Blokzincir Tabanlı Etkili Bir Denetim İzi Çerçevesi

Bilal S. T. Alagha

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Aralık 2024

A Blockchain-Based Efficient Audit Trail Framework

Bilal S. T. Alagha

MASTER OF SCIENCE THESIS

Department of Computer Engineering

December 2024

A Blockchain-Based Efficient Audit Trail Framework

A thesis submitted to the Eskisehir Osmangazi University
Graduate School of Natural and Applied Sciences
in partial fulfillment of the requirements for the degree of Master of Science
in Discipline of Information Security of the Department of Computer Engineering
by

Bilal S. T. Alagha

Supervisor: **Dr. Öğr. Üyesi İlker ÖZÇELİK**

December 2024

APPROVAL OF THE THESIS

The thesis titled "A Blockchain-Based Efficient Audit Trail Framework" and submitted by "Bilal S. T. Alagha" has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Department of Computer Engineering.

Supervisor : Dr. Öğr. Üyesi İlker ÖZÇELİK

Co-Supervisor : -

Examining Committee Members:

Member : Dr. Öğr. Üyesi İlker ÖZÇELİK

Member : Dr. Öğr. Üyesi Emrah ATILGAN

Member : Dr. Öğr. Üyesi Mehmet SEVRİ

Graduation of "Bilal S. T. Alagha" was approved by the Graduate School Board Decision on ..... with the decision number of .....

Prof.Dr. Fatma TÜMSEK
Director of the Institute

ETHICAL STATEMENT

I hereby declare that this thesis study titled "A Blockchain-Based Efficient Audit Trail Framework" has been prepared in accordance with the thesis writing rules of Eskişehir Osmangazi University Graduate School of Natural and Applied Sciences under academic consultancy of my supervisor Dr. Öğr. Üyesi İlker ÖZÇELİK . I hereby declare that the work presented in this thesis is original. I also declare that, I have respected scientific ethical principles and rules in all stages of my thesis study, all information and data presented in this thesis have been obtained within the scope of scientific and academic ethical principles and rules, all materials used in this thesis which are not original to this work have been fully cited and referenced, and all knowledge, documents and results have been presented in accordance with scientific ethical principles and rules.
27.12.2024

Bilal S. T. Alagha

ÖZET

Denetim izi, bir sistemdeki faaliyetleri kaydederek ve değişiklikleri izleyerek, kritik bir kontrol mekanizması olarak görev yapar. Muhasebe ve finans sistemlerinde yaygın olarak kullanılan denetim izleri, sahtekarlık önleme, uyumluluk ve veri yönetiminde önemli bir rol oynar. Ancak geleneksel denetim izi sistemleri genellikle verimsizdir, yüksek maliyetlidir ve güvenlik ihlalleri ile yetkisiz erişime karşı savunmasızdır. Bu kısıtlamalar, gerçek zamanlı doğrulama yeteneklerinin eksikliğiyle birleşerek, geleneksel denetim izlerinin sahtekarlığı önleme ve veri bütünlüğünü sağlama konusundaki etkinliğini sınırlamaktadır. Wirecard gibi büyük şirketlerin çöküşü ve Luckin Coffee'nin karşılaştığı zorluklar, daha güvenli, güvenilir ve verimli çözümlere olan acil ihtiyacı vurgulamaktadır.

Blokzincir teknolojisi, geleneksel denetim izlerine umut verici bir alternatif sunar. Merkeziyetsiz ve değiştirilemez defteri, şeffaf ve gerçek zamanlı işlem kayıtları sağlar. Blokzincir, işlem kayıtlarını otomatikleştirerek, manuel prosedürlerle ilişkili zamanı ve maliyeti azaltırken, denetim kalitesini önemli ölçüde artırır ve sahtekarlık risklerini düşürür. Bu avantajlara rağmen, blokzincir denetim sistemleri ölçeklenebilirlik ve arama konularında kritik zorluklarla karşı karşıyadır. İşlem hacmi arttıkça, belirli bilgileri almak ve doğrulamak giderek zaman alıcı hale gelmekte, blokzincir tabanlı denetim sistemlerinin gerçek dünya uygulamalarında pratikliğini sınırlamaktadır.

