

## **Systematic Literature Review On Replication Distributed Data**

**Manohar Lal Raksha\*1, Dr. Sanjay Singh Bhadoriya\*2**

*\*1(Department of Comp. Sci. & Application, Dr. A. P. J. Abdul Kalam University, Indore)*

*\*2(Department of Comp. Sci. & Application, Dr. A. P. J. Abdul Kalam University, Indore)*

**Abstract:** An introductory introduction to database replication and a summary of the benefits and drawbacks of conventional methods are presented in this paper. Specifically, we discuss the primary shortcomings of these techniques in order to differentiate between inescapable and preventable constraints. This brings us to the main ideas of the strategy this dissertation suggests, which removes the preventable and lessens the inherent drawbacks of conventional methods. A feature that all of these definitions have in common is that a distributed system always seems to consumers to be one very powerful computer.

In this paper an introductory introduction to database replication and a summary of the benefits and drawbacks of conventional methods are presented.

**Keywords:** *Distributed System, Data replication, partial replication, Techniques.*

### **INTRODUCTION**

A systematic literature review on replication and distributed databases has been conducted to analyse recent trends and strategies in data replication techniques. The paper emphasises how crucial data replication is to distributed systems, especially when it comes to guaranteeing fault tolerance and high availability. The difficulties in preserving data consistency and the requirement for effective replication techniques to strike a balance between cost and availability are also covered.

**Distributed System:** A distributed system consists of a set of independent computers connected by a network and distribution middleware that allows computers to coordinate their activities and share system resources so that users see the system as a single computing device.[1]

**Data Replication:** Data replication is the process of storing data in more than one location or node. This is necessary to improve the availability of information. This can be full replication where a copy of the entire database is stored in each location. Partial replication can also occur, where some database fragments are copied and others are not. With the growing popularity of distributed systems, it has become easier to access data anytime and anywhere, and data replication has become the most widely used approach in distributed systems. [2,3]

**Contention management:** Conflict control policies, which are mechanisms used in distributed information systems to reduce conflicts and resolve disputes between nodes. The purpose of the dispute resolution policy is to ensure that the system advances to a certain level of justice by making the right decision in case of conflict. The choice of competition management policy strongly affects system performance.[4]

**Data Replication Strategies:** The review discusses various strategies for data replication, including voting structures, hybrid models, and genetic programming-based approaches. These strategies aim to improve performance, increase information availability and reduce costs.

**Distributed Database Design Techniques:** The survey additionally covers appropriated data set plan procedures, including information fracture, information portion, and information replication. It features the significance of these strategies in overseeing huge scope dispersed data sets and guaranteeing information consistency.

**Recent Trends and Future Directions:** The review concludes with a discussion of recent trends and future directions in data replication and distributed databases. This highlights the need for innovative strategies to address data replication challenges in large distributed systems. Overall, the systematic literature review provides a comprehensive overview of the latest trends and strategies in data replication and distributed databases, and highlights the importance of these techniques to ensure high availability, failover and data consistency in distributed systems.

## **LITERATURE REVIEW**

Data replication ensures high availability and reliability of cloud services in several different locations. However, challenges arise due to dynamic cloud configurations, increasing data volumes and different application requirements. The study analyzes different reproduction strategies - hybrid, dynamic and static - evaluating their effectiveness, advantages and disadvantages. Data replication is the process of storing data in more than one location or node. This is necessary to improve the availability of information. This can be full replication where a copy of the entire database is stored in each location. Partial replication can also occur, where some database fragments are copied and others are not. With the growing popularity of distributed systems, anytime, anywhere access to data has become even more possible. Distributed database These two issues become the design of the global schema and the local physical databases at each location. The techniques used in these aspects are the same as in centralized databases. Researchers are constantly improving information distribution methods in distributed systems to handle massive amounts of data, network expansion, and user numbers. The research emphasizes optimal data handling and minimal connection costs through database fragmentation and partitioning.

Over the past years, and in addition to our work, many researchers have begun to explore the combination of fragmentation, database partitioning and distributed system solutions. [5,6]

**Noraziah et al.** (2021)[7]: With the rapid increase in data generation, replication is vital for data management in remote databases, enhancing accessibility and reliability. The study introduces the Binary Vote Assignment on Grid Quorum with Association Rule (BVAGQ-AR) algorithm for managing fragmented database synchronous replication, improving utilization, dependability, and effectiveness of distributed databases.

