Knowledge-Based Systems 310 (2025) 112871

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Data-driven approach for selection of on-chain vs off-chain carbon credits data storage methods

Nada Atetallah Alghanmi [a,c], Nouf Atiahallah Alghanmi [e], Samaher Atiatallah Alghanmi [b], Ming Zhao [d], Farookh Khadeer Hussain [a,*]

[a] University of Technology Sydney, 15 Broadway Ultimo, Sydney, 2007, NSW, Australia
[b] King Abdulaziz City for Science and Technology, King Abdullah Rd, Al Raid, Riyadh, 11442, Saudi Arabia
[c] University of Jeddah, 34 W, Al Faisaliyyah, Jeddah, 23445, Makkah, Saudi Arabia
[d] University of New South Wales, Canberra, 2600, ACT, Australia
[e] King Abdulaziz University, Rabigh, 21911, Makkah, Saudi Arabia

ARTICLE INFO

Keywords:
Blockchain
Carbon credits
Data storage
On-chain
Off-chain
Blockchain data management

ABSTRACT

As blockchain technology continues to revolutionize various industries, the efficient and secure storage of data has become a critical challenge. Inspired by blockchain's distributed peer-to-peer networks, security and information storage solutions have found widespread application in high-information environments due to their efficiency, security, traceability, and decentralized nature. However, as blockchain applications expand and the volume of data grows exponentially, the interplay between on-chain and off-chain storage is rapidly becoming a crucial factor in blockchain data management. The motivation for this research stems from the pressing need to address the scalability and performance issues inherent in blockchain systems as they handle increasingly large datasets. Carbon credits data, specifically from a provenance perspective, is multimodal in nature. It can be captured using many different modes and devices and could include images, text, and unstructured data.

Current literature lacks a comprehensive analysis of the trade-offs between on-chain and off-chain storage methods, particularly for multimedia data, in terms of their impact on system scalability, operational efficiency, and economic viability. This gap in knowledge hinders stakeholders from making informed decisions about optimal data disposition strategies for their blockchain implementations.

This paper aims to bridge this gap by delivering a comprehensive explanation of the data storage methods adopted by blockchain systems, focusing on scalability, efficiency, and economic viability as key evaluation criteria. We analyze the unique connotations and implications of both on-chain and off-chain storage, considering their effects on data retrieval times and overall system performance. Through our analysis, we derive valuable insights to guide the decision-making process in designing and implementing blockchain solutions tailored to specific needs and application scenarios.

By providing a thorough examination of these storage methods, this research enables stakeholders to make informed choices about blockchain data disposition, ensuring the long-term sustainability and efficiency of their blockchain implementations. Our findings will contribute to the optimization of blockchain systems, balancing performance, security, and cost-effectiveness in an era of rapidly expanding blockchain applications and data volumes.

1. Introduction

Since its creation, blockchain has been applied as a provenance mechanism. But a crucial property of blockchain technology, in addition to provenance, is that data cannot be stored or changed **without ** undergoing a multi stakeholder consensus procedure. This consensus creates mining time, which can be substantial. However, information can be saved off-chain as an alternative to avoid this mining time. This has its downsides. Blockchain has evolved into the game-changing innovation in the continuously changing digital technological landscape, and its immediate and profoundly transformative statement of relevance has been felt across various sectors due to its inherent

* Corresponding author.
E-mail addresses: Nada.alghanmi@uts.edu.au (N.A. Alghanmi), naalganmy@kau.edu.sa (N.A. Alghanmi), Salghanmi@kacst.gov.sa (S.A. Alghanmi), ming.zhao@unsw.edu.au (M. Zhao), farookh.hussain@uts.edu.au (F.K. Hussain).

https://doi.org/10.1016/j.knosys.2024.112871
Received 18 June 2024; Received in revised form 19 November 2024; Accepted 10 December 2024
Available online 26 December 2024

characteristics of security, traceability, efficiency, and decentralization in developing reinforced solutions towards storage and protection of information. With the increased number of **applications**, users, and volume of data, the interactions of off-chain storage methods used for blockchain systems become ever more critical to comprehend and manage [1].

Managing the scalability and effectiveness of blockchain technologies involves the strategic management of data. There are both benefits and limitations to each of these methods of storage that directly influence performance and sustainability. In this sense, off-chain storage touches the opposite side which equals cost reduction and reduction of the basic load on the blockchain [2,3].

The work presented in this article is an in-depth consideration of the mechanisms of data disposal in blockchain-based systems with respect to their scalability, efficiency, and financial viability as the basic considerations for the choice of storage solutions. We approach this analysis in a structured, reasoned way to look at trade-offs in on-chain and off-chain carbon-credit storage across key dimensions in performance, security, cost, and complexity. The research will throw light on practical insights that will enable stakeholders to base their decisions on the methods of blockchain data disposition for typical application scenarios and needs.

The provenance of carbon credits is a significant research issue at present. Researchers are working on using blockchain technology to address this problem. For the purpose of verifying carbon credits, data are collected from a diverse range of resources such as sensors, drones (images), and other sources. The collected data comprise different types, including videos, images, and both structured and unstructured data. There is a lack in the existing literature focusing on comparing on-chain and off-chain storage approaches for such multi-modal carbon credit data. This represents a crucial gap in the current body of knowledge.

This project aims to address this gap in a systematic manner. The theoretical contribution of this research lies in providing a systematic, data-driven approach to objectively compare on-chain versus off-chain methods. As this is the first work to focus on a systematic, data-driven, and objective comparison between on-chain and off-chain methods, practitioners can use our approach in the future to make reliable decisions in carbon credit data storage. Design decisions, such as whether data should be stored on-chain, off-chain, or through a hybrid approach, will be of profound significance. If, for example, in carbon trading, data has to be accessed in real time, then the right design decisions can be made to ensure efficiency and access.

This paper is structured as follows. Section 1 introduces the subject and sets the stage for discussion. Section 2 provides an exhaustive transversal literature review with reference to previous research regarding the role of blockchain in carbon credit schemes and comparatively further explains the on-chain and off-chain methodologies. The most important findings and new developments identified in this research study are detailed in Section 3. The detailed description of the data collection processes and the techniques applied to this study is further elaborated in Section 4. Section 5 is the proposed system design and its implementation to show the working of the theoretical model. Section 6 explains the position of assessment metrics for the performance of algorithms in on-chain and off-chain. Section 7 further elaborates on the comparative assessment in terms of efficiency and implication of these methods. Finally, Section 8 concludes the research work and provides directions for future research development on this technology. At the end of the paper, stakeholders will have complete information to make the right decision about the long-term sustainability and efficiency of their blockchain-based systems.

This research marks a significant advancement in the field of blockchain technology, particularly in its application to carbon credit systems. The primary contributions of this study are summarized as follows:

Comprehensive Empirical Evaluation: This study represents the first comprehensive empirical evaluation of on-chain versus off-chain storage approaches, specifically within the context of carbon credit systems. Unlike the existing literature that offers limited or segmented analysis, this research examines these approaches across a multitude of dimensions including security, scalability, latency, and cost. This holistic evaluation is crucial for stakeholders to make informed decisions and tailor blockchain solutions to meet specific market needs effectively.

Identification of a Critical Research Gap: A significant contribution of this work is the identification and **address** of a notable gap in the existing literature — the lack of a thorough and holistic evaluation of on-chain versus off-chain approaches across various parameters. This gap has been overlooked in previous studies, which have not fully explored the implications of these storage methods on decision-making processes within the carbon credit domain.

Innovative Framework for Data Management: Based on the empirical evaluation, this study proposes a novel framework for optimizing data management strategies in carbon credit markets. This framework is designed to guide stakeholders in adopting blockchain solutions that are not only efficient and scalable but also secure and cost effective.

Clarification on Blockchain Application in Carbon Credits: The research provides clarity on the application of blockchain technology in carbon credit systems, a topic that has been ambiguously addressed in previous studies. It delineates the use of on-chain and off-chain methods, offering insights into their advantages and limitations within this specific context.

Basis for Future Research and Implementation: By highlighting the strengths and weaknesses of on-chain and off-chain approaches and proposing an optimized framework, this study lays a solid foundation for future research and practical implementation of blockchain technology to improve the integrity and efficiency of carbon credit systems.

In conclusion, the contributions of this paper significantly enrich the academic discourse and practical understanding of blockchain applications in environmental sustainability efforts. It paves the way for a more informed and effective use of blockchain technology to address the challenges of carbon credit trading and environmental conservation.

2. Related works

In addressing the intricacies of blockchain technologies for carbon credit systems, our investigation clearly contrasts off-chain and on-chain solutions, pinpointing a notable gap in empirical examinations within this specialized field. Off-chain mechanisms, exemplified by the Lightning Network and analogous initiatives [4–7], offer scalable, efficient transaction capabilities beyond blockchain's inherent limitations. These strategies, leveraging smart contracts, significantly reduce operational costs and transaction times, enhancing interoperability across blockchain platforms.

