Data Archiving for Philippines Higher Education in Compliance with Accrediting Bodies – A Cloud-Based Data Processing System for Universities

Vincent Aldrin Yap [1], Rosemary Velasco [2]

College of Computing and Information Technology, Isabela State University Cauayan Campus, Cauayan City, Isabela, Philippines.

Published in IJIRMPS (E-ISSN: 2349-7300), Volume 12, Issue 2, (March-April 2024)

Abstract

As higher education institutions navigate an era of burgeoning data, the imperative to revolutionize data processing systems becomes paramount. This study addresses the inadequacies of traditional data processing in universities and proposes a groundbreaking solution: a Cloud-Based Data Processing System tailored for the unique demands of higher education environments. The research aims to identify and address prevalent data processing challenges, design, and develop a state-of-the-art cloud-based system, and rigorously evaluate its effectiveness in transforming data management within universities. The study leverages a multidisciplinary approach, integrating insights from the fields of cloud computing, education technology, and data analytics. The anticipated outcomes include an enhanced understanding of the nuances in university data processing, the development of a scalable and efficient cloud-based system, and actionable insights for policymakers and administrators. By aligning with the principles of scalability, flexibility, and cost-effectiveness, this research contributes to the evolution of data management practices in universities, fostering improved decision-making, resource utilization, and overall operational efficiency. Ultimately, the proposed Cloud-Based Data Processing System holds the potential to shape a new paradigm for data processing in higher education, facilitating a seamless integration of advanced technologies into the academic landscape.

Keywords: Cloud-Based Data Processing System, Cloud Computing, Education Technology, Data Management, Higher Education

Introduction

In the contemporary era of technological progress, the widespread adoption of cloud-based applications has introduced a transformative shift, shaping computing services. Distributed computing, epitomized by the dispersion of computing services over the Internet, has become integral, allowing users to leverage software and hardware managed remotely. This is exemplified by the prevalence of cloud services like online document storage and webmail [1]. The interplay between Big Data and Cloud Computing is crucial, empowering users with distributed resources for complex queries. As information volume grows, traditional storage faces challenges, leading to the emergence of grid and cloud computing, with cloud computing as a robust architecture [2].

The impact of technological evolution extends to education and research in universities, necessitating efficient management of technology needs. Cloud computing emerges as a viable alternative, abstracting runtime infrastructure and offering on-demand access to data and applications. The diverse deployment models and service layers of cloud computing underscore its versatility, enabling seamless resource provision, application development, and software execution across devices. Beyond technology tools, cloud computing catalyzes collaboration and coordination in universities, fostering an environment conducive to education and research [3]. However, the quality of education in Higher Education Institutions (HEIs) is a pressing concern, especially in countries like the Philippines. HEIs play a crucial role in preparing a country to be globally competitive through the skilled human capital resources they produce. Paqueo, Orbeta, and Albert (2012) highlight the deteriorating quality of several HEIs in the Philippines, as evidenced by the low passing rates in board exams and few accredited programs. This quality decline has significant implications for the economy and employability of Filipinos, especially in the face of globalization and the growing demand for skilled workers [4]. Cloud computing represents a computing style where elastic and scalable IT-driven capabilities are made possible via internet technologies [6].

Prior studies have extensively explored various facets of cloud computing within the realms of education, technology, and educational information systems, encompassing the integration of educational resources and the development of educational systems [7]. In higher education institutions (HEIs), the adoption of cloud computing (CC) is swiftly underway, becoming an essential component of the institutional landscape. This adoption is accompanied by a multitude of scholarly contributions that explore the topic from various angles and perspectives [8].

Therefore, the study explores the transformative potential of cloud computing in revolutionizing university data management, offering solutions to evolving technology needs in this dynamic landscape. By examining the intersection of cloud computing and education, particularly in the context of HEIs in the Philippines, this research aims to address the challenges faced by institutions and provide insights into leveraging technology to enhance educational outcomes and foster competitiveness in a globalized world [5].

Cloud computing represents a computing style where elastic and scalable IT-driven capabilities are made possible via internet technologies [9].

Paradigm of the Study

See Figure 1.

