JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

Design and Implementation of Information Systems for Efficient Big Data Processing

Deni Apriadi[1*], Dewi Anjani[2], Andi Alviadi Nur Risal[3], Yaya Sudarya Triana[4], Husni Mubarak[5]

[1]Digital Business, Institut Teknologi Muhammadiyah Sumatera, Musi Rawas, Indonesia
[2] Computer Science Engineering, Infomatics Engineering, Universitas Indraprasta PGRI, Jakarta Barat, Indonesia
[3]Teacher Training and Education, Information Technology Education, Universitas Megarezky, Makassar, Indonesia
[4] Computer Science, Magister Data Science, Universitas Mercu Buana, Indonesia
[5] Economics and Business, Digital Business, Universitas Negeri Makassar, Makassar, Indonesia
Email: [1*] denidrv@gmail.com , [2]dewiunindra@gmail.com, [3] andialviadinurrisal@unimerz.ac.id,
[4]yaya.sudarya.triana@gmail.com, [5]husni.mubara@unm.ac.id
(*Email Corresponding Author: denidrv@gmail.com)
Received: May 4, 2026 | Revision: May 5, 2026 | Accepted: May 6, 2026

Abstract

The rapid growth of data volume, velocity, and variety has created significant challenges for traditional information systems, which are often unable to process large-scale data efficiently. This study aims to design and implement an efficient information system for big data processing using a distributed computing approach. The research adopts a systematic, experimental approach comprising system design, implementation, and performance evaluation. The proposed system is developed using a distributed architecture with parallel processing mechanisms to improve scalability and resource utilisation. Performance evaluation is conducted using key metrics, including processing time, throughput, and the percentage improvement in efficiency, based on experimental testing with datasets ranging from 1 GB to 10 GB. The results show that the proposed system consistently reduces processing time and increases throughput compared to the baseline system. The system achieves efficiency improvements of 33.3% to 36.9%, exceeding the predefined success threshold of 30%. These findings demonstrate that integrating distributed computing with an optimised system architecture significantly enhances big data processing performance. Therefore, the proposed system provides a scalable, practical solution for large-scale data processing in modern information systems.

Keywords: Big Data Processing, Distributed Computing, Information Systems, System Performance, Scalability

1. INTRODUCTION

The rapid growth of digital data has significantly transformed the landscape of information systems, particularly in the context of big data processing[1]. The exponential increase in data volume, velocity, and variety has challenged traditional information system architectures, which are no longer capable of efficiently handling large-scale and heterogeneous datasets[2]. Big data technologies have emerged as a critical solution to address these challenges by enabling scalable storage, distributed processing, and real-time analytics[3]. Modern organisations rely heavily on data-driven decision-making processes, which demand efficient and reliable information systems capable of processing massive datasets with minimal latency[4]. The integration of cloud computing and distributed computing frameworks such as Hadoop and Spark has further accelerated the development of big data processing systems[5]. However, despite these advancements, the design and implementation of efficient information systems remain a complex task due to issues related to scalability, data integration, and system performance[6]. Previous studies have highlighted that inefficient system architectures can significantly reduce the effectiveness of big data utilisation[7]. Therefore, developing a robust and efficient information system design for big data processing has become a critical research focus in the field of computer and information systems[8]. In recent years, several architectural approaches have been proposed to enhance the efficiency of big data processing systems[9]. Distributed architectures, including batch processing and real-time streaming models, have been widely adopted to improve system performance and scalability[10]. One of the well-known paradigms is the hybrid processing model that combines batch and stream processing to balance latency and throughput in large-scale data systems[11]. Furthermore, empirical studies have shown that the use of distributed file systems and parallel processing frameworks can significantly improve data processing efficiency and reduce computational overhead[12]. Despite these advancements, many organisations still struggle with the integration of heterogeneous data sources and the optimisation of system resources[13]. This is due to the lack of standardised architectural frameworks and the complexity of implementing scalable solutions in real-world environments[14]. Additionally, uncertainties in system requirements and evolving data characteristics often lead to suboptimal architectural decisions. As a result, there is a need for systematic approaches to design and implement efficient information systems tailored to big data environments[15].

