International Journal of Computer Science and Engineering (IJCSE)
ISSN (P): 2278–9960; ISSN (E): 2278–9979
Vol. 10, Issue 1, Jan – Jun 2021; 119–138

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

DEVELOPING SCALABLE SEARCH INDEXING INFRASTRUCTURES FOR HIGH-VELOCITY E-COMMERCE PLATFORMS

Suraj Dharmapuram[1], Imran Khan[2], Murali Mohana Krishna Dandu[3], Prof. (Dr) Punit Goel[4], Prof. (Dr) Arpit Jain[5] & Er. Aman Shrivastav[6]

[1]Suraj Dharmapuram, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213

[2]Scholar, Visvesvaraya Technological University, College - MVJ College of Engineering, Bangalore, India

[3]Scholar, Texas Tech University, USA

[4]Maharaja Agrasen Himalayan Garhwal University, Uttarakhand, India

[5]KL University, Vijayawada, Andhra Pradesh, India

[6]ABESIT Engineering College, Ghaziabad, India

ABSTRACT

As e-commerce platforms grow in complexity and user demand, maintaining high-performance search functionality becomes increasingly challenging. Users expect real-time, relevant results, even as product catalogs expand and user traffic increases. This paper addresses the architecture and implementation of scalable search indexing infrastructures optimized for high-velocity e-commerce platforms, where query speed and result accuracy are essential for user retention and conversion rates.

The study first explores current limitations of traditional search indexing systems, which often struggle under high request volumes and frequent product updates. Key challenges in this domain include managing massive data flows, ensuring low-latency responses, handling distributed architectures, and supporting advanced query requirements. This paper proposes a layered approach to mitigate these issues, emphasizing scalability, resilience, and fault tolerance. The architecture incorporates distributed indexing, in-memory databases, and intelligent caching strategies to maintain high performance.A central aspect of the proposed solution is the use of partitioned indexing across distributed nodes, supported by sharding and replication to handle variable loads. By leveraging these methods, the indexing infrastructure can efficiently scale horizontally, ensuring that increased traffic or data volume does not degrade performance. Additionally, we explore adaptive caching mechanisms based on query frequency and data popularity, which significantly reduce load on primary databases and lower latency for common queries. Machine learning techniques, such as predictive pre-fetching and dynamic query optimization, are integrated to further improve response times and relevance, especially during high-traffic events.

Real-time data synchronization is another critical factor. This paper introduces a strategy for maintaining up-to-date search indices through streaming data pipelines that allow continuous product updates without interrupting service availability.

The infrastructure design is validated through a series of benchmarking tests simulating real-world e-commerce scenarios, including Black Friday-level traffic spikes and sudden inventory changes.

In conclusion, the paper demonstrates that a scalable, distributed search indexing infrastructure can meet the demands of high-velocity e-commerce environments. Future research directions include enhancing machine learning algorithms for personalized search ranking and exploring serverless architectures for additional scalability. This framework provides a roadmap for e-commerce companies aiming to improve search experience, drive conversions, and maintain performance during peak traffic.

KEYWORDS: *Scalable Search Indexing, High-Velocity E-Commerce, Distributed Architecture, in-memory Databases, Real-Time Data Synchronization, Adaptive Caching, Machine Learning In Search, Query Optimization*

Article History

Received: *06 Jun 2021* | ***Revised:*** *15 Jun 2021* | ***Accepted:*** *20 Jun 2021*