In contrast, on-chain solutions emphasize blockchain's core attributes of security and transparency in transactions across critical sectors like finance and insurance [8–11]. The decentralized nature of the blockchain boosts the security and efficiency of financial transactions. Nevertheless, scalability challenges persist, exacerbated by blockchain's limited capacity and network congestion. Innovations and optimizations, including block structure and storage rules alongside second-layer protocols such as the Lightning Network [12], are imperative to mitigate these issues.

Our contribution to the blockchain discourse, through a rigorous empirical analysis of on-chain versus off-chain storage methods in carbon credit systems, addresses a critical literature gap. This comprehensive evaluation, which covers the security, scalability, latency, and cost dimensions, marks a pioneering effort in the domain.

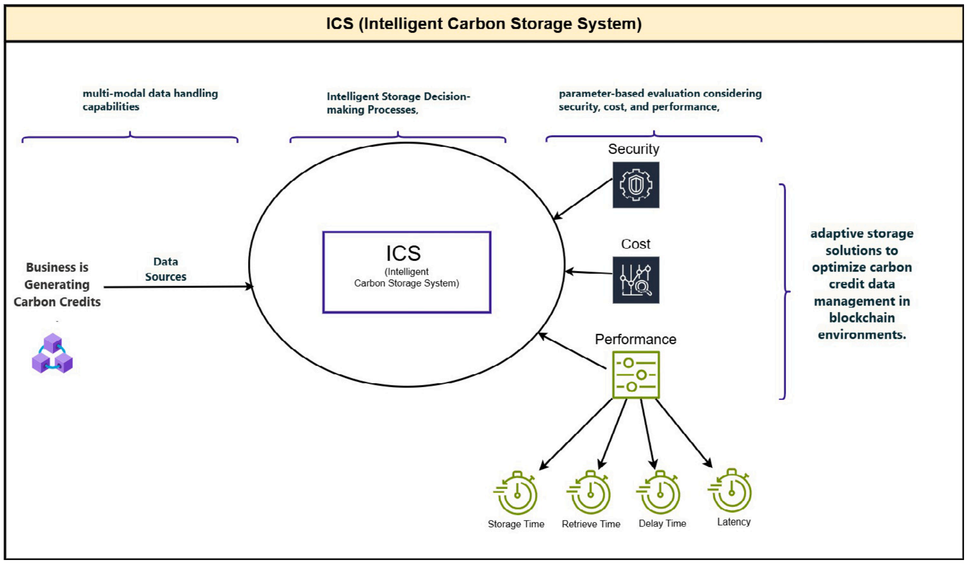

Fig. 1. ICS (Intelligent Carbon Storage System).

2.1. Blockchain in carbon projects

Blockchain's role in the carbon credit ecosystem is transformative, enhancing system transparency, accessibility, and efficiency [13]. The integration of this technology facilitates fraud-resistant carbon trading through decentralized governance and smart contracts [14], aligning with global climate change mitigation objectives. Moreover, the evolving maturity of blockchain projects within carbon markets highlights the urgent need for regulatory clarity to promote wider adoption and market globalization [15].

The applications of blockchain extend to the reduction efforts of carbon emissions in India [16], the development of eco-industrial parks [17], and the strides towards carbon neutrality [18]. Blockchain supports sustainable industrial practices, simplifies data management, and promotes governance. Furthermore, its role in facilitation of green investment within supply chains [19] and the environmental implications of cryptocurrency mining [20] present a nuanced view of the environmental impact of blockchain.

This narrative underscores blockchain's multifaceted applications in fostering environmental sustainability, particularly in carbon credit trading, advocating for a technology-driven, unified approach to climate change and sustainability challenges. Our study, through its novel empirical evaluation of on-chain versus off-chain approaches, aims to fill the existing research void, contributing significantly to the discourse on blockchain technology in environmental conservation efforts.

2.2. Data storage and integrity

Blockchain technology revolutionizes data integrity and security in digital ecosystems through its decentralized, tamperproof design, making it pivotal for applications such as legal documents, cloud storage, and off-chain transactions. This innovation underscores the critical role of data storage and integrity within blockchain environments to ensure network trustworthiness and reliability. The immutability of the blockchain ledger is essential to maintain the authenticity of the information, crucial to the integrity of legal documents [21]. Furthermore, the synergy between on-chain and off-chain data storage solutions offers a strategic approach to balance scalability with privacy, improving data governance frameworks [22].

However, the blockchain domain faces scalability challenges against the backdrop of increased data storage demands, complicating data integrity and security management across on-chain and off-chain platforms. Furthermore, the adaptation of customized on-chain governance mechanisms remains a hurdle that requires robust and integrated data management frameworks [22].

In conclusion, while blockchain serves as a formidable platform for data storage and integrity, addressing its scalability and governance challenges through ongoing innovation and collaboration is imperative to maximize its potential for secure, transparent, and efficient data management.

3. Conceptual framework of ICS (intelligent carbon storage system)

In this section, we introduce the Intelligent Carbon Storage (ICS) system, a novel approach to managing and storing carbon credit data. The ICS is designed to address the complex challenges associated with carbon credit provenance and data management in blockchain environments.

The ICS Fig. 1 serves as a central system for storing, managing, and accessing carbon credit data. It caters to various stakeholders, including individuals and businesses generating carbon credits, as well as those accessing and verifying this information. The system is designed to handle the multi-dimensional nature of carbon credit data, which can come from diverse sources such as drones, sensors, and other data collection methods. This data can take various forms, including videos, photographs, and both structured and unstructured data.

3.1. Key features of ICS

1. Multi-modal Data Handling: ICS is capable of processing and storing a wide range of data types associated with carbon credits, ensuring comprehensive coverage of all relevant information.
2. Intelligent Storage Decision-making: The core functionality of ICS lies in its ability to make informed decisions about data storage. It analyzes multiple factors to determine the optimal storage method (on-chain or off-chain) for each piece of data.
3. Parameter-based Evaluation: ICS evaluates storage options based on several critical parameters:
    - Security: Assessing the level of data protection required
    - Cost: Considering the financial implications of different storage solutions
    - Performance: Analyzing data retrieval speeds and overall system efficiency

- Adaptive Storage Solutions: Based on its analysis, ICS recommends the most suitable storage option, whether it is on-chain, off-chain, or a hybrid approach.

3.2. ICS workflow

1. Data Input: Carbon credit generators or businesses input their data into the ICS.
2. Parameter Analysis: The system evaluates the input data against predefined parameters (security, cost, performance).
3. Storage Decision: Based on the analysis, ICS determines the optimal storage method.
4. Implementation: The system executes the storage decision, placing data either on-chain or off-chain as appropriate.
5. Accessibility: Stakeholders can access the stored data as needed, with ICS managing retrieval based on the storage method used.

The ICS framework represents a significant advancement in blockchain-based carbon credit management. By providing a data-driven, objective approach to storage decisions, it addresses key challenges in the field:

- Optimizing resource utilization in blockchain environments
- Enhancing data security and accessibility
- Improving cost-efficiency in data storage
- Ensuring scalability as the volume of carbon credit data grows

3.3. Theoretical contribution

The ICS framework contributes to the theoretical understanding of blockchain data management in several ways:

- It provides a systematic approach to evaluating and deciding between on-chain and off-chain storage methods, filling a gap in current literature.
- It introduces a novel concept of adaptive storage solutions in the context of carbon credit data, potentially applicable to other domains.
- The framework offers a model for balancing security, cost, and performance in blockchain-based systems, which can inform future research and development in this area.

This framework forms the foundation for our subsequent performance evaluations, where we will assess the effectiveness of ICS across multiple dimensions, providing valuable insights for practitioners in the field of carbon credit management and blockchain data storage (see Table 1).

4. Performance evaluations along multiple dimensions

In response to the growing need for efficient and secure carbon data management solutions, our research initiative embarked on the development of prototype systems designed to meticulously evaluate the spectrum of methodologies available for carbon data capture and storage. The core objective of these prototypes is to dissect and understand the nuances that differentiate on-chain implementations, characterized by their integration within the blockchain infrastructure, from off-chain implementations, which operate outside of the blockchain. This distinction is crucial because it directly influences the system's overall performance, its security capabilities, and the economic viability of its operation.

The development of these prototype systems was influenced by a multifaceted strategy aimed at establishing a comprehensive testing environment to monitor, assess, and contrast the operational performance of each system. Key performance indicators such as storage time, retrieval time, delay time, and latency act as measurable benchmarks for evaluating the effectiveness and responsiveness of each prototype. Security assessment, with a specific emphasis on data integrity, scrutinizes the ability of each system to protect against unauthorized access and ensure the credibility and precision of the stored carbon data. Lastly, cost assessment examines the operational and transactional costs associated with each approach, offering insights into the financial considerations of implementing on-chain versus off-chain solutions in practical settings.

Through the use of these prototype systems, our study aims to shed light on the strengths and weaknesses of both on-chain and off-chain approaches to capturing and storing carbon data. The results are intended to offer valuable perspectives on how to best balance performance, security, and cost effectiveness, with the goal of improving the development of more efficient carbon management tactics. This effort not only enriches our comprehension of technology but also promotes the overall objective of promoting sustainable practices in carbon data handling.