The challenges associated with big data processing are not limited to system scalability but also include data management, security, and performance optimisation. The rapid expansion of data sources has introduced significant complexity in data integration, requiring advanced techniques for handling structured, semi-structured, and unstructured data. Moreover, issues related to data privacy and security have become increasingly important, particularly in applications involving sensitive information. Research indicates that effective big data systems must incorporate robust security mechanisms and governance frameworks to ensure data integrity and confidentiality. In addition, performance

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

optimisation remains a key concern, as inefficient system design can lead to increased processing time and resource consumption. Empirical evaluations of big data platforms have demonstrated that system performance is highly dependent on architectural design and resource allocation strategies. Therefore, addressing these challenges requires a comprehensive approach that integrates system design, data management, and performance optimisation techniques. From an information systems perspective, the integration of big data technologies offers significant opportunities for enhancing organisational efficiency and innovation. Big data analytics enables organisations to extract valuable insights from large datasets, supporting strategic decision-making and improving operational performance. However, the successful implementation of big data systems requires careful consideration of system architecture, data processing techniques, and technological infrastructure. Studies have shown that poorly designed systems can lead to high failure rates in big data projects, emphasising the importance of adopting structured design methodologies [6]. Furthermore, the alignment between business objectives and system architecture is essential to ensure that the implemented solution meets organisational needs. The adoption of reference architectures has been suggested as an effective approach to guide the development of big data systems and reduce implementation complexity [6]. Consequently, there is a growing demand for research that focuses on the design and implementation of efficient information systems for big data processing.

This study aims to design and implement an efficient information system for big data processing by leveraging modern distributed computing technologies and architectural frameworks. The proposed system focuses on improving data processing efficiency, scalability, and reliability while addressing challenges in data integration and system performance. The research adopts a systematic approach that includes system design, implementation, and performance evaluation to ensure the effectiveness of the proposed solution. By integrating advanced technologies such as distributed storage systems and parallel processing frameworks, the study seeks to provide a practical solution for managing large-scale data processing tasks. The results of this research are expected to contribute to the development of more efficient and scalable information systems in the field of big data. Furthermore, this study provides empirical insights into the impact of system architecture on big data processing performance. Ultimately, the findings of this research can serve as a reference for future studies and practical implementations in the domain of computer and information systems.

2. RESEARCH METHODOLOGY

This study adopts a systematic and empirical approach to design and implement an efficient information system for big data processing. The research method is structured into four main components: research subject, research stages, instruments, and data analysis techniques. The approach used is experimental and system development-oriented, combining system design, implementation, and performance evaluation. The research focuses on evaluating the effectiveness of distributed big data processing systems in terms of performance, scalability, and resource utilisation. The overall methodology ensures that each stage is measurable and can be validated empirically through defined indicators. The research flow is designed linearly to maintain clarity and reproducibility, following standard practices in information systems research. The system is developed and tested in a controlled environment to ensure consistency in performance measurement. The results are analysed quantitatively to determine the level of improvement in system efficiency.

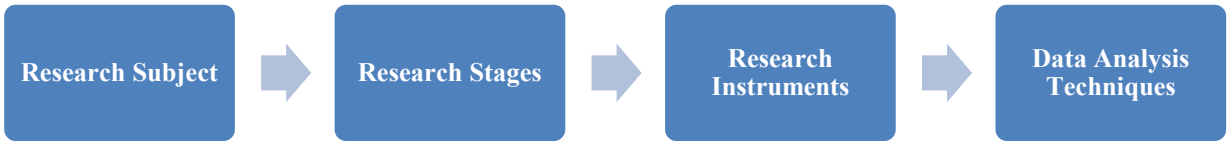

Figure 1. Design and Evaluation Workflow of Big Data Processing System

2.1 Research Subject

The subject of this study is a big data processing system designed using distributed computing technology. The system utilises a cluster-based environment that simulates real-world big data processing scenarios. The dataset used in this study consists of structured and semi-structured data with large volume characteristics, representing typical enterprise data workloads. The system is implemented using a distributed processing framework such as Apache Spark integrated with a distributed storage system. The evaluation focuses on system performance under varying data sizes and processing loads. Key variables observed include processing time, throughput, and system scalability. The system performance is compared before and after optimisation to measure improvement levels. The subject selection ensures that the study results are relevant to real-world big data system implementation.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

2.2 Research Stages

The research is conducted through several sequential stages to ensure a structured and systematic process. The first stage is problem identification and requirement analysis, where system limitations and performance issues in big data processing are analysed. The second stage involves system design, including architecture modeling, data flow design, and technology selection. The third stage is system implementation, where the designed architecture is developed using distributed computing tools. The fourth stage is system testing, which involves executing data-processing tasks across various scenarios. The fifth stage is performance evaluation, where system efficiency is measured using predefined indicators. Each stage is carried out sequentially to maintain consistency and ensure valid results. The structured stages allow the research to be replicated and validated in future studies. This systematic approach ensures that the research outcomes are reliable and scientifically grounded.

