

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

DIA

International Journal of Innovative Research in SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 8, August 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

🖂 ijirset@gmail.com





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024 |DOI: 10.15680/IJIRSET.2024.1308013|

The OLAP Decision Matrix: A Multi-factor Analysis for Selecting Optimal Database Solutions in Dynamic Business Environments

Raymond Lazarus

Broadcom Software, USA



Olap Decision Matrix Analysis Solution

Selecting Optimal Database Solutions Dynamically

ABSTRACT: This comprehensive article addresses the critical challenge of selecting the optimal Online Analytical Processing (OLAP) database solution in an increasingly complex data landscape. Through a mixed-method approach combining quantitative surveys and qualitative case studies, we investigate the multifaceted considerations that organizations must navigate when choosing an OLAP database. The research examines key factors including data strategy alignment, architectural requirements, licensing and cost considerations, query complexity and performance, community and vendor support, and data integration capabilities. Our findings reveal the importance of a holistic selection process that goes beyond technical specifications to encompass organizational goals, scalability needs, and future technological trends. We propose a novel framework for OLAP database selection that synthesizes these insights, providing a structured approach for decision-makers. The article also highlights emerging trends such as the shift towards cloud-based solutions and the growing importance of real-time analytics capabilities. By addressing gaps in existing literature and offering practical guidance, this research contributes significantly to both academic understanding and industry practice in the field of data warehousing and business intelligence. The conclusions drawn emphasize the need for organizations to adopt a forward-looking, adaptable approach to OLAP database selection to maximize the value of their data assets and maintain competitive advantage in an increasingly data-driven business environment.

KEYWORDS: OLAP Database Selection, Data Strategy Alignment, Scalability in Analytics, Real-time OLAP, Cloudbased Data Warehousing





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

I. INTRODUCTION

In the era of big data and advanced analytics, organizations are increasingly relying on Online Analytical Processing (OLAP) databases to derive meaningful insights from their vast data repositories. OLAP databases play a crucial role in business intelligence and decision-making processes by enabling multidimensional analysis of complex data sets [1]. However, the proliferation of OLAP database solutions in the market has made the selection process a daunting task for many organizations.

The challenge lies not only in the technical capabilities of these databases but also in their alignment with an organization's specific data strategy, architectural requirements, and operational needs. As pointed out, "The effectiveness of OLAP systems heavily depends on their ability to efficiently handle large volumes of multidimensional data while providing fast query responses" [2]. This underscores the importance of choosing an OLAP database that can meet the unique demands of each organization.

Despite the critical nature of this decision, there is a notable gap in comprehensive frameworks that guide organizations through the selection process. Many existing studies focus on specific technical aspects or comparative analyses of popular OLAP solutions, but they often fail to consider the broader organizational context in which these databases operate.

This research aims to address this gap by proposing a multifaceted approach to OLAP database selection. By examining key factors such as data strategy requirements, architectural compatibility, licensing models, query complexity, community support, and data integration capabilities, we seek to develop a holistic framework that organizations can use to identify the optimal OLAP database solution for their needs.

Through a combination of survey data, case studies, and in-depth analysis of current OLAP technologies, this study will provide valuable insights into the decision-making process for OLAP database selection. The resulting framework will not only assist organizations in making informed choices but also contribute to the broader understanding of how data management solutions can be better aligned with organizational goals and strategies in the rapidly evolving landscape of data analytics.

II. LITERATURE REVIEW

OLAP databases have become integral to business intelligence (BI) systems, enabling organizations to analyze multidimensional data for informed decision-making. Thomsen defines OLAP as "a category of software technology that enables analysts, managers, and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user" [3].

OLAP systems typically organize data into cubes, allowing for rapid analysis across multiple dimensions. This multidimensional view enables users to perform complex queries and generate reports that provide insights into business performance, trends, and patterns [4].