4.1. Parameter selection and analysis

To thoroughly evaluate the various methodologies utilized in carbon data capture and storage, our research project has resulted in the creation of two prototype systems. These systems were carefully developed with the primary goals of testing and comparison in mind, specifically to highlight the subtle differences in performance metrics, security measures, and overall cost effectiveness which set on-chain implementations apart from off-chain ones. Through a systematic approach, our aim is to analyze the inherent benefits and possible limitations associated with each approach, thus offering a comprehensive overview of their practicality in real-world settings. This initiative not only aims to enhance the existing knowledge base but also to steer future implementations towards more efficient, secure, and financially sustainable carbon data management solutions.

4.2. Performance evaluation metrics

To comprehensively assess the efficiency and responsiveness of our proposed carbon data storage and retrieval system, we employed the following performance metrics, evaluated using real-world carbon credit data:

1. **Storage Time:** Measured in milliseconds (ms), this metric quantifies the average time required to write a specified amount of carbon credit data to the system. Lower storage times, as observed with our data set of real carbon credit entries, indicate faster data ingestion, which is crucial for managing large volumes of data in practical applications. Factors that influence storage time include data size, the chosen storage technology, and available network bandwidth.
2. **Retrieve Time:** Measured in milliseconds (ms), this metric evaluates the average time it takes to access and retrieve a specific carbon credit data element upon request. It is crucial to differentiate between retrieving the entire dataset and accessing a specific data point, as observed with real carbon credit data. Lower retrieval times, as witnessed in our evaluation, signify faster data accessibility, improving system responsiveness to user queries for specific carbon credit information.
3. **Delay Time:** Measured in milliseconds (ms), this metric captures the total time elapsed from initiating a data request to receiving the retrieved data. An analysis of the real carbon credit data revealed several components contributing to delays, including network latency, server processing time for data retrieval and transformation, and potential queueing delays within the system. Optimizing delay time requires addressing each contributing factor to improve the overall system response in real-world scenarios.

Table 1
Research methodology stages.

Research methodology stages	Description
Objective Initialization	Define evaluation criteria (scalability, efficiency, economic viability), set clear objectives, and outline anticipated outcomes for analysis.
Prototype System Development	Build and implement prototype systems to test different blockchain storage methods, focusing on key performance indicators like operational performance, security, and cost.
Parameter Selection for Evaluation	Choose relevant parameters (storage time, retrieval time, delay time, latency) based on real-world significance and impact on system performance.
Performance Evaluation	Measure and analyze performance metrics (storage time, retrieval time, delay time, latency) using real-world carbon credit data sets.
Security Evaluation	Compare intrinsic security of on-chain storage with variable security of off-chain systems.
Cost Evaluation	Analyze operational and transaction costs to determine the overall cost effectiveness of different systems.
Comparative Analysis	Evaluate trade-offs and identify the optimal system based on a comprehensive analysis considering performance, security, and cost factors.
Conclusion and Recommendations	Summarize key findings, offer strategic recommendations based on empirical evidence, and highlight comparative advantages of on-chain and off-chain systems in carbon credit management.

4. **Latency:** Measured in milliseconds (ms), this metric signifies the observed delay in specific system operations, such as writing or reading carbon credit data. An analysis of real data showed that lower latency translates to faster response times for individual operations, enhancing real-time interaction and user experience when working with carbon credit information.

Trade-offs and Considerations: As observed with real carbon credit data, optimizing one metric may be beneficial (e.g., minimizing storage time), but it is crucial to consider the impact on other aspects, such as potentially increased latency or slower retrieval times. Therefore, striking a balance between these metrics is essential to achieve optimal overall system performance in practical applications that involve real-world carbon credit data.

Limitations: This performance evaluation is limited to the specific dataset and test environment used. Further evaluation with diverse carbon credit datasets and broader testing scenarios is recommended for a more comprehensive understanding of the system's performance under various real-world conditions.

4.2.1. Security evaluation

The evaluation of security measures in data storage, especially in the context of blockchain technology, involves a critical analysis between on-chain and off-chain storage solutions. This section delves into the inherent security features of on-chain data storage provided by blockchain technology and contrasts these with the variable security levels achievable with off-chain storage solutions. Our discussion is particularly relevant to applications in environmental sustainability, such as the storage and management of carbon credits on the Ethereum blockchain.

- On-Chain Data Storage Security

On-chain data storage is distinguished by several intrinsic security features that are foundational to blockchain technology:

Immutability: The blockchain's immutable ledger ensures that once data, such as carbon credit transactions, are committed, they cannot be altered or erased. This permanence is safeguarded by cryptographic hash functions linking blocks in a tamper-evident chain [23].

Decentralization: Unlike traditional centralized databases, blockchain operates on a distributed ledger technology paradigm. This decentralization eliminates single point of failure and distributes data across a network, significantly mitigating the risks associated with central data repositories.[24]

Cryptography: Blockchain leverages sophisticated cryptographic algorithms to secure data transactions. These algorithms ensure the confidentiality, integrity, and authentication of data transactions, making unauthorized access and alterations computationally infeasible.[23]

Consensus Mechanisms: The integrity of the data on the blockchain is maintained through consensus algorithms such as Proof of Work or Proof of Stake. These mechanisms require network-wide agreement for transaction validation, effectively preventing fraudulent activities and unauthorized data manipulation [25]

Transparency and Trust: Blockchain's transparent nature offers an additional layer of security. The ability for any network participant to verify transactions in real time fosters an environment of trust and deters malicious activities.

- Off-Chain Data Storage Security

The security of off-chain storage solutions is more variable and dependent on the specific technologies and protocols implemented. Factors influencing the security level of off-chain storage include:

Centralization versus Decentralization: Off-chain solutions may opt for centralized or decentralized storage models. While decentralized models can offer similar security benefits to blockchain by distributing data, centralized models are potentially more susceptible to targeted cyber attacks and data breaches.

Access Control: The effectiveness of off-chain storage security is heavily dependent on robust access control systems. Implementations that employ advanced authentication methods, including multifactor authentication and role-based access controls, provide superior security.

Encryption Standards: Off-chain storage security is also dependent on the use of strong encryption protocols for data at rest and in transit. The choice of encryption algorithms and key management practices is paramount in protecting against unauthorized data access.

Backup and Disaster Recovery: The resilience of off-chain storage is significantly enhanced by comprehensive backup and disaster recovery strategies. These measures ensure data durability and availability, even in the event of system failures or security breaches.

Regulatory Compliance: Adherence to regulatory standards and industry best practices is a critical component of off-chain storage security. Compliance with frameworks such as ISO/IEC 27001, GDPR, and others ensures that off-chain solutions meet established security and privacy benchmarks.

In conclusion, the security evaluation of on-chain versus off-chain data storage methodologies reveals distinct advantages and considerations. On-chain storage, with its inherent security properties stemming from blockchain technology, provides a robust framework for secure and immutable data storage. Off-chain solutions, while more varied in their security capabilities, offer flexibility and efficiency advantages when implemented with stringent security measures. This analysis underscores the importance of selecting appropriate data storage solutions based on the specific security, performance, and operational requirements of applications such as carbon credit management on the Ethereum blockchain.

4.2.2. Cost evaluation

Cost analysis encompasses:

- **Operational Costs:** Operational costs encompass the broad spectrum of expenses incurred to ensure the smooth functioning of data storage systems. These costs can vary significantly between on-chain and off-chain systems due to the intrinsic differences in their architectures and technologies.
  On-chain systems inherently require higher operational costs due to the nature of blockchain technology. Every transaction, including data storage and retrieval, requires computational power for execution and validation across the network. This decentralized verification process, while ensuring security and immutability, demands significant energy and resources, translating into higher operational costs.
  Off-chain systems, such as SQLite databases used in conjunction with blockchain technologies for hybrid solutions, present a more cost-effective approach to data management. Operational costs are lower, as these systems rely on traditional centralized servers for data storage and retrieval, which are typically more efficient in handling large volumes of data without the need for the extensive computational work required by blockchain networks.
- **Transaction Costs:** Research on the impact of transaction costs in on-chain systems, particularly on the Ethereum network, has revealed a complex interplay between these costs and economic activity [26]. This relationship is further explored in the context of operational transactions and smart contracts, with a focus on computational costs and the design of smart contract transactions in supply chains [27]. To address the issue of rising transaction fees, a middleware platform has been proposed to support flexible and secure transaction executions, including off-chain states and batching of user requests [28]. Finally, the optimization of smart contract execution costs on the blockchain network has been discussed, with a focus on reducing gas costs through the structure and content of smart contracts [29]. These studies collectively highlight the need for cost-effective solutions in on-chain systems, particularly in the context of blockchain technology.
  Off-chain systems generally do not incur these blockchain-specific transaction costs for data storage and retrieval operations. However, integrating these systems with the blockchain, for instance, to ensure data integrity or for specific transaction validations, might introduce minimal transaction costs associated with the blockchain interactions [30]. However, these costs are significantly lower than those of fully on-chain solutions.