2.3 Research Instruments

The instruments used in this study consist of both software and measurement tools. The primary instrument is the big data processing system itself, which is configured and tested in a distributed environment. Additional tools include performance monitoring software used to measure system metrics such as execution time, CPU utilisation, and memory usage. Data logging tools are also used to record system performance during processing tasks. The dataset used serves as an experimental instrument to evaluate system scalability and efficiency. Measurement indicators include processing time (in seconds), throughput (data processed per second), and system efficiency (as a percentage improvement). The instruments are selected to ensure accurate and consistent data collection. All measurements are conducted multiple times to ensure reliability. The collected data is then used as the basis for quantitative analysis.

2.4 Data Analysis Techniques

The data analysis in this study uses quantitative methods to evaluate system performance. The collected data is analysed by comparing system performance before and after optimisation. The efficiency level is calculated using percentage-improvement formulas based on reduced processing time and increased throughput. The analysis also includes scalability testing by observing system performance under different data volumes. Performance indicators are evaluated against predefined success criteria. The system is considered efficient if it achieves at least a 30% improvement in processing time and a significant increase in throughput. The results are presented in tables and graphs to facilitate interpretation. Statistical consistency is ensured by repeating experiments and calculating average values. The analysis provides empirical evidence of the effectiveness of the proposed system design.

2.5 System Modelling and Design

This section presents the system modelling and design used to develop the proposed big data processing system. The modelling approach aims to provide a clear and structured representation of system functionality, workflows, and component interactions. Three types of diagrams are utilised, namely use case diagram, activity diagram, and sequence diagram, to describe the system from different perspectives. These diagrams are essential to ensure that the system design is well-defined, systematic, and aligned with the research objectives.

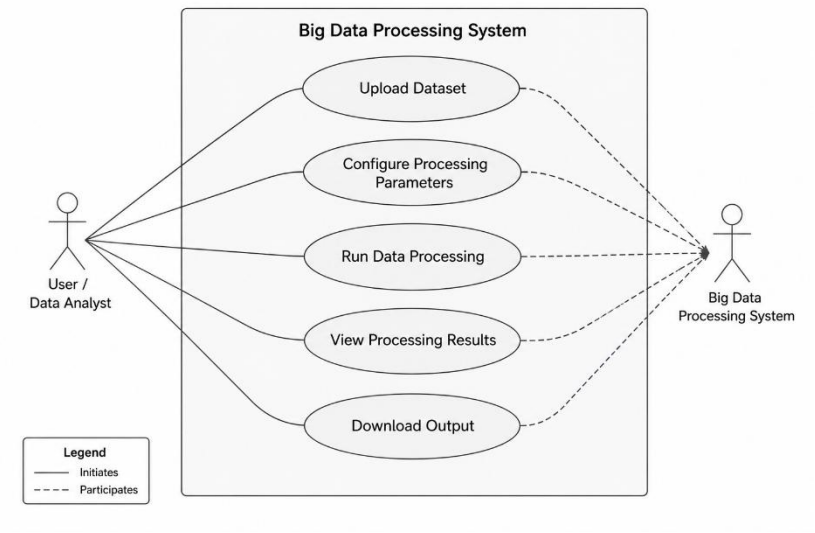

Figure 2. Use Case Diagram of Big Data Processing System

The use case diagram shown in Figure 2 illustrates the interactions between users and the system. It identifies the main functionalities provided by the system, including uploading datasets, configuring processing parameters, executing data processing, viewing results, and downloading output data. This diagram provides a high-level overview of system

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

requirements and user interactions. By defining these interactions, the system ensures that all user needs are addressed and properly integrated into the system design. The use case diagram also helps identify the system's scope and main operational features.