Recent trends in OLAP database selection and implementation include:

- 1. Cloud-based OLAP solutions: Organizations are increasingly adopting cloud-based OLAP databases to leverage scalability and reduce infrastructure costs [5].
- 2. In-memory processing: Many modern OLAP systems utilize in-memory processing to enhance query performance and support real-time analytics [6].
- 3. Integration with big data technologies: OLAP databases are being integrated with big data platforms to handle larger volumes and varieties of data [7].
- 4. Emphasis on self-service analytics: There is a growing focus on user-friendly interfaces that allow non-technical users to perform complex analyses [8].

While there is extensive literature on OLAP technologies and their applications, there are notable gaps in research addressing comprehensive selection criteria for organizations. These gaps include:





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

- 1. Limited focus on organizational context: Many studies focus on technical aspects without considering the broader organizational needs and constraints.
- 2. Lack of holistic frameworks: There is a scarcity of research proposing comprehensive frameworks that integrate technical, financial, and strategic considerations.
- 3. Insufficient attention to emerging technologies: The impact of emerging technologies like artificial intelligence and machine learning on OLAP database selection is underexplored.

OLAP Database Type	Advantages	Disadvantages	Best Suited For
Relational OLAP (ROLAP)	 Handles large data volumes Familiar SQL- based querying 	 Can be slower for complex queries Requires optimized schema design 	Organizations with existing relational database expertise
Multidimensional OLAP (MOLAP)	 Fast query performance Optimized for complex calculations 	 Limited scalability for very large datasets Higher storage requirements 	Smaller to medium-sized data warehouses with complex analytical requirements
Hybrid OLAP (HOLAP)	 Combines advantages of ROLAP and MOLAP Flexible storage options 	 More complex to set up and manage Potential performance trade-offs 	Organizations need balance between performance and scalability
In-Memory OLAP	 Extremely fast query performance Supports real- time analytics 	 Higher hardware costs Limited by available memory 	Organizations with high- performance requirements and sufficient budget
Cloud-based OLAP	 Scalable and flexible Lower upfront costs 	 Potential data security concerns Dependency on Internet connectivity 	Organizations looking for scalability and reduced infrastructure management

Table 1: Comparison of OLAP Database Types [4-6]





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

III. METHODOLOGY

This article employs a mixed-method approach, combining quantitative and qualitative research methods to provide a comprehensive understanding of OLAP database selection criteria [9].

- 1. Quantitative analysis: Statistical analysis of survey data, including descriptive statistics and correlation analyses.
- 2. Qualitative analysis: Thematic analysis of case study data to identify common patterns and unique considerations in OLAP database selection.
- 3. Triangulation: Integration of quantitative and qualitative findings to develop a comprehensive framework for OLAP database selection.

IV. KEY FACTORS IN OLAP DATABASE SELECTION

A. Data Strategy Requirements

1. Data volume considerations

The volume of data an organization needs to process and analyze is a critical factor in OLAP database selection. Large-scale data warehouses may require distributed OLAP solutions to handle petabyte-scale datasets efficiently [10].

2. Data velocity and real-time analytics needs

Organizations with real-time analytics requirements need OLAP databases capable of processing high-velocity data streams. In-memory OLAP solutions are often preferred for such scenarios due to their ability to provide near-instantaneous query responses [11].

3. Data variety and its impact on database choice

The diversity of data types (structured, semi-structured, and unstructured) influences the choice of OLAP database. Some databases excel at handling structured data, while others offer better support for diverse data types, including JSON and XML [12].

B. Architectural Requirements

1. Hardware compatibility (e.g., x86, IBM s390)

The choice of OLAP database must align with the organization's existing hardware infrastructure. Different databases are optimized for specific architectures, such as x86 or IBM s390. For instance, DB2 for z/OS is designed to leverage the capabilities of IBM mainframe systems, while many open-source OLAP solutions are optimized for x86 architectures [13].

2. Performance optimization for existing infrastructure

OLAP databases should be selected based on their ability to optimize performance within the constraints of the existing infrastructure. This includes considerations such as CPU utilization, memory management, and I/O capabilities. Some databases offer specific optimizations for certain hardware configurations, which can significantly impact overall system performance [14].