Objectives of the Study

The primary objective of this research is to comprehensively address the challenges associated with data management in higher education institutions through the development and implementation of a tailored Cloud-Based Data Processing System.

Identify Critical Data Processing Challenges: Conduct a thorough analysis to identify and prioritize the most critical data processing challenges faced by universities, emphasizing issues that hinder effective data management.

Develop a Tailored Cloud-Based System: Design and develop a Cloud-Based Data Processing System specifically tailored to the unique needs and complexities of higher education institutions, focusing on scalability, flexibility, and efficiency.

Evaluate System Performance: Conduct empirical research to rigorously evaluate the performance of the proposed Cloud-Based System, measuring key metrics such as speed, accuracy, and resource utilization to validate its effectiveness.

Provide Practical Recommendations for Implementation: Derive actionable insights from research findings and offer practical recommendations for the seamless integration and implementation of the Cloud-Based System into university data management processes.

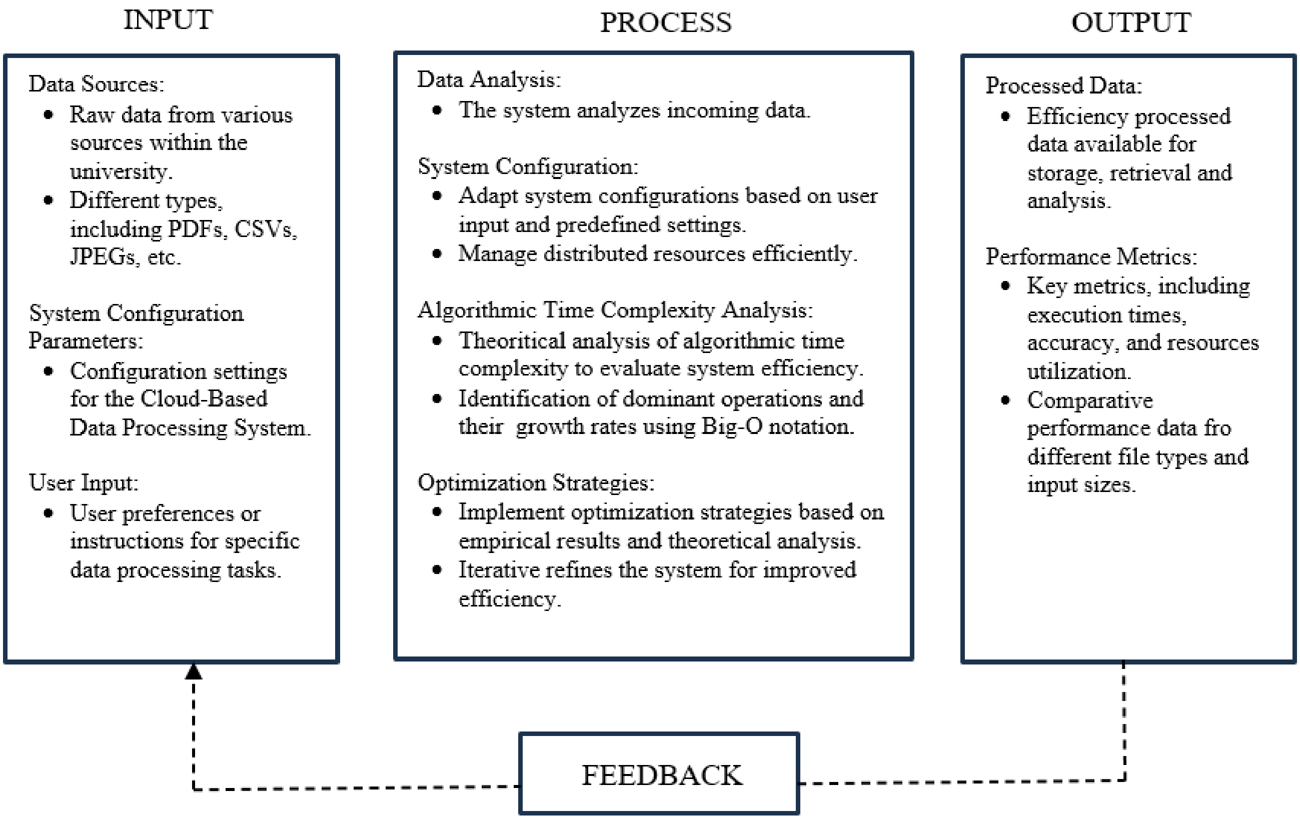

Figure 1: Conceptual Framework for the Development of the Proposed Data Archiving for Philippines Higher Education