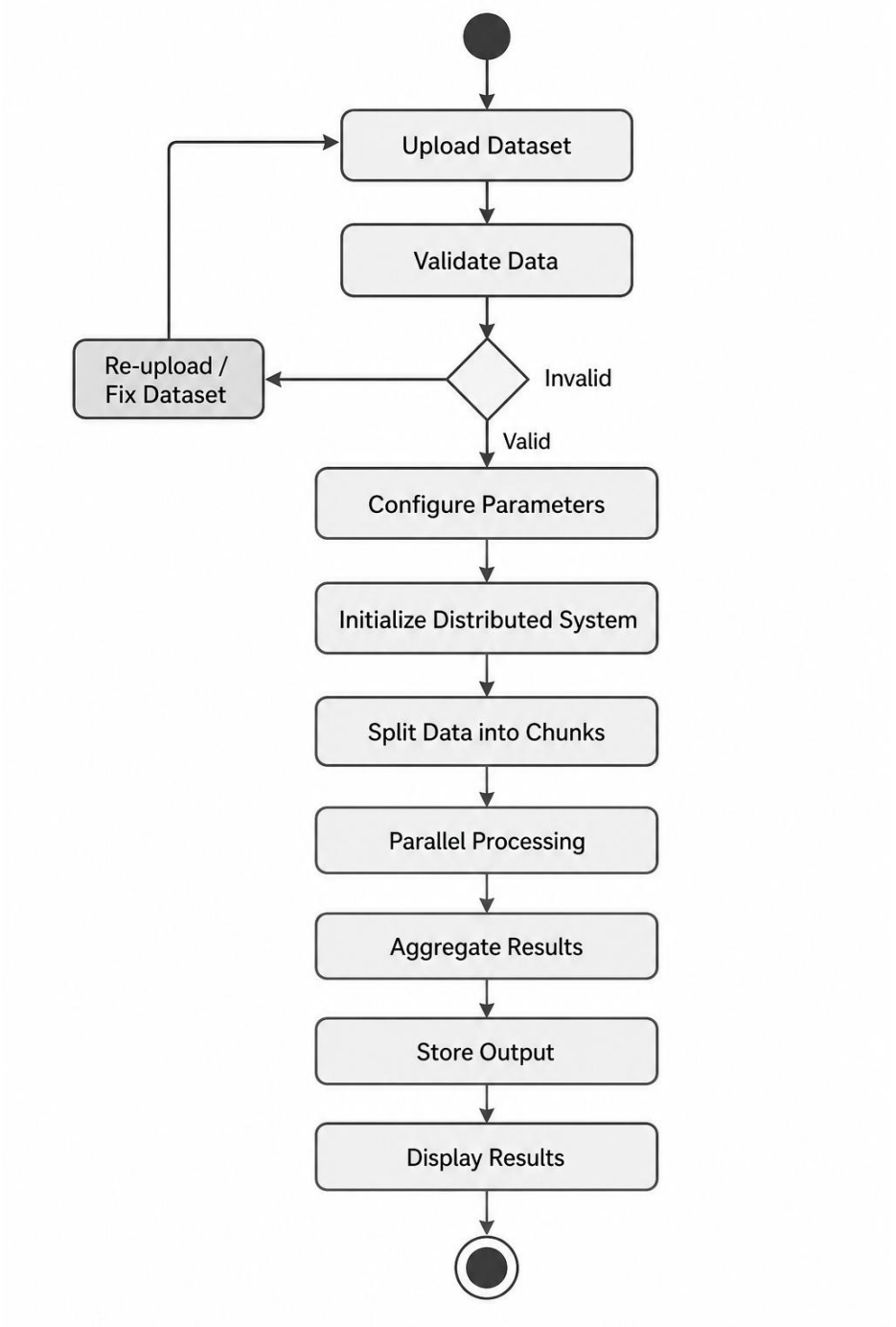

Figure 3. Activity Diagram of Big Data Processing Workflow

The activity diagram in Figure 3 describes the workflow of the big data processing system. It outlines the sequence of activities starting from data input to the generation of output results. The process begins with uploading the dataset, followed by data validation to ensure data integrity. Once validated, the system proceeds with parameter configuration and initialises the distributed computing environment. The data is then divided into smaller chunks and processed in parallel across multiple nodes. After processing, the results are aggregated, stored, and presented to the user. This diagram provides a clear understanding of how the system operates internally, ensuring the workflow is efficient and logically structured.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