In the context of blockchain technologies, particularly Ethereum, costs are inherently associated with the computational effort required to execute transactions, which includes data storage and retrieval operations [27]. These costs are influenced by several factors, including network congestion, data size, and the computational complexity of the smart contracts involved.

For carbon credits, a system that ensures both the integrity and accessibility of data while minimizing operational costs is crucial. On-chain data storage involves directly storing data on the blockchain. This method offers unparalleled security and data immutability but comes at higher costs due to the computational resources required for validation and consensus mechanisms.

Conversely, off-chain data storage mechanisms, such as utilizing databases like SQLite in conjunction with Ethereum, offer a more cost-effective solution [31]. While this approach can significantly reduce transaction costs by offloading data storage to more conventional databases, it necessitates additional mechanisms to ensure data integrity and security, bridging the gap between blockchain security and off-chain flexibility.

5. System design and implementation

5.1. Data selection process

The dataset utilized in this study was sourced from an open-access file, provided by the Berkeley Carbon Trading Project, available online at the University of California, Berkeley's website [32]. This comprehensive compilation of data is relevant to carbon credit projects. The selection criterion was meticulously aligned with the research objective, focusing on extracting pertinent data that offer insightful perspectives on carbon credit projects.

5.2. Structure overview

The selected dataset is structured into two primary nodes: *Project Block* and *Credit Nodes*, each encompassing critical data points, outlined as follows:

- **Project Block**: This is the central entity within the dataset, encapsulating essential information about carbon credit projects and associated credit details. It includes the following attributes:
  Project Node:
  - *Project ID*: A unique identifier assigned to each project.
  - *Project Name*: The official title of the project.
  - *Voluntary Registry*: The registry where the project is listed.
  - *ARB Project*: Denotes compliance with the Air Resources Board regulations.
  - *Voluntary Status*: The project's status in the voluntary carbon market.
  - *Scope*: The emission coverage scope of the project.
  - *Type*: The nature of the carbon offset project, e.g., reforestation, renewable energy.
  - *Methodology/Protocol*: The adopted methodology or protocol for carbon reduction or removal.
  - *Region*: The geographical region of the project.
  - *Country*: The country where the project is located.
  - *State*: The state or province of the project's location.
  - *Project Site Location*: The specific location of the project.
  - *Project Developer*: The organization or individual developing the project.
  - *First Year of Project*: The initial year when the project commenced.
  - *Issuance Date*: The date when the project was initiated and credits were issued.
  - *Verification Status*: The current verification status of the project.
  - *SDGs Met*: Sustainable Development Goals addressed by the project.
  - *Volume*: The quantity of carbon offset represented by each credit.
  - *Total Credits Issued*: The total number of credits issued by the project.
  - *Total Buffer Pool Deposits*: Credits deposited in the buffer pool for risk mitigation.
  - *Reversals Covered by Buffer Pool*: The number of reversals (loss of carbon sequestered) covered by the buffer pool.
- **Credit Nodes**: Nested within the Project Block, these nodes provide detailed information about each carbon credit.

- *Credit ID*: A unique identifier for each carbon credit.
- *Issuance Date*: The specific date of issuance for each credit.
- *Retirement Status*: Indicates whether the credit is active or retired.
- *Total Credits Retired*: The total number of credits that have been retired.
- *Total Credits Remaining*: The number of credits yet to be retired.
- *Reversals Not Covered by Buffer*: The number of reversals not covered by the buffer pool.

This dataset provides a multifaceted view of carbon credit projects, enabling a thorough analysis of various aspects such as compliance, methodology, regional distribution, and the impact on Sustainable Development Goals. The detailed structure of the dataset facilitates a comprehensive understanding of the dynamics involved in the issuance and management of carbon credits.

This section describes the design and implementation of the systems used to evaluate on-chain and off-chain data storage methods. The design focuses on creating robust, scalable, and secure systems for storing and retrieving blockchain data, accommodating various performance metrics and ensuring data integrity.

5.3. On-chain development techniques

5.3.1. Installation of solidity compiler (solc)

Solidity, the predominant language for Ethereum smart contracts, requires the installation of its compiler, solc. Multiple installation methods are available, including npm, binaries, or Docker. The compiler's source is hosted on its GitHub repository[1].

5.3.2. Development frameworks

For efficient development, frameworks such as Truffle and Hardhat are recommended:

- **Truffle**: A well-known framework, available on GitHub[2].
- **Hardhat**: A robust development environment, whose source code is on GitHub[3].

5.3.3. Deployment and interaction with smart contracts

These frameworks facilitate the compilation, deployment, and testing of smart contracts on either local test networks or public testnets.

5.4. Off-chain ethereum setup

Utilizing SQLite for Data Management SQLite, a lightweight and efficient database management system, is integrated to facilitate data storage and retrieval operations. This self-contained, serverless database engine provides robustness and simplicity, making it an ideal choice for managing transactional data within the project. SQLite's architecture ensures minimal setup, making it well-suited for environments with limited resources, while still offering the full capabilities of a SQL database management system.

SQLite's source code and installation instructions are accessible via its official website[4]. Its integration simplifies data handling, allowing for the efficient querying, updating, and management of blockchain transaction records or any other relevant data, thereby enhancing the overall efficacy and responsiveness of the application.

5.4.1. Selection of an ethereum client

The first step in the implementation process involves selecting an appropriate Ethereum client, considering factors such as programming language support and installation methods. Two primary options are:

- **Geth**: The Go implementation of the Ethereum protocol. It can be installed via package managers or pre-compiled binaries depending on the operating system. The source code is accessible via its GitHub repository[5].
- **OpenEthereum (formerly Parity)**: The Rust implementation of the Ethereum protocol, available in its GitHub repository[6].

5.4.2. Synchronization with the blockchain

After client installation, synchronization with the blockchain is essential. Options include a full sync, which is comprehensive but time-consuming, or a fast or light sync, offering quicker setup at the expense of not acquiring the entire blockchain history.

5.4.3. Interaction with the ethereum network

Post-synchronization interactions such as transactions, contract deployments, and potentially mining (subject to Ethereum's upcoming transition to Proof-of-Stake) are possible.

6. Algorithm for performance evaluation of on-chain and off-chain data storage methods in carbon credit systems

The algorithm incorporates performance metrics and cost analyses for a comprehensive range of carbon credit set sizes, spanning from 10 to 3000. This range is carefully selected to cover the entire spectrum of typical carbon credit transactions in the market. The smallest set size of 10 represents individual or small business transactions, often used for personal carbon offsetting or micro-business sustainability efforts. The mid-range sets (100, 300, 500) correspond to small to medium-sized business transactions or recurring purchases by larger entities. The larger sets (1000, 1500, 2000, 2500) are indicative of significant corporate offsetting efforts or bulk purchases by major institutions. Finally, the largest set size of 3000 represents major corporate or governmental transactions, often part of large-scale sustainability initiatives or compliance with stringent environmental regulations.

This algorithm takes into account both on-chain and off-chain architectures to provide a holistic view of system performance and cost-effectiveness across these varied transaction sizes. The on-chain component evaluates the efficiency and cost of storing carbon credit data directly on the blockchain, considering factors such as transaction speed and gas fees for each set size. The off-chain architecture assessment examines alternative data storage methods, analyzing their speed, cost, and scalability as the number of carbon credits increases from 10 to 3000. By comparing these two approaches across the entire range of set sizes, the algorithm aims to identify optimal storage strategies for different scales of operation, enabling a more nuanced and efficient approach to carbon credit management in blockchain-based systems.

Input:

- `sets`: A list of sets with sizes 10, 100, 300,500, 1500, 2000, 2500, 3000.
- `performance_metrics`: A dictionary with performance metrics for each set size, considering both on-chain and off-chain transactions
- `cost`: A dictionary with cost information (gas costs) for each set size, considering both on-chain and off-chain transactions
- `blockchain_config`: A dictionary with Ethereum blockchain configuration parameters
- `offchain_config`: A dictionary with off-chain storage configuration parameters.

Steps involved in performance evaluation:

1. Initialize an empty list `results` to store the results.

[1] https://github.com/ethereum/solidity
[2] https://github.com/trufflesuite/truffle
[3] https://github.com/nomiclabs/hardhat
[4] https://www.sqlite.org/index.html
[5] https://github.com/ethereum/go-ethereum
[6] https://github.com/openethereum/openethereum

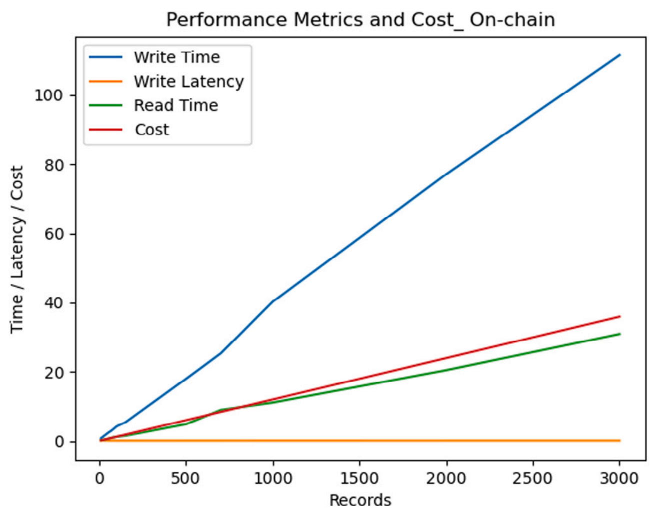

Fig. 2. Performance metrics and cost on-chain.