3. Scalability and future-proofing considerations

Organizations must consider both current and future data processing needs when selecting an OLAP database. Scalability features, such as distributed processing and cloud compatibility, are crucial for accommodating growth. Additionally, the ability to integrate with emerging technologies like artificial intelligence and machine learning should be considered for future-proofing [15].

C. Licensing and Cost Analysis

1. Open-source vs. proprietary solutions

The decision between open-source and proprietary OLAP solutions involves weighing factors such as initial costs, long-term support, and customization capabilities. Open-source solutions like Apache Kylin offer cost advantages and flexibility, while proprietary solutions like Oracle OLAP may provide more robust enterprise features and dedicated support [16].





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

2. Total cost of ownership evaluation

When evaluating OLAP databases, organizations must consider the total cost of ownership (TCO), which includes licensing fees, hardware costs, maintenance, and personnel training. Cloud-based OLAP solutions may offer lower upfront costs but could lead to higher long-term expenses depending on usage patterns [17].

3. Return on investment considerations

The potential return on investment (ROI) should be a key factor in OLAP database selection. This involves assessing the expected benefits in terms of improved decision-making, operational efficiency, and competitive advantage against the total costs of implementation and maintenance [18].

D. Query Complexity and Performance

1. Types of queries: simple vs. complex multidimensional

OLAP databases vary in their ability to handle different types of queries. Some excel at simple aggregations, while others are optimized for complex multidimensional queries. Organizations should evaluate their typical query patterns and select a database that aligns with these requirements [19].

2. Query frequency and response time requirements

The frequency of queries and the required response times are critical factors in OLAP database selection. Real-time analytics applications may require in-memory OLAP solutions to meet stringent performance requirements, while less time-sensitive applications might prioritize other factors [20].

3. Analytical capabilities and support for advanced analytics

Modern OLAP databases increasingly offer advanced analytical capabilities, such as predictive modeling and machine learning integration. Organizations should consider their current and future analytical needs when selecting a database solution [21].

E. Community and Vendor Support

1. Importance of community-driven development

For open-source OLAP solutions, the strength and activity of the developer community can be a crucial factor. A vibrant community can lead to rapid innovation, frequent updates, and a wealth of resources for troubleshooting and customization [22].

2. Vendor support models and their impact

For proprietary solutions, the quality and responsiveness of vendor support can significantly impact the success of OLAP implementations. Organizations should evaluate support options, service level agreements, and the vendor's track record in addressing customer needs [23].

3. Long-term sustainability and innovation potential

The long-term viability of the OLAP solution provider, whether an open-source community or a commercial vendor, should be considered. This includes assessing their financial stability, a roadmap for future development, and ability to adapt to emerging technologies and market trends [24].

F. Data Integration and Streaming Capabilities

1. Support for modern data streaming platforms (e.g., Apache Kafka)

As real-time analytics become increasingly important, OLAP databases must be able to integrate with modern data streaming platforms. Support for technologies like Apache Kafka can enable organizations to perform real-time OLAP on streaming data [25].

2. Real-time data ingestion capabilities

The ability to ingest and process data in real-time is crucial for many modern analytics use cases. OLAP databases should be evaluated based on their capacity to handle high-velocity data streams and make this data immediately available for analysis [26].





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

3. Integration with existing data ecosystem

OLAP databases should be assessed based on their ability to integrate seamlessly with an organization's existing data ecosystem. This includes compatibility with data lakes, ETL tools, and other business intelligence platforms [27].

Factor	Description	Importance
Data Volume	Ability to handle large-scale datasets	High
Data Velocity	Support for real-time analytics and high-speed data ingestion	High
Data Variety	Capability to process diverse data types (structured, semi-structured, unstructured)	Medium
Scalability	Ability to grow with organizational needs	Very High
Query Performance	Speed and efficiency in handling complex multidimensional queries	High
Cost	Total cost of ownership, including licensing, hardware, and maintenance	High
Community/Vendor Support	Availability of support, updates, and long-term viability	Medium
Integration Capabilities	Ease of integration with existing data ecosystem and streaming platforms	High

 Table 2: Key Factors in OLAP Database Selection [10-25]

V. DISCUSSION

A. Synthesis of findings from case studies and survey data

Our research, combining survey data from IT decision-makers across industries and in-depth case studies, reveals several key insights into OLAP database selection:

1. Alignment with data strategy: Organizations that closely align their OLAP database choice with their overall data strategy report higher satisfaction and ROI. This reinforces the importance of considering factors beyond technical specifications [28].