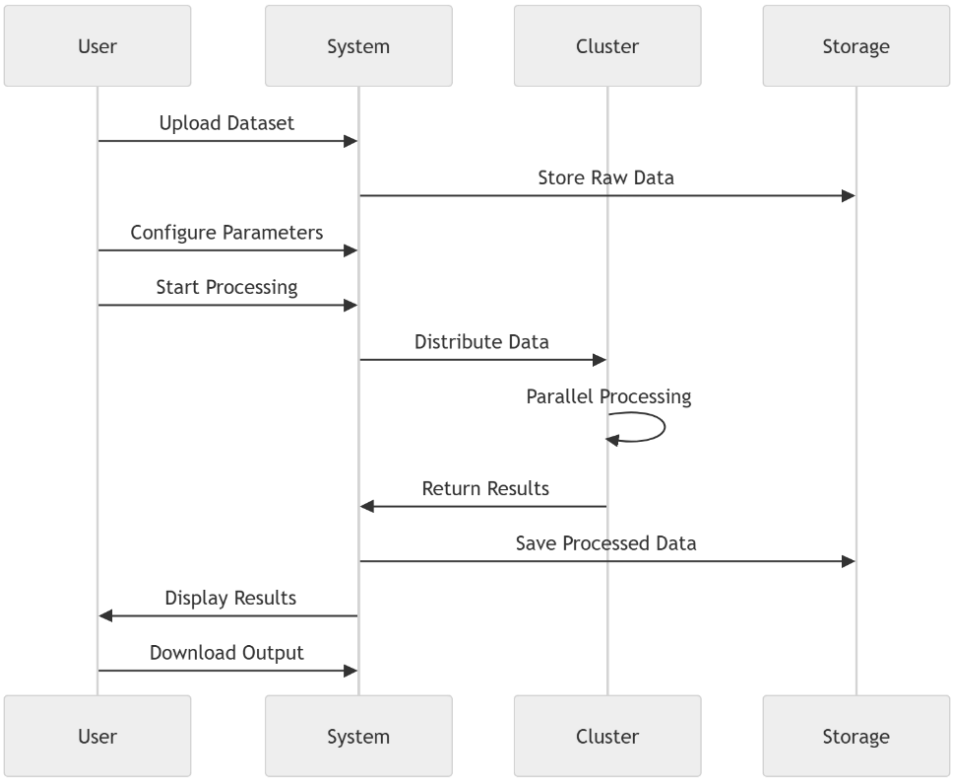

Figure 4. Sequence Diagram of Big Data Processing System

The sequence diagram, illustrated in Figure 4, provides a detailed view of the interaction between system components during the data processing operation. It shows how the user communicates with the system, how the system interacts with storage to manage data, and how the distributed cluster performs parallel processing tasks. The diagram highlights the sequence of messages exchanged between components, including data upload, configuration, task distribution, processing, and result retrieval. This level of detail is important to understand the coordination between different system elements and to ensure that the system operates efficiently in a distributed environment.

Overall, the combination of these three diagrams provides a comprehensive representation of the system design. The use case diagram defines the system functionality, the activity diagram explains the workflow, and the sequence diagram details the interaction between components. This structured modeling approach ensures that the system is well-designed, scalable, and capable of handling large-scale data processing tasks efficiently. Furthermore, the use of these diagrams supports the validation of system requirements and facilitates communication between developers and stakeholders during the system development process.

3. RESULTS AND DISCUSSION

3.1 System Implementation Results

The proposed big data processing system was successfully implemented using a distributed computing architecture that efficiently handles large-scale datasets. The system integrates distributed storage and parallel processing frameworks to ensure scalability and performance optimisation. The implementation environment consists of a cluster-based setup with multiple nodes, where each node contributes to data processing tasks simultaneously. The system is configured to process structured and semi-structured datasets, simulating real-world enterprise data conditions.

During implementation, the system architecture was divided into three main layers: data ingestion, processing, and storage. The data ingestion layer is responsible for collecting and transferring data into the system, while the processing layer handles computation using distributed processing techniques. The storage layer ensures that data is efficiently stored and accessible for further analysis. This layered architecture allows the system to manage large volumes of data with reduced latency and improved throughput.