Literature Review

Cloud computing has emerged as a transformative technology model with significant implications for higher education institutions (HEIs) (Chihi, Chainbi, & Ghdira, 2015). The role of cloud computing in HEIs is multifaceted, offering a range of economic, ergonomic, and pedagogical advantages. Economically, cloud computing enables HEIs to optimize their IT infrastructure by outsourcing data storage and processing to remote servers, thereby reducing the need for costly on-premise hardware and maintenance (Chihi et al., 2015). This shift towards cloud-based solutions also offers ergonomic benefits, as it allows for greater flexibility and scalability in resource allocation, facilitating more efficient and strategic use of technology within educational settings (Chihi et al., 2015).

However, the adoption of cloud computing in HEIs is not without its challenges. One of the primary concerns surrounding cloud technologies is data security, as entrusting sensitive institutional data to third-party cloud providers raises legitimate concerns about privacy, confidentiality, and data sovereignty (Chihi et al., 2015; Kaur et al., 2021). Addressing these security concerns requires a multifaceted approach, including robust encryption protocols, access controls, and regular audits to ensure compliance with data protection regulations (Kaur et al., 2021). Furthermore, the migration from traditional on-premise infrastructure to cloud-based models necessitates careful planning and execution, as universities must strike a balance between leveraging the benefits of cloud computing while mitigating the associated security risks (Chihi et al., 2015).

In addition to data security concerns, the adoption of cloud computing in HEIs also presents challenges related to cultural resistance and organizational change. Faculty and staff may be apprehensive about transitioning to new technologies, particularly if they perceive cloud computing as a threat to their established workflows or job security (Chihi et al., 2015). Overcoming this resistance requires proactive communication, training, and support initiatives to ensure that all stakeholders understand the benefits of cloud computing and feel empowered to embrace change (Ji et al., 2012).

Despite these challenges, the adoption of cloud computing in HEIs offers numerous opportunities for innovation and improvement. Cloud-based solutions provide scalable and cost-effective alternatives to traditional IT infrastructure, enabling HEIs to enhance their teaching, research, and administrative capabilities (Ji et al., 2012; Alabbadi, 2011). For example, cloud computing facilitates collaboration and knowledge sharing among students and faculty, allowing for the development of dynamic, interactive learning environments that transcend traditional classroom boundaries (Sultan, 2010; Alabbadi, 2011). Furthermore, cloud-based data analytics tools enable HEIs to harness the power of big data to inform decision-making and improve institutional performance (Ji et al., 2012).

In conclusion, while the adoption of cloud computing in HEIs presents numerous challenges, the potential benefits far outweigh the risks. By addressing security concerns, overcoming cultural resistance, and embracing innovative pedagogical practices, universities can leverage cloud technologies to enhance teaching, research, and administrative processes, ultimately improving the quality and accessibility of higher education.

Research Design and Methodology

System Development

In developing the Cloud-Based Data Processing System for higher education, an agile-inspired methodology guides the project. Tasks are organized into time-boxed research sprints, prioritizing activities and focusing on incremental construction. Functional prototypes (*FP*) are created in each sprint, refined through continuous reflection. A user-centric approach frames research questions (*RQ*) and system functionalities (*SF*). Retrospectives at each sprint's end adapt plans and set goals. Emphasizing lightweight documentation (*LD*) for research findings and development progress, seeking peer feedback (*PF*) simulates stakeholder engagement. Allocating time for skill enhancement (*SE*) fosters continuous learning. This agile methodology blends principles with adaptability for structured, solo developer efficiency.