**E. Mehmood et al.** (2020)[8]: Efficient execution of analytical tasks is crucial for business responsiveness to market shifts. This research proposes a method to assess data access patterns for large-scale analytics, aiding in data partitioning and replication. The approach benefits query optimization in modern data management systems.

**Yoga Andrian et al.** (2019) [9]: The Indonesian National Institute of Aeronautics and Space developed the Space Weather Information and Forecast Services (SWIFtS) system, which faced challenges due to a centralized data storage model.

**Sashi Tarun et al.** (2019) [10]: Researchers continuously improve methodologies for data distribution in distributed systems to manage massive data, network expansion, and user numbers. The study emphasizes

optimal data processing and minimal connection costs through database fragmentation, allocation, and replication, highlighting strategies and frameworks for efficient distributed systems.

**S. Annal et al.** (2015) [11]: Data replication ensures high availability and dependability in cloud computing across multiple locations. However, challenges arise due to dynamic cloud settings, increasing data volumes, and diverse application requirements. The study analyzes various replication strategies—hybrid, dynamic, and static assessing their effectiveness, benefits, and drawbacks. It also addresses challenges like consistency, scalability, fault tolerance, and security, aiming to optimize replication for better performance and resilience.

## CONCLUSION

The paper has given a review of the issues connected with accomplishing adaptation to internal failure by replication in disseminated frameworks. The growth of the Internet and distributed computing technology has encouraged the development of sophisticated database systems that enable applications such as electronic commerce, remote airline check-in, and telephone banking. In these environments, data replication is critical for both contention control and efficiency, allowing available sites to take over when servers fail and reduce response time by reading local copies.

## REFERENCES

1. *M. Baumer. Integrating synchronous partial replication into the PostgreSQL database engine.*
2. *Master's thesis, Department of Computer Science, ETH Zurich, Switzerland, 1999.*
3. *W. Bausch. Integrating synchronous update-everywhere replication into the PostgreSQL database engine. Master's thesis, Department of Computer Science, ETH Zurich, Switzerland, 1999.*
4. *C. Beeri, P.A. Bernstein, and N. Goodman. A model for concurrency in nested transactions systems. Journal of the ACM, 36(2):230–269, 1989.*
5. *C. Beeri, P.A. Bernstein, and N. Goodman. A model for concurrency in nested transactions systems. Journal of the ACM, 36(2):230–269, 1989.*
6. *T. A. Anderson, Y. Breitbart, H. F. Korth, and A. Wool. Replication, consistency, and practicality: Are these mutually exclusive? In Proc. of the ACM SIGMOD Int. Conf. on Management of Data, pages 484–495, Seattle, Washington, June 1998.*
7. *Noraziah ,Ainul Azila Che Fauzi, Sharifah Hafizah Sy Ahmad Ubaidillah;Basem Alkazemi;Julius Beneoluchi Odili(2021)BVAGQ-AR for Fragmented Database Replication ManagementIEEE AccessYear: 2021.*
8. *E. Mehmood and T. Anees, "Challenges and Solutions for Processing Real-Time Big Data Stream: A Systematic Literature Review," in IEEE Access, vol. 8, pp. 119123-119143, 2020, doi: 10.1109/ACCESS.2020.*
9. *Yoga Andrian;Essaid Maryam;Kim DaeYong;Soo Hoon Maeng;Hongtaek Ju(2019)Space Weather Data Management System and Monitoring in Decentralized Storage Environment 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS) Year: 2019*
10. *Sashi Tarun;Ranbir Singh Batth;Sukhpreet Kaur(2019)A Review on Fragmentation, Allocation and Replication in Distributed Database Systems2019 International Conference on Computational*

*Intelligence and Knowledge Economy (ICCIKE) Year: 2019*

11. S. Annal Ezhil Selvi , Dr. R. Anbuselvi, "An Analysis of Data Replication Issues and Strategies on Cloud Storage System" Conference: *International Journal of Engineering Research & Technology (IJERT)* [www.ijert.org](http://www.ijert.org) NCICN-2015 Conference Proceedings At: Kongunadu College of engineering and technologie Volume: Voume 4 Year: 2015