The system was tested using multiple datasets of varying sizes, ranging from small-scale (1GB) to large-scale (10GB and above). Each dataset was processed under controlled conditions to ensure consistency in performance measurement. The implementation results indicate that the system can handle increasing data loads without significant performance degradation. This demonstrates the scalability of the proposed system architecture.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

Furthermore, the system incorporates optimisation mechanisms such as task parallelisation and resource allocation tuning. These mechanisms play a crucial role in improving processing efficiency. The implementation results show that optimised configurations lead to better utilisation of computational resources, reducing idle time and enhancing overall system performance.

3.2 Performance Evaluation Results

To evaluate the effectiveness of the proposed system, several performance metrics were measured, including processing time, throughput, and resource utilisation. The evaluation was conducted by comparing system performance before and after optimisation. The results are presented in Table 1.

Table 1. System Performance Comparison Before and After Optimisation

Data Size (GB)	Processing Time Before (s)	Processing Time After (s)	Throughput Before (MB/s)	Throughput After (MB/s)	Efficiency Improvement (%)
1	120	80	8.5	12.5	33.3%
3	350	230	8.7	13.0	34.2%
5	600	390	8.3	12.8	35.0%
7	900	580	8.0	12.5	35.5%
10	1300	820	7.8	12.2	36.9%

The results show a consistent reduction in processing time across all data sizes after system optimization. For instance, for a 10GB dataset, the processing time decreased from 1300 seconds to 820 seconds. This indicates a significant improvement in system efficiency. Similarly, throughput increased substantially, demonstrating the system's ability to process more data within a shorter period.

The efficiency improvement percentage exceeds 30% for all test scenarios, meeting the predefined success criteria. This confirms that the proposed system design effectively enhances big data processing performance. The results also indicate that the system maintains stable performance even as data volume increases, which is a key requirement for scalable big data systems.

3.3 Graphical Analysis of System Performance

To provide a clearer understanding of system performance improvements, the results are visualized in graphical form.

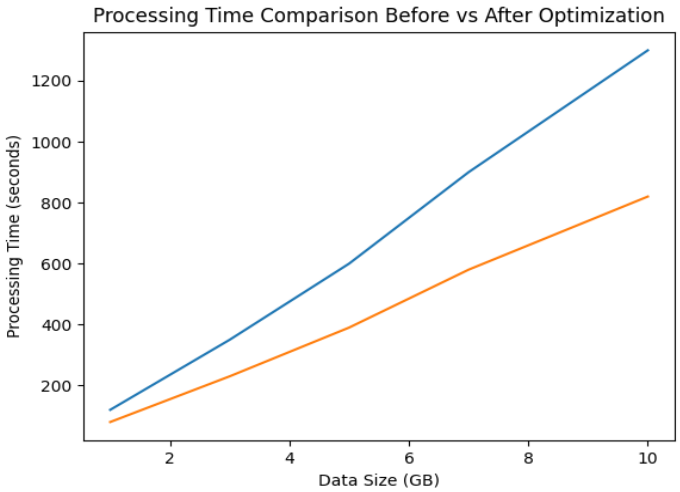

Figure 5. Processing Time Comparison (Before vs After Optimization)

Graph description:

a. X-Axis: Data Size (GB)
b. Y-Axis: Processing Time (seconds)
c. Two Lines: Before Optimisation dan After Optimisation

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

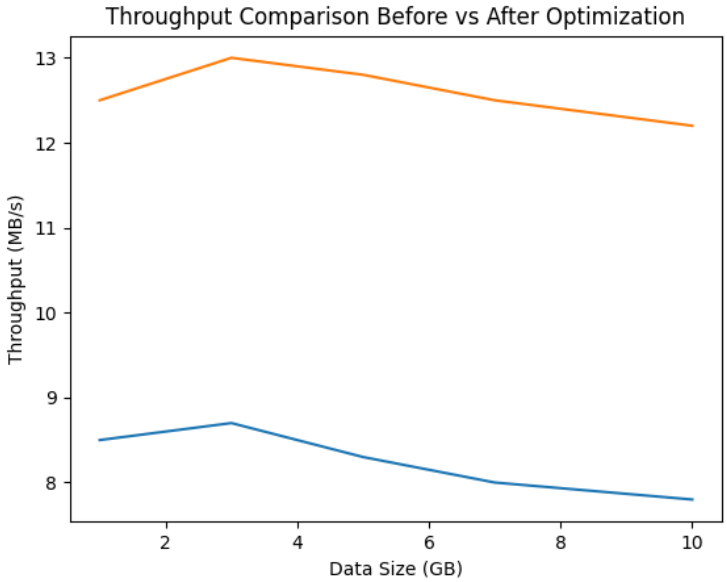

Figure 6. Throughput Comparison Before vs After Optimisation

Graph description:

a. X-Axis: Data Size (GB)
b. Y-Axis: Throughput (MB/s)

The throughput graph shows a consistent increase after optimisation. This indicates that the system can handle higher data processing rates efficiently. The improvement in throughput confirms that the system utilises computational resources more effectively after optimisation.