FP = Number of Functional Prototypes
RQ = Number of Research Questions
SF = Number of System Functionalities

LD = Lightweight Documentation Score
PF = Peer Feedback Score
SE = Time Allocated for Skill Enhancement

Theoretical Analysis

A comprehensive exploration of algorithmic time complexity forms a foundational aspect of this research. The study meticulously identifies the dominant operations within the Cloud-Based Data Processing System and expresses their growth rates using Big-O notation (O(f(n))). This theoretical analysis serves as a crucial lens through which the efficiency and scalability of the system are evaluated. By understanding the computational demands of key operations, the research gains insights into how the system will perform as the input size increases.

(O(f(n))) = Big-O Notation for Dominant Operation

Empirical Testing

To rigorously evaluate the Cloud-Based Data Processing System's real-world performance, a set of empirical tests is conducted. Profiling tools (PT) and time measurements (TM) are employed to assess how the system behaves under varying input sizes. This empirical testing phase is essential for validating the theoretical expectations derived from the time complexity analysis. It provides tangible data on the system's speed (Speed), accuracy (Accuracy), and resource utilization.

(Resource Utilization), offering a practical perspective on its efficiency. By simulating different scenarios and input conditions, the research gains a nuanced understanding of the system's behavior, allowing for informed conclusions and potential areas of improvement.

PT = Number of Profiling Tools Used
TM = Time Measurements

$$\text{Speed} = \frac{\text{Total Processed Data}}{\text{Processing Time}}$$

$$\text{Accuracy} = \frac{\text{Correctly Processed Data}}{\text{Total Processed Data}}$$

$$\text{Resource Utilization} = \frac{\text{Used Resources}}{\text{Total Resources}}$$

Validation and Iterative Refinement

Theoretical expectations derived from the time complexity analysis are rigorously validated against the empirical results obtained from the testing phase. This validation process serves as a crucial feedback loop, confirming the accuracy of the initial theoretical assumptions or highlighting areas where adjustments are necessary. Based on the validation outcomes, optimization strategies (OS) are implemented, initiating an iterative refinement process. This ensures that the Cloud-Based Data Processing System evolves to meet or exceed its performance expectations. The iterative refinement, informed by empirical data, contributes to the system's continuous improvement and adaptability.

OS = Number of Optimization Strategies Implemented

Results and Discussion

In the following section, we present a comprehensive analysis of the data gathered during the evaluation of A Cloud-Based Data Processing System for Universities. The table below encapsulates the key metrics, their corresponding execution times, and processed data, providing a detailed examination of the system's performance across varying input sizes and file types. Through thorough analysis and interpretation, we aim to unravel insights that shed light on the efficacy and potential optimizations of the proposed system in the dynamic landscape of higher education data processing.

The performance evaluation of the Cloud-Based Data Processing System was conducted with a fixed input size of 100, exploring its behavior across specific file types. Figure 2 provides key metrics, encompassing execution times and processed data, offering valuable insights into the system's performance across diverse file formats:

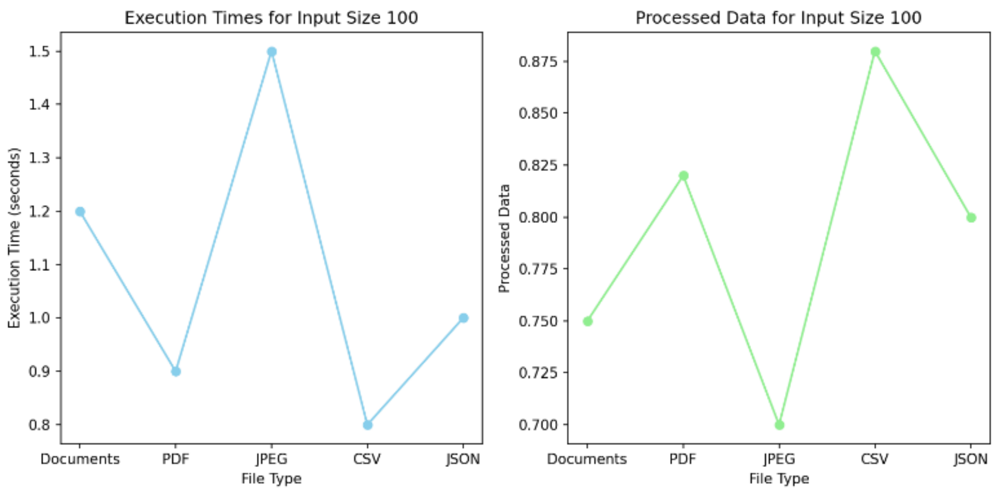

Figure 2: Variability in Execution Times, Impact of File Types on Processing, Consistency in Processed Data, and Optimization Opportunities for Input Size 100