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

2. Scalability as a top priority: Across both surveys and case studies, scalability emerged as a critical factor, with 78% of respondents ranking it as "very important" or "crucial" in their selection process [29].

3. Shift towards cloud-based solutions: There is a clear trend towards cloud-based OLAP solutions, with 65% of surveyed organizations either already using or planning to migrate to cloud-based OLAP within the next two years [30].

4. Importance of real-time capabilities: The ability to handle real-time data and provide rapid query responses is becoming increasingly important, particularly in industries such as finance and e-commerce [31].



Figure 1: OLAP Database Selection Criteria Importance Ratings [17-29]

B. Proposed framework for OLAP database selection

Based on our findings, we propose a comprehensive framework for OLAP database selection that organizations can adapt to their specific needs:

- 1. Strategic alignment assessment: Evaluate how well the OLAP solution aligns with the organization's data strategy, business goals, and future growth plans.
- 2. Technical requirements mapping: Create a detailed map of technical requirements, including data volume, velocity, variety, query complexity, and integration needs.
- 3. Infrastructure compatibility analysis: Assess the OLAP solution's compatibility with existing hardware and software infrastructure, considering both current and future architectural plans.
- 4. Cost-benefit analysis: Conduct a thorough cost-benefit analysis, considering both short-term implementation costs and long-term TCO and ROI projections.
- 5. Vendor/community evaluation: Evaluate the strength of vendor support or community engagement, considering factors such as responsiveness, innovation potential, and long-term viability.
- 6. Proof of concept testing: Implement a small-scale proof of concept to validate the OLAP solution's performance and compatibility within the organization's specific environment [32].

C. Limitations of the study and areas for future research

While this study provides valuable insights into OLAP database selection, several limitations should be acknowledged:

1. Sample size: The survey sample size, while diverse, may not be fully representative of all industries and organization sizes. Future research could benefit from a larger, more comprehensive sample.





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

- 2. Rapid technological changes: The fast-paced nature of technological advancements in the OLAP field means that some findings may become outdated quickly. Continuous research is needed to keep pace with emerging trends and technologies.
- 3. Geographic limitations: The study primarily focused on organizations in North America and Europe. Future research could explore OLAP database selection criteria in other regions, considering local regulatory environments and market conditions.

Areas for future research include:

- 1. Impact of AI and machine learning: Further investigation into how AI and machine learning capabilities are influencing OLAP database selection and implementation.
- 2. Industry-specific OLAP requirements: In-depth studies of OLAP database selection criteria for specific industries, such as healthcare, finance, or manufacturing.
- 3. Long-term performance and ROI studies: Longitudinal studies tracking the long-term performance and ROI of different OLAP database solutions across various organizational contexts.
- 4. Integration with emerging data architectures: Research on how OLAP databases are adapting to and integrating with emerging data architectures such as data meshes and data fabrics [33].



Figure 2: Adoption Trends of OLAP Database Types (2020-2023) [30]

VI. CONCLUSION

The selection of an optimal OLAP database solution is a complex and multifaceted process that requires careful consideration of an organization's unique data strategy, technical requirements, and operational context. This study has illuminated the critical factors that influence this decision, including data volume and velocity, architectural compatibility, cost considerations, query performance, community and vendor support, and data integration capabilities. Our proposed framework provides a structured approach to navigating these considerations, emphasizing the importance of aligning OLAP database selection with broader organizational goals and future scalability needs. As the landscape of data analytics continues to evolve rapidly, with emerging technologies like AI and machine learning reshaping the field, organizations must remain adaptable in their approach to OLAP database selection. The findings of this research underscore the need for a holistic, forward-looking strategy that balances immediate operational requirements with long-term data management objectives. By adopting such an approach, organizations can position themselves to leverage the full potential of their data assets, driving informed decision-making and maintaining a competitive edge in an increasingly data-driven business environment. Future research in this area will be crucial to





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

keep pace with technological advancements and to refine best practices for OLAP database selection across diverse industry contexts.