2. For each set size `n` in `sets`:
   a. Calculate the performance metrics for the set of size `n` using the `performance_metrics` dictionary, considering both on-chain and off-chain transactions.
   b. Calculate the cost for the set of size `n` using the `cost` dictionary, considering both on-chain and off-chain transactions.
   c. Calculate a weighted `score` based on performance metrics and cost, taking into account the on-chain and off-chain architecture.

7. Comparative analysis

7.1. On-chain results

The graph in Fig. 2 illustrates the increasing trend of write time and costs associated with the on-chain approach. The on-chain data handling approach demonstrates distinct behavior in relation to transaction times and costs.

The write time increases substantially with the number of records, suggesting a non-linear scalability that may present challenges in high-volume transaction environments. Similarly, write latency exhibits a slight but consistent increase, hinting at potential network strain as the transaction volume grows.

Read time also increases with larger datasets, reflecting the inherent complexities of data retrieval from a blockchain. Furthermore, the cost associated with on-chain transactions escalates significantly with the number of records, which could be prohibitive for operations at scale. These observations suggest that while the on-chain approach provides robust security and data integrity, it may be less suited for applications requiring high throughput or cost efficiency.

7.1.1. On-chain approach metric histograms

The histograms for the on-chain approach (Fig. 3) reveal the distribution of key performance metrics. The on-chain approach histograms provide insights into the distribution of metrics related to write time, write latency, read time, and cost:

Write Time Histogram: The on-chain write time histogram shows a skewed distribution, indicating that most of the write operations are relatively quick, with a tail extending towards longer times. This suggests that while the on-chain approach can handle transactions efficiently most of the time, there are instances where the write time can significantly increase, possibly due to the variable nature of block creation and transaction confirmation times.

Write Latency Histogram: The histogram for on-chain write latency is quite interesting, with a majority of operations occurring within a narrow time frame, but unlike write time, there is no long tail. This could indicate that once a transaction begins processing, the time it takes to be added to the blockchain is relatively consistent, but the start of that process can vary more.

Read Time Histogram: The read time histogram for on-chain is tightly clustered, reflecting consistent performance for reading data from the blockchain. The lack of a long tail suggests that read times are not as affected by the number of transactions as write times.

Cost Histogram: For on-chain costs, the histogram demonstrates that costs are mostly concentrated in the lower range, but with some occurrences of higher costs. This distribution likely reflects the complexity of transactions and their associated fees, which can vary based on the computational effort required to process and validate transactions on the blockchain.

Overall, the histograms for the on-chain approach suggest that while there is a base level of performance and cost that is generally consistent, there are outliers that can drive up the time and cost significantly. These outliers could be a crucial factor for entities considering blockchain for their operations, as they may affect the predictability and reliability of transaction processing on the chain.

7.1.2. On-chain approach metric trends

The trends in on-chain metrics (Fig. 4) show how the performance scales with an increasing number of records.

Write Time Trend: The on-chain write time trend shows a clear linear increase in the time required to write records as the number of records grows. This suggests that as transactions accumulate, the process of recording them on-chain becomes progressively slower, which could be due to the increased computational work needed to validate and add each transaction to the blockchain.

Write Latency Trend: The on-chain write latency trend presents an initial spike which quickly stabilizes as the number of records increases. The initial latency could be related to the start-up overhead or initial block formation that stabilizes with subsequent transactions.

Read Time Trend: The on-chain read time trend exhibits a steady increase, although it is less steep than the write time. This trend indicates that retrieving data from the blockchain becomes gradually slower with more records, likely due to the immutable and sequential nature of blockchain data retrieval.

Cost Trend: The on-chain cost trend shows a direct and strong positive correlation with the number of records. This is expected since more transactions mean more resource use on the blockchain, which increases the cost linearly.

7.2. Off-chain results

The graph in Fig. 5 demonstrates the relatively stable write time and lower costs for the off-chain method. In contrast, the off-chain approach yields different performance characteristics. Write time remains relatively stable and grows marginally with the number of records, indicating a more efficient handling of transaction writing which could be advantageous for scalability. Write latency, similar to write time, is comparatively low and does not show substantial growth with increased data volumes, implying an efficient transaction processing capability.

The trend with read time, while increasing, remains consistently lower than that of the on-chain approach across all record sizes. The cost analysis also indicates that the off-chain approach is more economical, with a less steep cost curve as the number of records grows. These results suggest that the off-chain approach, with its lower transaction times and costs, may be more suitable for applications that demand rapid processing and are sensitive to operational expenses.

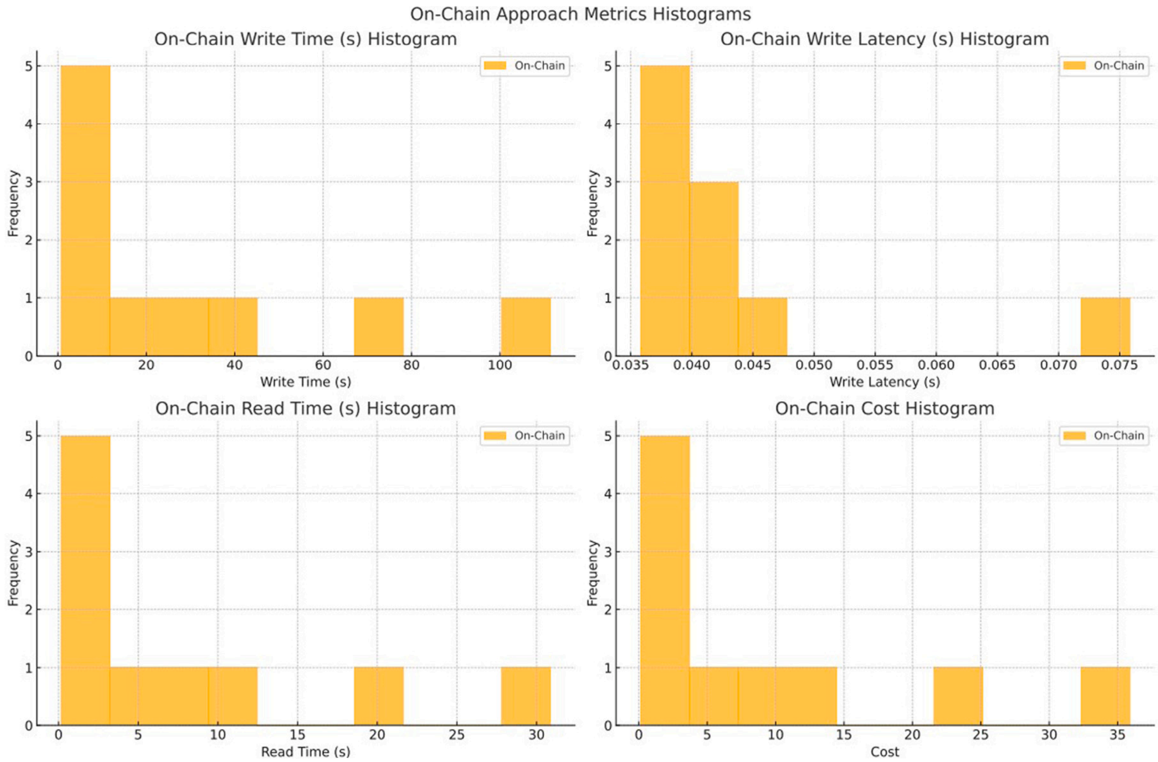

Fig. 3. On-chain approach metric histograms.