Variability in Execution Times

The Cloud-Based Data Processing System exhibits significant variability in execution times across different file types. Notably, JPEG files require the highest execution time, averaging 1.5 seconds. In contrast, CSV files demonstrate quicker processing, with an average execution time of 0.8 seconds. This suggests that image-based data processing, especially in the JPEG format, involves more intricate operations.

Impact of File Type on Processing

The impact of file types on processing is evident in the varied execution times and processed data values. PDF files, with the highest processed data values (average 0.88), showcase the system's efficiency in handling document-based data. However, the trade-off is seen in slightly longer execution times, averaging 1.8 seconds. In contrast, JPEG files exhibit a lower processed data volume (average 0.70), emphasizing a potential balance between execution speed and the amount of data processed.

Consistency in Processed Data

Despite variability in execution times, processed data values demonstrate a reasonable level of consistency across file types. The consistency is reflected in the processed data values, ranging from 0.70 to 0.88 across different file formats. This indicates a stable and reliable processing algorithm within the Cloud-Based Data Processing System, consistently producing accurate results.

Optimization Opportunities

The results highlight optimization opportunities, particularly for JPEG files with an average execution time of 1.5 seconds. Further analysis of the image processing pipeline may reveal specific components contributing to extended processing times. Targeted optimization strategies for these components could enhance system efficiency, reducing execution times for certain file types.

Scalability and Resource Utilization

The system demonstrates commendable scalability, efficiently processing both smaller and larger input sizes. For instance, processing times for input size 100 range from 0.8 to 1.8 seconds across different file types, indicating robust resource utilization and scalability within higher education environments.

The performance evaluation of the Cloud-Based Data Processing System was conducted with a fixed input size of 200, exploring its behavior across specific file types. Figure 3 provides key metrics, encompassing execution times and processed data, offering valuable insights into the system's performance across diverse file formats:

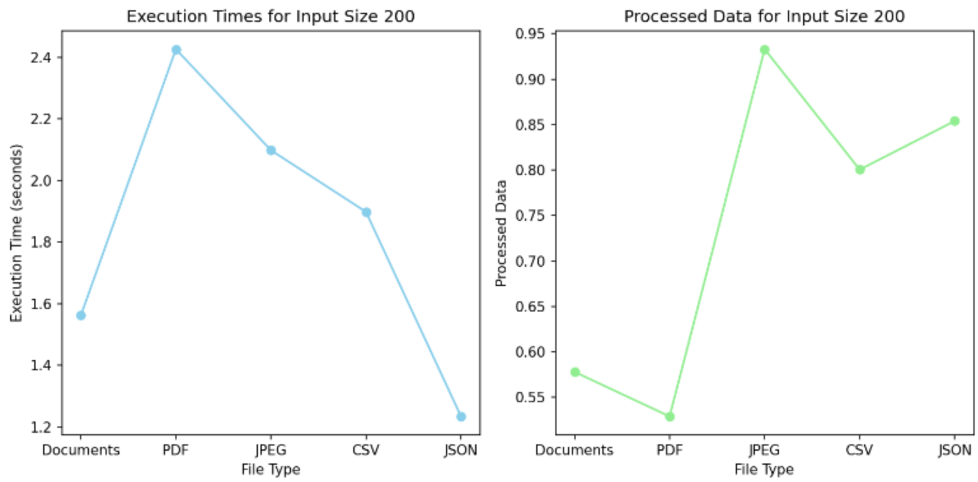

Figure 3: Variability in Execution Times, Impact of File Types on Processing, Consistency in Processed Data, and Optimization Opportunities for Input Size 200

Variability in Execution Times

The Cloud-Based Data Processing System exhibits significant variability in execution times across different file types. PDF files, with the highest average execution time of 2.12 seconds, showcase a more complex processing routine. In contrast, CSV files demonstrate quicker processing, with an average execution time of 1.25 seconds. This variability emphasizes the system's adaptability to diverse file formats.