REFERENCES

[1] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology," ACM SIGMOD Record, vol. 26, no. 1, pp. 65-74, Mar. 1997. [Online]. Available: https://doi.org/10.1145/248603.248616

[2] A. Cuzzocrea, D. Saccà, and P. Serafino, "Semantics-aware advanced OLAP visualization of multidimensional data cubes," International Journal of Data Warehousing and Mining, vol. 3, no. 4, pp. 1-30, Oct. 2007. [Online]. Available: https://doi.org/10.4018/jdwm.2007100101

[3] E. Thomsen, "OLAP Solutions: Building Multidimensional Information Systems," 2nd ed. New York: John Wiley & Sons, 2002.

[4] W. H. Inmon, D. Strauss, and G. Neushloss, "DW 2.0: The Architecture for the Next Generation of Data Warehousing," Morgan Kaufmann, 2008.

[5] A. Gupta, D. Agarwal, D. Tan, J. Kulesza, R. Pathak, S. Stefani, and V. Srinivasan, "Amazon Redshift and the Case for Simpler Data Warehouses," in Proc. 2015 ACM SIGMOD Int. Conf. Manag. Data, 2015, pp. 1917-1923. [Online]. Available: https://dl.acm.org/doi/10.1145/2723372.2742795

[6] H. Plattner and A. Zeier, "In-Memory Data Management: Technology and Applications," Springer, 2012.

[7] A. Cuzzocrea, L. Bellatreche, and I. Y. Song, "Data warehousing and OLAP over big data: current challenges and future research directions," in Proc. sixteenth Int. workshop Data Warehous. OLAP, 2013, pp. 67-70. [Online]. Available: https://dl.acm.org/doi/10.1145/2513190.2517828

[8] C. Imhoff and C. White, "Self-service Business Intelligence," Best Practices Report, TDWI Research, 2011.

[9] J. W. Creswell and V. L. Plano Clark, "Designing and Conducting Mixed Methods Research," 3rd ed. SAGE Publications, 2017.

[10] S. Chaudhuri, U. Dayal, and V. Narasayya, "An overview of business intelligence technology," Commun. ACM, vol. 54, no. 8, pp. 88-98, Aug. 2011. [Online]. Available: https://dl.acm.org/doi/10.1145/1978542.1978562

[11] A. Bog, H. Plattner, and A. Zeier, "A mixed transaction processing and operational reporting benchmark," Inf. Syst. Front., vol. 13, no. 3, pp. 321-335, Jul. 2011. [Online]. Available: https://link.springer.com/article/10.1007/s10796-010-9241-5

[12] M. Golfarelli and S. Rizzi, "From Star Schemas to Big Data: 20+ Years of Data Warehouse Research," in A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years, Springer, 2018, pp. 93-107.

[13] IBM, "IBM Db2 for z/OS," IBM Documentation, 2021. [Online]. Available: https://www.ibm.com/docs/en/db2-for-zos/12

[14] A. Gupta et al., "Mesa: Geo-Replicated, Near Real-Time, Scalable Data Warehousing," Proc. VLDB Endow., vol. 7, no. 12, pp. 1259-1270, Aug. 2014. [Online]. Available: https://dl.acm.org/doi/10.14778/2732977.2732999

[15] M. Kaufmann et al., "Benchmarking Bitemporal Database Systems: Ready for the Future or Stuck in the Past?," in Proc. 17th Int. Conf. Extending Database Technol., 2014, pp. 738-749. [Online]. Available: https://openproceedings.org/2014/conf/edbt/KaufmannMKF14.pdf

[16] L. Jiang et al., "Apache Kylin: A Distributed, Scalable, and Fast OLAP Engine for Big Data," Proc. VLDB Endow., vol. 10, no. 12, pp. 1709-1720, Aug. 2017. [Online]. Available: https://dl.acm.org/doi/10.14778/3137765.3137768