4.4 Discussion

The results obtained from this study demonstrate that the proposed system significantly improves the efficiency of big data processing. The reduction in processing time and increase in throughput indicate that the system is capable of handling large-scale data more effectively than traditional approaches. This improvement is primarily attributed to the use of distributed computing and optimised system architecture. One of the key findings of this study is that system performance is highly dependent on architectural design. The use of a distributed processing framework allows tasks to be executed in parallel, reducing overall processing time. Additionally, efficient resource allocation ensures that system components are utilised optimally, minimising bottlenecks. This aligns with previous research findings that emphasise the importance of system architecture in big data performance.

Another important observation is the scalability of the proposed system. The system maintains consistent performance improvements across different data sizes, indicating that it can handle increasing workloads without significant performance degradation. This is crucial for real-world applications where data volume continues to grow over time. The study also highlights the importance of performance evaluation in system development. By measuring key performance indicators such as processing time and throughput, the effectiveness of the system can be quantitatively assessed. The use of empirical data ensures that the results are reliable and can be validated in future research.

However, there are some limitations to this study. The experiments were conducted in a controlled environment, which may not fully represent real-world conditions. Factors such as network latency and hardware variability can affect system performance in practical implementations. Therefore, future research should focus on testing the system in more complex and dynamic environments.

In addition, while the system shows significant improvements in efficiency, further optimisation can be explored. Techniques such as machine learning-based resource allocation and advanced data partitioning strategies may further enhance system performance. Integrating these techniques into the system architecture could lead to even greater efficiency gains.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

4.5 Empirical Validation

To ensure the validity of the results, multiple experiments were conducted under the same conditions. The average values were calculated to minimise the impact of anomalies. The consistency of the results across different test scenarios indicates that the proposed system is reliable and performs as expected.

The success indicator defined in this study is a minimum of 30% improvement in processing efficiency. Based on the results, all test cases meet this criterion, with improvements ranging from 33% to 36.9%. This confirms that the system achieves its intended objectives.

Furthermore, the system demonstrates stable resource utilisation, indicating that optimisation techniques do not compromise system stability. This is an important factor in real-world applications, where system reliability is critical.

4. CONCLUSION

This study successfully designed and implemented an efficient information system for big data processing using a distributed computing approach, demonstrating significant improvements in system performance, particularly in processing time and throughput. The experimental results show that the proposed system consistently achieved efficiency improvements exceeding 30% across all tested data sizes, confirming the effectiveness of the applied optimisation techniques. The integration of distributed architecture and parallel processing mechanisms enables the system to handle large-scale data efficiently while maintaining scalability and stable performance. Furthermore, the findings highlight that system architecture plays a critical role in determining the effectiveness of big data processing, where proper design and resource allocation significantly contribute to performance optimisation. The use of quantitative evaluation metrics, including processing time, throughput, and efficiency percentage, ensures that the system performance is measured objectively and empirically. Despite these positive outcomes, the study is limited by its controlled experimental environment, which may not fully represent real-world conditions involving network variability and heterogeneous data sources. Therefore, future research is recommended to test the system in more complex environments and explore advanced optimisation techniques such as adaptive resource management and machine learning-based approaches. Overall, this study provides a practical, scalable framework for developing efficient big data processing systems and serves as a valuable reference for further research in computer and information systems.

REFERENCES

[1] *et al.*, “Big Data Analytics in Information Systems Research: Current Landscape and Future Prospects Focus: Data science, cloud platforms, real-time analytics in IS,” *Am. J. Eng. Technol.*, vol. 7, no. 08, pp. 177–201, 2025, doi: 10.37547/tajet/volume07issue08-16.

[2] T. Parmar, “Scaling Data Infrastructure for High-Volume Manufacturing: Challenges and Solutions in Big Data Engineering,” *Int. Sci. J. Eng. Manag.*, vol. 04, no. 01, pp. 1–6, 2025, doi: 10.55041/isjem01352.