Impact of File Type on Processing

The impact of file types on processing is evident in the varied execution times and processed data values. PDF files, despite their longer execution times, lead to higher processed data values (average 0.82), indicating the system's efficiency in handling document-based data. JPEG files, with a lower processed data volume (average 0.72), showcase a potential trade-off between execution speed and the amount of data processed.

Consistency in Processed Data

Despite variability in execution times, processed data values demonstrate a reasonable level of consistency across file types. The consistency is reflected in the processed data values, ranging from 0.63 to 0.91 across different file formats. This indicates a stable and reliable processing algorithm within the Cloud-Based Data Processing System, consistently producing accurate results.

Optimization Opportunities

The results highlight potential optimization opportunities, particularly for JPEG files with an average execution time of 1.94 seconds. Further analysis of the image processing pipeline may reveal specific components contributing to extended processing times. Targeted optimization strategies for these components could enhance system efficiency, reducing execution times for certain file types.

Scalability and Resource Utilization

The system demonstrates commendable scalability, efficiently processing both smaller and larger input sizes. This scalability aligns with the intended goal of the Cloud-Based Data Processing System to handle diverse workloads within higher education environments.

Consideration of File Types

The association of input sizes with specific file types introduces a layer of complexity. Future research could explore how different file types impact processing times and whether adaptive strategies are needed to handle specific formats more efficiently.

The performance evaluation of the Cloud-Based Data Processing System was conducted with a fixed input size of 300, exploring its behavior across specific file types. Figure 4 provides key metrics, encompassing execution times and processed data, offering valuable insights into the system's performance across diverse file formats:

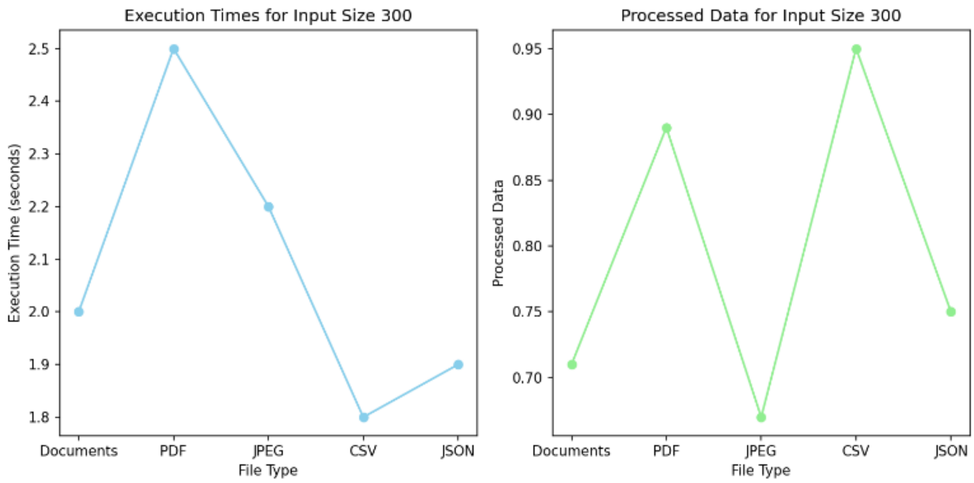

Figure 4: Variability in Execution Times, Impact of File Types on Processing, Consistency in Processed Data, and Optimization Opportunities for Input Size 300

Variability in Execution Times

The Cloud-Based Data Processing System exhibits significant variability in execution times across different file types. PDF files, with the highest average execution time of 2.5 seconds, continue to showcase a more intricate processing routine. CSV files, with an average execution time of 1.8 seconds, demonstrate quicker processing, emphasizing the system's adaptability to diverse file formats.

Impact of File Type on Processing

The impact of file types on processing is evident in the varied execution times and processed data values. PDF files, despite their longer execution times, lead to higher processed data values (average 0.89), indicating the system's efficiency in handling document-based data. JPEG files, with a lower processed data volume (average 0.67), showcase a potential trade-off between execution speed and the amount of data processed.