Bu tez, blokzincir içindeki belirli bilgilerin doğrulanması için yeni bir çözüm olarak Kriptografik Biriktirici'yi kullanarak bu zorlukları ele almaktadır. Kriptografik Biriktirici, işlem verilerini tek bir kompakt değerde toplayarak, tüm blokzincir'e erişmeden üyelik doğrulamasını etkin bir şekilde sağlar. Bu yaklaşım, blokzincir tabanlı denetim sistemlerinin verimliliğini ve ölçeklenebilirliğini önemli ölçüde artırarak, temel özellikleri olan değiştirilemezlik, güvenlik ve kullanılabilirliği korurken, büyük ölçekli uygulamalar için pratik olmalarını sağlar.

Bu yaklaşımı doğrulamak için araştırma, RSA tabanlı Kriptografik Biriktirici'nin performansını Sawtooth platformunda değerlendirerek, doğrusal işlem arama, veritabanı destekli işlem arama ve blokzincir platformuna özgü çözümler gibi mevcut yöntemlerle karşılaştırmıştır. Test sonuçları, RSA Kriptografik Biriktirici'nin karmaşıklığı ile işlem doğrulama süresini önemli ölçüde azalttığını göstermiştir. Arama ve ölçeklenebilirlik sınırlamalarını ele alarak, bu araştırma, güvenli ve verimli denetim çözümlerinin benimsenmesine katkıda bulunmaktadır.

Önerilen çerçeve, yalnızca doğrulama sürecini geliştirmekle kalmaz, aynı zamanda mevcut blokzincir tabanlı denetim sistemlerine sorunsuz bir şekilde entegre olur. İşlem verilerinin hızlı ve güvenilir bir şekilde doğrulanmasını sağlayarak, blokzincir denetim sistemlerinin kritik bir yönünü ele alır. Arama yerine doğrulamaya odaklanarak, bu araştırma, gerçek dünya uygulamaları için pratik bir çözüm geliştirmek üzere kapsamını daraltır.

Bu araştırmanın önemi, denetim uygulamalarında devrim yaratma potansiyelinde yatmaktadır. Kriptografik Biriktirici'yi kullanarak kuruluşlar, sağlam, değiştirilemez ve ölçeklenebilir denetimler elde edebilir, şeffaflığı ve uyumluluğu sağlayabilir ve finansal sahtekarlık risklerini azaltabilir. Bu yaklaşım, geleneksel ve blokzincir tabanlı denetim sistemlerinin karşılaştığı zorluklar için ölçeklenebilir ve verimli bir çözüm sunarak, blokzincir teknolojisinin denetim, sahtekarlık önleme ve uyumluluk alanlarında daha geniş çapta benimsenmesini sağlar. Bu çalışmanın bulguları, yalnızca denetim izlerinin güvenilirliğini ve verimliliğini artırmakla kalmaz, aynı zamanda güvenli veri yönetimi ve işlem doğrulamada kriptografik tekniklerin daha geniş anlamda anlaşılmasına da katkıda bulunur.

Sonuç olarak, bu tez, arama ve ölçeklenebilirlik gibi kritik zorlukları ele alarak, blokzincir denetim alanında mevcut durumu ileriye taşımaktadır. RSA tabanlı Kriptografik Biriktirici, etkin ve güvenilir işlem doğrulama için güçlü bir araç olarak ortaya çıkmakta, blokzincir tabanlı denetim sistemlerini büyük ölçekli uygulamalar için daha pratik ve ölçeklenebilir hale getirmektedir. Teorik gelişmeleri gerçek dünya gereksinimleriyle köprüleyerek, bu araştırma, geniş bir uygulama yelpazesinde güvenli, verimli ve ölçeklenebilir çözümler sunan yeni nesil denetim izi sistemleri için temel oluşturmaktadır.

Anahtar Kelimeler: Denetim İzi, Blokzincir, Verimlilik, Kriptografik Biriktirici, Doğrulama.

SUMMARY

An audit trail records activities and tracks changes within a system, serving as a crucial tool for fraud prevention and data management. However, traditional audit trail systems often suffer from inefficiencies, high costs, and vulnerability to security breaches and unauthorized access, limiting their effectiveness in real-time fraud detection. These limitations highlight the urgent need for more reliable, secure, and efficient solutions. Blockchain technology, with its decentralized and immutable ledger, provides a promising solution. By offering tamper-proof, transparent, and real-time transaction records, blockchain significantly enhances the reliability, efficiency, and security of audit trails, revolutionizing auditing practices and increasing stakeholder trust.

