IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 1, August 2025

Impact Factor: 7.67

Building Resilient, Globally Distributed Systems with Azure Cosmos DB: Optimizing Performance and Latency

Giriraj Agarwal
Sr. Manager - Projects – Cognizant
https://orcid.org/0009-0006-1042-6568

Abstract: *In today's increasingly connected digital environment, modern applications demand ultra-low latency, high availability, and seamless global scalability. These requirements have accelerated the shift toward cloud-native, distributed database architectures. Azure Cosmos DB, Microsoft's fully managed NoSQL database service, has emerged as a leading solution empowering developers to build fault-tolerant, globally distributed applications.*

This study explores how Cosmos DB's architecture enables high performance, continuous availability, and low-latency data access across multiple geographies. It analyzes features such as multi-region replication, multiple consistency models, multi-region writes, and automatic failover, all contributing to its fault-tolerant and high-throughput capabilities. By replicating data across multiple Azure regions, Cosmos DB ensures 99.999% uptime and sub-10-millisecond latencies for both read and write operations—an essential requirement for mission-critical applications in finance, e-commerce, IoT, and international logistics.

The paper also highlights advanced features like AI-driven performance optimization and intelligent data partitioning, which enhance scalability without compromising consistency or reliability. Through a qualitative, case study–based research approach, the study evaluates real-world scenarios where Cosmos DB has enabled global organizations to improve user experience through guaranteed uptime, fast data delivery, and dynamic control over throughput.

Additionally, the research offers a comparative analysis between Cosmos DB and traditional NoSQL solutions, focusing on its strengths in automatic horizontal scaling and multi-model API support (including SQL, MongoDB, Cassandra, Gremlin, and Table APIs). Findings indicate that Azure Cosmos DB not only meets but often exceeds the architectural expectations of next-generation applications through its inherent resilience and performance guarantees.

The paper concludes with practical recommendations for deploying Cosmos DB in globally distributed systems, addressing cost management, latency optimization, and consistency model selection. As global commerce continues to prioritize user experience and operational reach, adopting platforms like Azure Cosmos DB becomes essential for building robust, future-ready digital infrastructure.

Keywords: Azure Cosmos DB, global availability, low latency, distributed systems, NoSQL databases, high performance, resilience, cloud computing, real-time applications, database scalability

DOI: 10.48175/IJARSCT-28620

ISSN 2581-9429 IJARSCT