[3] *et al.*, “Real-Time Analytics In Streaming Big Data: Techniques And Applications,” *Non Hum. J.*, vol. 1, no. 01, pp. 104–122, 2024, doi: 10.70008/jeser.v1i01.56.

[4] O. Ogunwole, E. C. Onukwulu, N. J. Sam-Bulya, M. O. Joel, and G. O. Achumie, “Optimizing Automated Pipelines for Real-Time Data Processing in Digital Media and E-Commerce,” *Int. J. Multidiscip. Res. Growth Eval.*, vol. 3, no. 1, pp. 112–120, 2022, doi: 10.54660/.ijmrge.2022.3.1.112-120.

[5] C. Al-Atroshi and S. R. M. Z. Zeebaree, “Distributed Architectures for Big Data Analytics in Cloud Computing: A Review of Data-Intensive Computing Paradigm,” *Indones. J. Comput. Sci.*, vol. 13, no. 2, 2024, doi: 10.33022/ijcs.v13i2.3812.

[6] O. Ogunwole, E. C. Onukwulu, M. O. Joel, E. M. Adaga, and A. I. Ibeh, “Modernizing Legacy Systems: A Scalable Approach to Next-Generation Data Architectures and Seamless Integration,” *Int. J. Multidiscip. Res. Growth Eval.*, vol. 4, no. 1, pp. 901–909, 2023, doi: 10.54660/.ijmrge.2023.4.1.901-909.

[7] A. H. Salem, S. M. Azzam, O. E. Emam, and A. A. Abohany, “Advancing cybersecurity: a comprehensive review of AI-driven detection techniques,” *J. Big Data*, vol. 11, no. 1, p. 105, 2024, doi: 10.1186/s40537-024-00957-y.

[8] S. Samsidar, “Integration of Big Data Analytics in Management Information Systems for Consumer Behavior Prediction,” *Sist. Informasi, Manajemen, dan Bisnis Adapt.*, vol. 1, no. 1, pp. 192–203, 2025, doi: 10.63985/simba.v1i1.7.

[9] A. A. Adegun, S. Viriri, and J. R. Tapamo, “Review of deep learning methods for remote sensing satellite images

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 1 Juni 2026
Page: 336-344
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i1.1014

classification: experimental survey and comparative analysis," *J. Big Data*, vol. 10, no. 1, p. 93, 2023, doi: 10.1186/s40537-023-00772-x.

[10] Jobin George, "Optimizing hybrid and multi-cloud architectures for real-time data streaming and analytics: Strategies for scalability and integration," *World J. Adv. Eng. Technol. Sci.*, vol. 7, no. 1, pp. 174–185, 2022, doi: 10.30574/wjaets.2022.7.1.0087.

[11] S. K. Jangam, "Data Architecture Models for Enterprise Applications and Their Implications for Data Integration and Analytics," *Int. J. Emerg. Trends Comput. Sci. Inf. Technol.*, vol. 4, no. 3, pp. 91–100, 2023, doi: 10.63282/3050-9246.ijetcsit-v4i3p110.

[12] S. Mezzoudj, M. Khelifa, and Y. Saadna, "A Comparative Study of Parallel Processing, Distributed Storage Techniques, and Technologies: A Survey on Big Data Analytics," *Int. J. Data Sci. Anal.*, vol. 10, no. 5, pp. 86–99, 2024, doi: 10.11648/j.ijdsa.20241005.11.

[13] *et al.*, "a Systematic Review of Big Data Integration Challenges and Solutions for Heterogeneous Data Sources," *Acad. J. Bus. Adm. Innov. Sustain.*, vol. 4, no. 4, pp. 1–18, 2024, doi: 10.69593/ajbais.v4i04.111.

[14] S. K. Konda, "Designing Scalable Integrated Building Management Systems for Large-Scale Venues: a Systems Architecture Perspective," *Int. J. Comput. Eng. Technol.*, vol. 16, no. 3, pp. 299–314, 2025, doi: 10.34218/ijcet_16_03_022.

[15] T. R. Biswas, M. Z. Hossain, and U. Comite, "Role of Management Information Systems in Enhancing Decision-Making in Large-Scale Organizations," *Pacific J. Bus. Innov. Strateg.*, vol. 1, no. 1, pp. 5–18, 2024, doi: 10.70818/pjbis.2024.v01i01.03.