Despite these advantages, a significant challenge remains: as the volume of transactions increases, retrieving and verifying specific information becomes increasingly time-consuming and resource-intensive. This issue, often overlooked in research, limits the scalability and practicality of blockchain-based audit systems in real-world applications.

This thesis focuses on developing an efficient approach for verifying specific information within a blockchain, addressing the limitations of searchability and scalability in current audit trail systems. The proposed approach introduces the Cryptographic Accumulator as a novel solution to these challenges. By accumulating all transaction hashes into a compact data structure, the cryptographic accumulator enables auditors to verify transaction authenticity without exhaustive searches or supplementary systems.

Our study evaluated this approach on the Sawtooth platform, comparing its performance against linear transaction search in blockchain, database-supported transaction search in blockchain, and blockchain platform-specific solutions. Test results prove that RSA-based cryptographic accumulator significantly enhances audit trail information verification efficiency with $O(1)$ time and space complexity, making blockchain-based audit trail systems more practical and scalable for large-scale applications. This research contributes to the advancement of blockchain technology, paving the way for its adoption in domains requiring secure, efficient, and scalable audit solutions.

Keywords: Audit Trail, Blockchain, Efficiency, Cryptographic Accumulator, Verification.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisor, Dr. Öğr. Üyesi İlker Özçelik, for his invaluable guidance, continuous support, and encouragement throughout every stage of this thesis. His extensive knowledge and expertise have been instrumental in shaping the outcome of this work, and I am profoundly thankful for the opportunity to learn under his guidance.

I am immensely grateful to my beloved family, especially my father, Samir, and my mother, Mirvat, for their patience and unwavering support in every aspect of my life. Their encouragement and belief in me have been my greatest source of strength.

Finally, I would like to extend my heartfelt thanks to Türkiye and Eskişehir Osmangazi University for providing me with the resources and opportunities to pursue my studies. Their support has been vital in making this academic journey possible.

LIST OF CONTENTS

Page

ÖZET ..... **vi**

SUMMARY ..... **viii**

ACKNOWLEDGMENT ..... **ix**

LIST OF CONTENTS ..... **x**

LIST OF FIGURES ..... **xii**

LIST OF TABLES ..... **xiii**

LIST OF ABBREVIATIONS AND SYMBOLS ..... **xiv**

1. INTRODUCTION AND PURPOSE **1**

2. LITERATURE REVIEW **5**

2.1. Historical Context and Evolution of Audit Trails 5

2.1.1. Traditional Methods of Creating and Maintaining Audit Trails 6

2.1.2. Impact of Regulations and Technological Advancements 6

2.2. Blockchain in Audit Trails 7

2.2.1. Auditing and Remote Audits (Governance) 7

2.2.2. Financial Audits 8

2.2.3. Enterprise Applications 9

2.2.4. Data and Records 10

2.2.4.1. Data Management and Sharing 10

2.2.4.2. Record Keeping 11

2.3. Blockchain's Impact on Audit Trails 13

2.3.1. Comparative Analysis: Pre- and Post-Blockchain Integration 13

2.4. Research Gaps 14

3. BACKGROUND **15**

3.1. Blockchain Technology 15

3.1.1. The Advent of Blockchain Technology 16

3.1.2. Sawtooth & Its Architecture 17

3.1.2.1. Sawtooth Transaction Process 20

3.1.3. Searching in Blockchain 21

3.1.3.1. Searching Using Identifiers 21
3.1.3.2. Searching Based on Database Structures 26
3.1.3.3. Searching Based on Authenticated Index Structures 27
3.1.3.4. Searching Using Merkle Trees 29
3.1.3.5. Other Search Approaches 30
3.1.4. Transaction Search Approaches 32
3.1.4.1. Linear Search 33
3.1.4.2. Platform Specific Solutions 33
3.1.4.3. Searching Using Databases 34
3.1.4.4. Searching Using a Merkle Hash 35
3.2. InterPlanetary File System (IPFS) 35
3.3. Cryptographic Accumulators 40
3.3.1. RSA-Based Cryptographic Accumulator 43