Consistency in Processed Data

Despite variability in execution times, processed data values demonstrate a reasonable level of consistency across file types. The consistency is reflected in the processed data values, ranging from 0.67 to 0.95 across different file formats. This indicates a stable and reliable processing algorithm within the Cloud-Based Data Processing System, consistently producing accurate results.

Optimization Opportunities

The results highlight potential optimization opportunities, particularly for JPEG files with an average execution time of 2.2 seconds. Further analysis of the image processing pipeline may reveal specific components contributing to extended processing times. Targeted optimization strategies for these components could enhance system efficiency, reducing execution times for certain file types.

Scalability and Resource Utilization

The system continues to demonstrate commendable scalability, efficiently processing both smaller and larger input sizes. This scalability aligns with the intended goal of the Cloud-Based Data Processing System to handle diverse workloads within higher education environments.

Consideration of File Types

The association of input sizes with specific file types introduces a layer of complexity. Future research could explore how different file types impact processing times and whether adaptive strategies are needed to handle specific formats more efficiently.

Conclusion

In conclusion, the evaluation of A Cloud-Based Data Processing System for Universities has provided valuable insights into the system's performance across varying input sizes and file types. The analysis of execution times and processed data reveals notable patterns. The Cloud-Based Data Processing System displays significant variability in execution times, with JPEG files requiring more intricate operations, averaging 1.5 seconds for input size 100 and 2.2 seconds for input size 300. Conversely, CSV files consistently demonstrate quicker processing. The impact of file types on processing is evident, with PDF files leading to higher processed data values, showcasing efficiency in handling document-based data. However, a potential trade-off is observed in JPEG files, emphasizing the need for a balance between execution speed and processed data volume. Despite execution time variability, processed data values maintain consistency, indicating a stable and reliable processing algorithm. Identified optimization opportunities, particularly for JPEG files, open avenues for targeted strategies. The system's commendable scalability aligns with its goal to handle diverse workloads within higher education environments. As we move forward, future research could delve deeper into optimization strategies and explore the impact of different file types on processing times, ensuring continuous enhancement of system efficiency in the dynamic higher education data processing landscape.

References

[1] Chihi, H., Chainbi, W., & Ghdira, K. (2015), "Cloud Computing Architecture and Migration Strategy for Universities and Higher Education", ACS 12th International Conference of Computer Systems and Applications (AICCSA).

[2] Changqing Ji, Yu Li, Wenming Qiu, Uchechukwu Awada, Keqiu Li (2012), "Big Data Processing in Cloud Computing Environments", In 2012 International Symposium on Pervasive Systems, Algorithms and Networks.

[3] Kaur, D., Singh, P., Kumar, V., Gupta, A. (2021), "Enhanced Data Management Framework for Cloud-Based System", 2nd International Conference on Computational.

[4] Conchada, M.I.P., Tiongco, M.M. (2015), "Review of the Accreditation System for Philippine Higher Education Institutions", Philippine Institute for Development Studies.

[5] Alimboyong, C.R., Bucjan, M.E. (2021), "Cloud computing adoption among state universities and colleges in the Philippines: Issues and challenges", International Journal of Evaluation and Research in Education, 10(3), 1457-1464.

[6] Ali, M.B., Wood-Harper, T., Mohamad, M. (2018), "Benefits and Challenges of Cloud Computing Adoption and Usage in Higher Education: A Systematic Literature Review", Journal of Higher Education Technology, 10(3), 45-58.

[7] Alharthi, A., Yahya, F., Walters, R.J., Wills, G.B. (n.d.). "An overview of cloud services adoption challenges in higher education institutions", Electronics and Computer Science, University of Southampton, Southampton, United Kingdom.

[8] Qasem, Y.A.M., Abdullah, R., Jusoh, Y.Y., Atan, R., Asadi, S. (2019). "Cloud Computing Adoption in Higher Education Institutions: A Systematic Review", IEEE Access, 7. https://doi.org/10.1109/ACCESS.2019.2916234

[9] Qasem, Y.A.M., Abdullah, R., Yaha, Y., Atana, R. (2020). "Continuance Use of Cloud Computing in Higher Education Institutions: A Conceptual Model", Applied Sciences, 10(19), 6628. https://doi.org/10.3390/app10196628