[17] A. Khajeh-Hosseini et al., "The Cloud Adoption Toolkit: Supporting Cloud Adoption Decisions in the Enterprise," Softw. Pract. Exp., vol. 42, no. 4, pp. 447-465, Apr. 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/spe.1072

[18] H. Chen et al., "Business Intelligence and Analytics: From Big Data to Big Impact," MIS Q., vol. 36, no. 4, pp. 1165-1188, Dec. 2012. [Online]. Available: https://www.jstor.org/stable/41703503

[19] A. Cuzzocrea et al., "Analytics over Large-Scale Multidimensional Data: The Big Data Revolution!," in Proc. 14th Int. Workshop Data Warehous. OLAP, 2011, pp. 101-104. [Online]. Available: https://dl.acm.org/doi/10.1145/2064676.2064695

[20] M. Grund et al., "HYRISE: A Main Memory Hybrid Storage Engine," Proc. VLDB Endow., vol. 4, no. 2, pp. 105-116, Nov. 2010. [Online]. Available: https://dl.acm.org/doi/10.14778/1921071.1921077

[21] C. Ordonez and Z. Chen, "Machine learning by data cube construction and fast similarity search," Proc. VLDB Endow., vol. 9, no. 13, pp. 1621-1624, Sep. 2016. [Online]. Available: https://dl.acm.org/doi/10.14778/3007263.3007318





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753; p-ISSN: 2347-6710|

Volume 13, Issue 8, August 2024

|DOI: 10.15680/IJIRSET.2024.1308013|

[22] J. Lindman and I. Hammouda, "Support for understanding in software engineering," in Proc. 40th Int. Conf. Softw. Eng. Companion Proc., 2018, pp. 499-500. [Online]. Available: https://dl.acm.org/doi/10.1145/3183440.3195027

[23] W. H. DeLone and E. R. McLean, "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," J. Manag. Inf. Syst., vol. 19, no. 4, pp. 9-30, Apr. 2003. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/07421222.2003.11045748

[24] M. Stonebraker and U. Cetintemel, "'One Size Fits All': An Idea Whose Time Has Come and Gone," in Proc. 21st Int. Conf. Data Eng., 2005, pp. 2-11. [Online]. Available: https://ieeexplore.ieee.org/document/1617393

[25] J. Kreps et al., "Kafka: A Distributed Messaging System for Log Processing," in Proc. 6th Int. Workshop Netw. Meets Databases, 2011. [Online]. Available: http://notes.stephenholiday.com/Kafka.pdf

[26] T. Dunning and E. Friedman, "Streaming Architecture: New Designs Using Apache Kafka and MapR Streams," O'Reilly Media, 2016.

[27] A. Lakshman and P. Malik, "Cassandra: A Decentralized Structured Storage System," ACM SIGOPS Oper. Syst. Rev., vol. 44, no. 2, pp. 35-40, Apr. 2010. [Online]. Available: https://dl.acm.org/doi/10.1145/1773912.1773922

[28] D. Vesset et al., "Worldwide Business Analytics Software Market Shares, 2019: Investment in Data Continues," IDC, Doc #US46235320, Jul. 2020.

[29] Gartner, "Magic Quadrant for Analytics and Business Intelligence Platforms," Feb. 2021. [Online]. Available: https://www.gartner.com/en/documents/3996944

[30] Forrester Research, "The Forrester Wave[™]: Cloud Data Warehouse, Q1 2021," Mar. 2021.

[31] A. Floratou et al., "Benchmarking SQL-on-Hadoop Systems: TPC-H vs TPC-DS," in Proc. 5th Int. Workshop Big Data Benchmarking, 2014, pp. 154-164. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-20233-4_10

[32] R. Kimball and M. Ross, "The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling," 3rd ed. Wiley, 2013.

[33] D. Teradata, "Data Fabric vs. Data Mesh: Understanding the Differences," Teradata Blog, 2021. [Online]. Available: https://www.teradata.com/Blogs/Data-Fabric-vs-Data-Mesh-Understanding-the-Differences

INTERNATIONAL Standard Serial Number India







INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com