4. METHODOLOGY 46
4.1. System Design 46
4.2. Framework Development 50
4.3. Evaluation Metrics 51

5. RESULTS AND DISCUSSION 52
5.1. Experiments (Existing & Proposed) 52
5.2. Results (Existing & Proposed) 53
5.3. Discussion 57
5.3.1. Comparison with Existing Systems 60
5.3.2. Implications for Engineering Applications 61

6. CONCLUSION AND RECOMMENDATIONS 63
6.1. Summary of Findings 65
6.2. Limitations 67
6.3. Future Research Directions 67

REFERENCES 69

LIST OF FIGURES

Figure		Page
3.1.	Blockchain construction	16
3.2.	Sawtooth network	18
3.3.	A Sawtooth node	19
3.4.	Validating transactions and updating the global state in Sawtooth.	21
3.5.	Linear search	33
3.6.	Database-based search	34
3.7.	B-tree structure	34
3.8.	Storing documents in IPFS	36
3.9.	Document retrieval and sharing in IPFS	37
3.10.	Key concept of cryptographic accumulators	40
4.1.	Overview of BEATS, illustrating the user-system interactions.	46
4.2.	Detailed workflow of BEATS	47
4.3.	Registering in the system	48
4.4.	Adding documents to accounts	49
4.5.	The audit process	50
5.1.	Average transaction download time required for linear search.	53
5.2.	Average CPU time required for linear transaction search.	54
5.3.	Average transaction retrieval time for platform-specific solutions.	54
5.4.	Average CPU time required for transaction retrieval in a database-supported blockchain system.	55
5.5.	Average CPU time required for retrieving the accumulator value from the blockchain and performing verification.	56
5.6.	Average CPU time required for verification.	56
5.7.	Average RAM consumption required for retrieving the accumulator value from the blockchain and performing verification.	57

LIST OF TABLES

Table		Page
2.1.	Comparing audit trails: before and after blockchain	14
3.1.	Key features of cryptographic accumulators	41
3.2.	Types of cryptographic accumulators based on set modification	42
3.3.	Types of cryptographic accumulators based on membership testing	42
3.4.	Applications of cryptographic accumulators	43

LIST OF ABBREVIATIONS AND SYMBOLS

Symbols	**Explanations**
O()	Big O
S	The set of elements to be accumulated
i	An integer index indicating the position of an element in the set S
j	An integer representing the total number of elements in the set S
n	RSA modulus
p	A large prime number
q	Another large prime number
$\mathbb{Z}_n^*$	The multiplicative group consisting of integers coprime with n
g	A generator chosen within $\mathbb{Z}_n^*$
$\in$	Indicates membership, meaning an element belongs to a set
A	The accumulator value representing the set S
x_i	An individual element of the set S
$\prod_{i=1}^{k}$	Product notation
mod	The modulo operation
w	The witness for an element x
x'	A new element being added to the set S
x'^{-1}	The modular inverse of x', used when removing an element from S
X	Placeholder
K	Thousand (commonly used to represent 1,000)
M	Million
s	Seconds
ms	Milliseconds
n	Number of Transactions
log(n)	Logarithm of n
c	Constant Coefficient
m	Number of Blocks in Blockchain
k	Average Number of Transactions per Block

Abbreviations	**Explanations**
ACC	Accumulator
AccVersion	Accumulator Version
ACFE	Association of Certified Fraud Examiners

ADS	Authenticated Data Structures
AI	Artificial Intelligence
AMD	Advanced Micro Devices, Inc.
API	Application Programming Interface
BEATS	Blockchain-Based Efficient Audit Trail System
BoBs	Blockchain-of-Blockchains
B-tree	Balanced/Bayer Tree
C	C Programming Language
C++	C++ Programming Language
CBDC	Central Bank Digital Currency
CCPA	California Consumer Privacy Act
CID	Content Identifier
CLI	Client
CompanyID	Company Identifier
CPA	Certified Public Accountant
CPU	Central Processing Unit
dApps	Decentralized Applications
DB	Database
DC	Data Consumer
DDoS	Distributed Denial-of-Service
DHT	Distributed Hash Table
DO	Data Owner
Doc	Document
DocID	Document Identifier
DocPID	Document Prime Representative Identifier
DocWit	Document Witness
DoS	Denial-of-Service
DRE	Direct Recording Electronic
eHealth	Electronic Health
EHRs	Electronic Health Records
EMR	Electronic Medical Record
EMRs	Electronic Medical Records
ESPAC	Enabling Security and Patient-centric Access Control
GAA-FQ	Granular Access Authorization supporting Flexible Queries
GB	GigaByte
GDPR	General Data Protection Regulation
GPUs	Graphics Processing Units
ID	Identifier