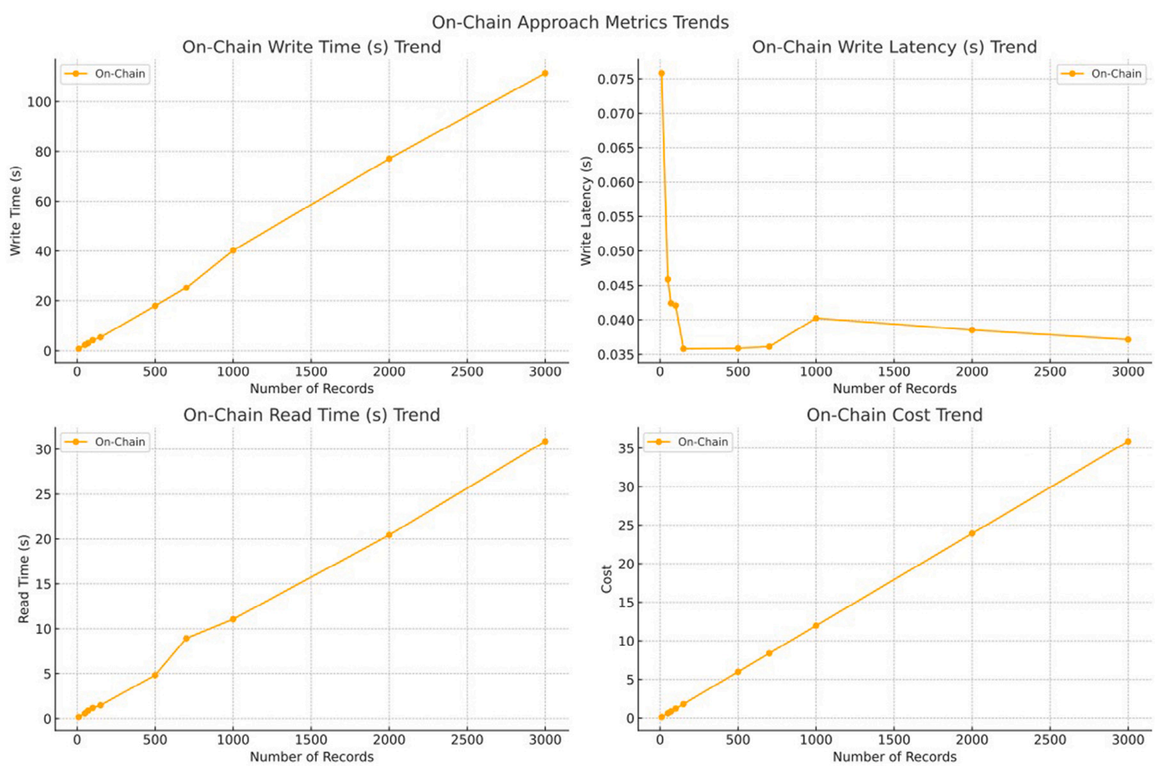

Fig. 4. On-chain approach metric trends.

7.2.1. *Off-chain approach metric histograms*

Fig. 6 illustrates the histograms for the off-chain approach:

Write Time Histogram: The off-chain write time histogram indicates that most of the write operations are clustered within a lower time bracket, with few occurrences of higher write times. This distribution shows that the off-chain approach is generally fast, with occasional outliers potentially due to network or operational anomalies.

Write Latency Histogram: The off-chain write latency histogram is similarly concentrated, with the majority of transactions occurring with low latency. The tight clustering implies consistent performance,

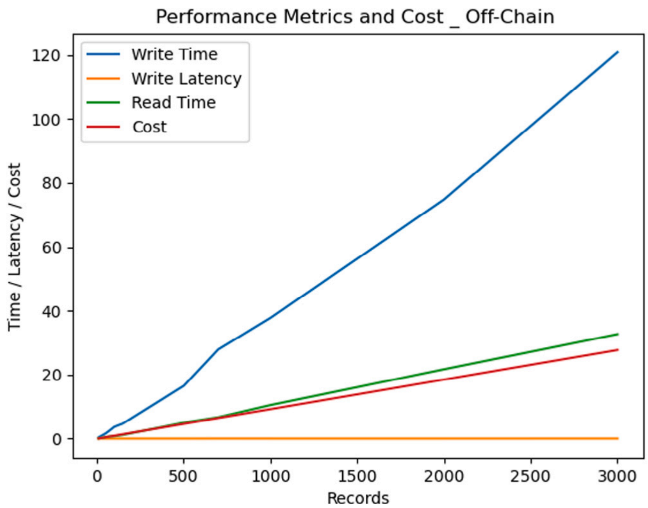

Fig. 5. Performance metrics and cost off-chain.

although a few transactions take significantly longer, as shown by the spread towards the higher end.

Read Time Histogram: The off-chain read time histogram also shows a concentration of read times at the lower end, suggesting that data retrieval is typically fast. However, the presence of bars at higher read times might indicate occasional delays, possibly due to data caching or synchronization issues.

Cost Histogram: For off-chain costs, the histogram is heavily skewed towards the lower end, highlighting the cost effectiveness of the off-chain approach. The lower frequency of higher costs underscores its economic advantage, especially when handling large datasets.

7.2.2. Off-chain approach metric trends

As depicted in Fig. 7, the trend lines for the off-chain metrics indicate the following:

Write Time Trend: The off-chain write time trend demonstrates a gradual increase, signifying that the time taken to write records scales moderately with the number of records. This trend suggests the good scalability of the off-chain system with respect to write operations.

Write Latency Trend: The trend for off-chain write latency is relatively flat with some variability. This suggests that the off-chain approach generally maintains consistent latency, barring a few exceptions where latency might spike.

Read Time Trend: The trend in off-chain read time exhibits a consistent and gradual upward trajectory, implying that while read operations slow down as the dataset size increases, the increase is proportionate and predictable.

Cost Trend: Lastly, the off-chain cost trend line climbs steadily, indicating that while costs increase with the number of records, the rise is more linear and controlled, which aligns with a predictable cost model that scales with usage.

In conclusion, the histograms and trends for both the on-chain and off-chain approaches reveal distinct characteristics in terms of performance and cost. The on-chain approach shows a linear increase in write time and costs, which could be problematic at scale, whereas the off-chain approach demonstrates overall more stable and lower write and read times, as well as costs, suggesting a more scalable solution for handling a larger number of transactions.

7.3. Comparison analysis

Fig. 8 presents a side-by-side comparison of the two approaches, highlighting the contrast between on-chain and off-chain.

Write Time Analysis: In comparing write time, we note that the on-chain approach generally takes a significantly longer time to write records as the number of records increases. This could be attributed to the nature of on-chain transactions which involve consensus mechanisms that are inherently slower. In contrast, the off-chain approach maintains a comparatively lower write time across all record counts, suggesting a more efficient process perhaps due to the absence of consensus-related delays. As data size grows, the time differential between on-chain and off-chain writes becomes more pronounced, indicating that off-chain solutions may offer better scalability with respect to write operations.

Write Latency Analysis: When examining write latency, it appears that the on-chain approach has a slight increase in latency with a larger number of records, which may be reflective of the increasing burden on the network as it processes more data. Off-chain solutions, however, show a relatively constant and lower latency, reinforcing the idea that off-chain systems can manage larger loads with minimal latency impact. This consistency is advantageous for applications that require predictable transaction processing times.

Read Time Analysis: The analysis of read time shows that on-chain approaches exhibit a growing read time with an increase in records, which can be a concern for applications requiring quick data retrieval. Off-chain approaches also display an increase, yet the read time remains below that of on-chain across the board. The gap in read times highlights the potential limitations of on-chain systems in terms of data retrieval speed, particularly at higher volumes of data.

Cost Analysis: a cost comparison reveals a direct correlation between the number of records and the cost for both approaches. On-chain, while secure and decentralized, shows a steeper increase in cost with the growth of records, which can be prohibitive for operations at scale. Off-chain maintains a lower cost trajectory, suggesting it is a more cost-effective solution, especially at scale. The disparity in costs could be decisive for organizations when considering the total cost of ownership over time.

In conclusion, while on-chain provides a decentralized and secure environment for transactions, it does so at the cost of increased write/read times and financial cost, particularly as scale increases. Off-chain solutions, by comparison, offer a more efficient and cost-effective approach for handling transactions, albeit potentially at the expense of the security and trustlessness that blockchain technology offers. The choice between on-chain and off-chain approaches should, therefore, be guided by the specific requirements and constraints of the application in question, including considerations of security, performance, and cost.

7.4. Framework for data management

The framework for data management depicted in Fig. 8 suggests a decision-making process that considers several key factors when choosing the appropriate data management approach. These factors include:

7.4.1. Data sensitivity and security requirements

The framework starts with an assessment of the data sensitivity and security requirements, which can be either moderate or high. This determines the overall direction of the decision-making process.

7.4.2. Scalability needs

Depending on the data sensitivity and security requirements, the framework considers the scalability needs, which can be either high or moderate.

7.4.3. Cost constraints

Another important factor is the cost constraints of blockchain, which can be either high or moderate.

