniques, identity and access management, and auditing mechanisms as essential components of a robust multi-cloud data security strategy.The research also evaluates cost optimization strategies, recognizing that the dynamic nature of multi-cloud pricing models can impact the overall cost of data transfer and processing. It examines approaches for workload placement, resource allocation, and predictive cost modeling to minimize operational expenses while maximizing performance.Moreover, this study provides insights into real-world case studies and best practices adopted by organizations that have successfully navigated the challenges of multi-cloud big data management. It presents a comparative analysis of various multi-cloud management platforms and tools available in the market.

Keywords : multi-cloud environments, big data workloads, data transfer optimization, data processing strategies
Conference Title : ICDMBDDDS 2023 : International Conference on Data Mining, Big Data, Database and Data System
Conference Location : New York, United States
Conference Dates : December 11-12, 2023