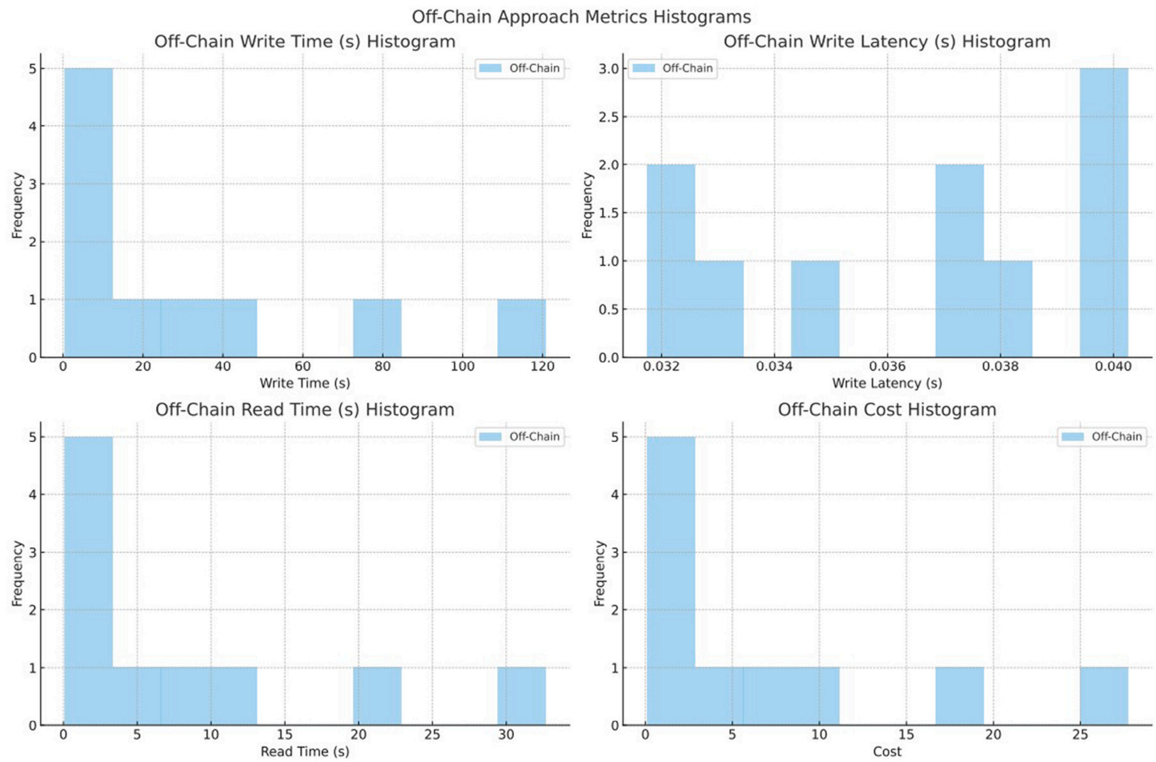

Fig. 6. Off-chain approach metric histograms.

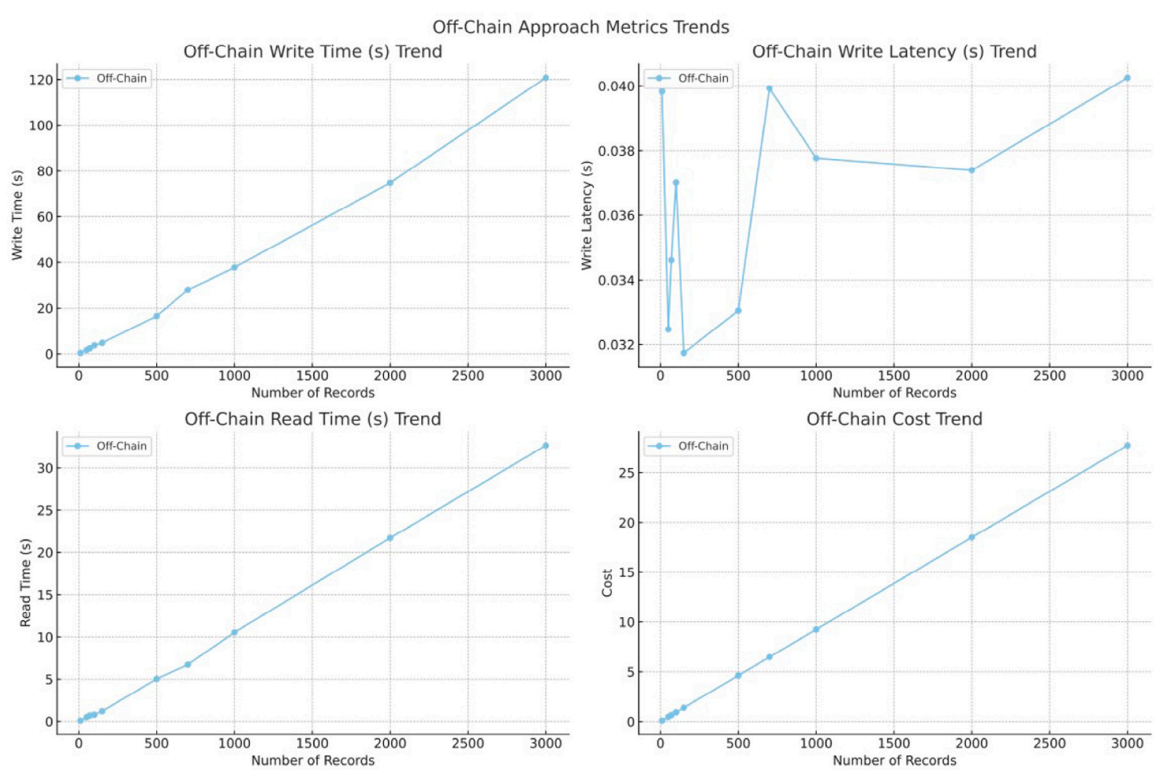

Fig. 7. Off-chain approach metric trends.

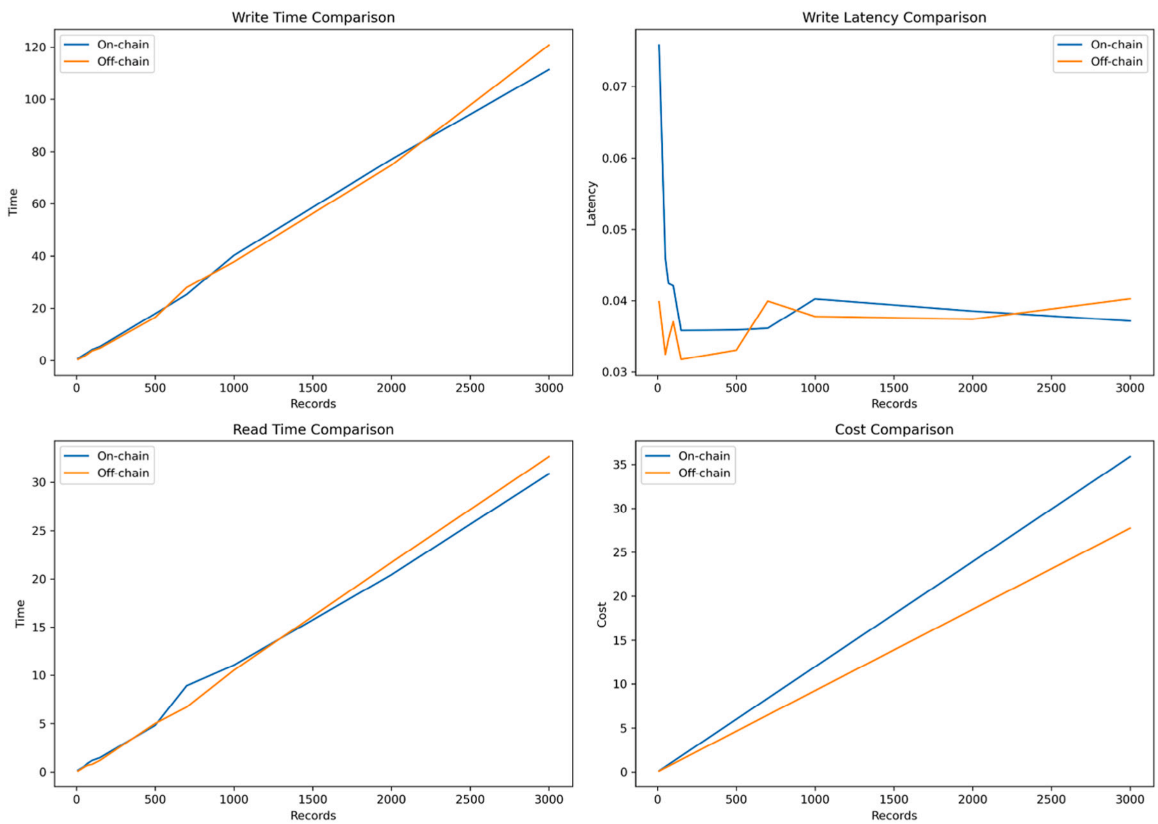

Fig. 8. Comparative analysis of on-chain and off-chain approaches.

7.4.4. Latency tolerance

The framework also takes into account the latency tolerance, which can be either high or low.

7.4.5. Regulatory compliance

Finally, the regulatory compliance requirements, which can be either low or moderate, are also factored into the decision-making process.

Based on the combination of these factors, the framework suggests several data management approaches, including on-chain and off-chain solutions. The specific approach selected will depend on the unique requirements and constraints of the particular use case or application.

The decision tree in Fig. 8 can be outlined as follows:

1. Start
2. Is data sensitivity high?
   - No → Off-chain
   - Yes → Is data high scalability?
     - No → Off-chain
     - Yes → Is cost high concern?
       - No → Off-chain
       - Yes → Is data high latency tolerance?
         - No → Off-chain
         - Yes → Is regulation compliance moderate?
         - No → Off-chain
         - Yes → On-chain

By considering this comprehensive set of factors, organizations can make informed decisions about the most suitable data management approach to meet their specific needs and requirements.

To illustrate how the decision tree can be applied, here are some hypothetical scenarios:

Scenario 1: Low Sensitivity, Low Scalability

- **Data Sensitivity:** Low
- **Outcome:** Off-chain

Scenario 2: High Sensitivity, Low Scalability

- **Data Sensitivity:** High
- **Scalability:** Low
- **Outcome:** Off-chain

Scenario 3: High Sensitivity, High Scalability, Low Cost

- **Data Sensitivity:** High
- **Scalability:** High
- **Cost Concern:** Low
- **Outcome:** Off-chain

Scenario 4: High Sensitivity, High Scalability, High-Cost Concern, Low Latency Tolerance

- **Data Sensitivity:** High
- **Scalability:** High
- **Cost Concern:** High
- **Latency Tolerance:** Low

- **Outcome:** Off-chain

Scenario 5: High Sensitivity, High Scalability, High Cost, High Latency Tolerance, Moderate Compliance

- **Data Sensitivity:** High
- **Scalability:** High
- **Cost Concern:** High
- **Latency Tolerance:** High
- **Regulation Compliance:** Moderate
- **Outcome:** On-chain

By applying the decision tree to these scenarios, the framework effectively guides the selection of the most appropriate data management approach based on the specific characteristics and requirements of the data and the application.

8. Conclusion and recommendations

This section delves into the implications of the observed results from the on-chain and off-chain data handling approaches. The on-chain approach's rising costs and transaction times with increased data volumes raise questions about its practicality for large-scale operations, despite its robust security benefits. On the other hand, the off-chain approach's lower costs and improved transaction efficiency offer a compelling alternative for large-scale applications, albeit with potential trade-offs in security and data integrity.

The results are in line with the expectations that blockchain-based solutions on-chain would incur higher overheads due to the consensus mechanisms required for validating transactions. This is a well-known trade-off in the design of blockchain systems, where security, transparency, and decentralization often come at the cost of efficiency. Conversely, off-chain solutions tend to optimize for performance and cost, which makes them attractive for use cases where rapid processing is critical, and trust assumptions are different.

In comparing the two approaches, it is essential to consider the context of application requirements. For instances where immutable audit trails and security are paramount, the on-chain approach may be justified despite the higher costs. Conversely, in scenarios where performance and cost are critical, and the trust model permits, off-chain solutions could provide significant benefits.

The findings from this analysis contribute to the ongoing discourse on the scalability of blockchain technologies and the exploration of hybrid systems that leverage both on-chain and off-chain approaches to balance trade-offs. Such hybrid systems could potentially harness the security benefits of on-chain transactions while delegating high-volume operations to more efficient off-chain processes.

CRediT authorship contribution statement

Nada Atetallah Alghanmi: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Nouf Atiahallah Alghanmi:** Software, Methodology. **Samaher Atiatallah Alghanmi:** Software, Methodology. **Ming Zhao:** Validation. **Farookh Khadeer Hussain:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, 2008, http://dx.doi.org/10.2139/ssrn.3440802.
[2] Andreas M. Antonopoulos, Mastering bitcoin: Unlocking digital crypto-currencies, 2014, https://api.semanticscholar.org/CorpusID:115315482.
[3] Jacob Eberhardt, Stefan Tai, On or off the blockchain? Insights on off-chaining computation and data, in: European Conference on Service-Oriented and Cloud Computing, 2017, https://api.semanticscholar.org/CorpusID:32306156.
[4] Hirotsugu Seike, Transaction queueing delay to ensure data integrity by blockchain, 2023, http://dx.doi.org/10.1109/ICCCBDA56900.2023.10154727.
[5] CryptoMaze: Privacy-preserving splitting of off-chain payments, IEEE Trans. Dependable Secure Comput. (2023) http://dx.doi.org/10.1109/tdsc.2022.3148476.
[6] Yuanhang Wu, Yanli Ren, Balanced off-chain payment channel network routing strategy based on weight calculation, Comput. J. (2023) http://dx.doi.org/10.1093/comjnl/bxad029.
[7] Faisal Nawab, Zk-oracle: Trusted off-chain compute and storage for decentralized applications, 2023, http://dx.doi.org/10.1109/ICBC56567.2023.10174864.
[8] Application of blockchain in the world of finance and business, 2023, http://dx.doi.org/10.31235/osf.io/jnzke.
[9] Applications and implementations of blockchain technology across the various sectors, in: Advances in Business Information Systems and Analytics Book Series, 2023, http://dx.doi.org/10.4018/978-1-6684-7808-0.ch001.
[10] Enhancing security feature in financial transactions using multichain based blockchain technology, 2023, http://dx.doi.org/10.1109/ICIEM59379.2023.10166589.
[11] Role of blockchain technology in the financial market, 2023, http://dx.doi.org/10.1108/978-1-80455-562-020231007.
[12] LEAF: Let's efficiently make adaptive forwarding in payment channel networks, IEEE Access (2023) http://dx.doi.org/10.1109/access.2023.3235828.
[13] Wenxuan Qiao, Zhengyi Li, Pengju Yang, Huanhao Li, A blockchain-based carbon transaction model for power industry, 2023, http://dx.doi.org/10.1117/12.2680811.
[14] Vaani Rawat, Nitin Dahiya, Siddhant Rai, Anshul Arora, A blockchain-based decentralized framework for carbon accounting, Trad. Govern. (2022) http://dx.doi.org/10.1145/3543712.3543755.
[15] Adam Sipthorpe, Sabine Brink, Tyler Van Leeuwen, Iain Staffell, Blockchain solutions for carbon markets are nearing maturity, One Earth (2022) http://dx.doi.org/10.1016/j.oneear.2022.06.004.
[16] Application of blockchain and IoT in cap and trade of carbon emissions, 2023, http://dx.doi.org/10.1109/raeeucci57140.2023.10134302.
[17] S.N.A. Ahmad Termizi, Sharifah Rafidah Wan Alwi, Zainuddin Abdul Manan, Petar Sabev Varbanov, Potential application of blockchain technology in eco-industrial park development, Sustainability (2022) http://dx.doi.org/10.3390/su15010052.
[18] Jiwen Yang, Chaoran Zheng, Hu Liu, Digital transformation and rule of law based on peak CO2 emissions and carbon neutrality, Sustainability (2022) http://dx.doi.org/10.3390/su14127487.
[19] Di Wang, Daozhi Zhao, Fang Chen, Research on financing strategy of green energy-efficient supply chain based on blockchain technology, Energies (2023) http://dx.doi.org/10.3390/en16072985.
[20] Xuejia Sang, Xiao Peng Leng, Linfu Xue, Xiangjin Ran, Based on the time-spatial power-based cryptocurrency miner driving force model, establish a global co2 emission prediction framework after china bans cryptocurrency, Sustainability (2022) http://dx.doi.org/10.3390/su14095332.
[21] Arpit Karnatak, G. Suseela, Preserving data integrity in legal documents using blockchain and NFTs, Adv. Sci. Technol. 124 (2023) 628–634, http://dx.doi.org/10.4028/p-cy5v3f.
[22] Shoufeng Cao, Thomas Miller, Marcus Foth, Warwick Powell, Integrating on-chain and off-chain governance for data storage, 2021, arXiv: Cryptography and Security.
[23] Shravani Amar, Blockchain mutability: Challenges and proposed solutions, Ind. Scient. J. Res. Eng. Manag. 06 (10) (2022) http://dx.doi.org/10.55041/ijsrem16668.
[24] Madhuri B. Thorat, Marmik S. Patil, Decentralization using blockchain as a service, Int. J. Sci. Technol. Eng. (2022) http://dx.doi.org/10.22214/ijraset.2022.47681.
[25] Jie Xu, Cong Wang, Xiaohua Jia, A survey of blockchain consensus protocols, ACM Comput. Surv. (2023) http://dx.doi.org/10.1145/3579845.
[26] L. Ante, A. Saggu, Time-varying bidirectional causal relationships between transaction fees and economic activity of subsystems utilizing the Ethereum blockchain network, J. Risk Finan. Manag. 17 (1) (2024) 19.
[27] A. Jabbar, S. Dani, Investigating the link between transaction and computational costs in a blockchain environment, Int. J. Prod. Res. 58 (11) (2020) 3423–3436.
[28] Y. Wang, Y. Tang, Enabling cost-effective blockchain applications via workload-adaptive transaction execution, 2022, arXiv preprint arXiv:2210.04644.
[29] D. Čeke, S. Kunosić, N. Buzađija, Smart contract execution costs optimisation on blockchain network, in: 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology, MIPRO, IEEE, 2022, pp. 1442–1447.

[30] T. Hepp, M. Sharinghousen, P. Ehret, A. Schoenhals, B. Gipp, On-chain vs. off-chain storage for supply-and blockchain integration, it-Inf. Technol. 60 (5–6) (2018) 283–291.
[31] X. Liu, R. Chen, Y.W. Chen, S.M. Yuan, Off-chain data fetching architecture for ethereum smart contract, in: 2018 International Conference on Cloud Computing, Big Data and Blockchain, ICCBB, IEEE, 2018, pp. 1–4.
[32] University of California, Berkeley, Berkeley Carbon Trading Project: Offsets Database, Goldman School of Public Policy, 2023, Available: https://gspp.berkeley.edu/research-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/offsets-database. (Accessed 8 December 2023).