IIoT	Industrial Internet of Things
IoT	Internet of Things
IPFS	InterPlanetary File System
IT	Information Technology
I-Voting	Digital Voting (I-Voting)
K8s	Kubernetes
KBs	kilobytes
LMDB	Lightning Memory-Mapped Database
MerkleDag	Merkle Directed Acyclic Graph
ML	Machine Learning
NAK	National Archives of Korea
NIST	National Institute of Standards and Technology
No	Number
P2P	Peer-to-peer
pAuditChain	Privacy-preserving Audit Chain
PBFT	Practical Byzantine Fault Tolerance
PHR	Patient Health Records
PIPEDA	Personal Information Protection and Electronic Documents Act
PIR	Private Information Retrieval
PKI	Public Key Infrastructure
PoC	Proof of Concept
PoET	Proof of Elapsed Time
PoS	Proof of Stake
PoW	Proof of Work
R&D	Research and Development
RAM	Random Access Memory
RDBMS	Relational Database Management System
REDCap	Research Electronic Data Capture
REST	Representational State Transfer
RSA	Rivest-Shamir-Adleman
SDK	Software Development Kit
SEBDB	Semantics Empowered Blockchain Database
SGX	Software Guard Extensions
SPIs	Service Provider Interfaces
SPOF	Single Point of Failure
SQL	Structured Query Language
SSE	Searchable Symmetric Encryption
SSE	Secure Searchable Encryption

TCP	Transmission Control Protocol
TEE	Trusted Execution Environment
TRMS	Trustworthy and Reliable Multi-keyword Search
TX	Transaction
URL	Uniform Resource Locator
VQL	Verifiable Query Layer
VVPAT	Voter Verifiable Paper Audit Trail
Web3j	Web 3 Java
ZeroMQ	ZeroMessageQueue

1. INTRODUCTION AND PURPOSE

An audit trail is a record of all events and activities that occur within a system. It is widely used in accounting and financial systems to track data changes. However, recent audit failures, such as the collapse of Wirecard (Baranidharan et al., 2023; Aßländer and Burkatzki, 2023) and the issues faced by Luckin Coffee (Li, 2023), underscore the urgent need for more reliable audit trails. Traditional audit methods reveal significant shortcomings: they often rely on outdated systems vulnerable to security breaches, unauthorized access, and manual procedures prone to errors and delays. Furthermore, these systems are typically time-consuming and costly (Newray, 2023), with the absence of real-time verification allowing fraudulent activities to remain undetected.

The significance of a reliable audit trail has been highlighted by numerous instances of financial fraud and mismanagement. According to the Association of Certified Fraud Examiners, businesses lose approximately 5% of their revenue annually due to fraud, amounting to over $4.7 trillion globally (ACFE, 2022). To combat this, reliable audit trails play a crucial role in ensuring transparency, compliance, and fraud prevention. Both government agencies and organizations use audit trails to meet regulatory requirements, such as those mandated by the GDPR (GDPR, 2019) and PIPEDA (Edwards et al., 2004), which emphasize error correction, auditing, and internal controls (Ringelstein and Staab, 2009).

To address these challenges, blockchain technology offers a promising solution for maintaining reliable audit trails and enhancing audit practices. A survey by Deloitte (Deloitte, 2019) indicates that 53% of respondents believe blockchain improves audit quality and reduces fraud risks. Blockchain's decentralization, immutability, and transparency make it particularly suitable for creating reliable and tamper-proof audit trails (Schulman, 2019). Transactions recorded on the blockchain are immutable and visible to all participants, significantly reducing the risk of fraud and enhancing the overall efficiency of the audit process. Additionally, blockchain automates transaction recording, lowering the time and cost associated with traditional manual record-keeping.

Blockchain is widely used to securely store information, offering a reliable audit trail. However, a significant challenge arises as the volume of data increases over time. While blockchain inherently ensures the immutability of stored data, verifying the existence of a piece of information becomes increasingly time-consuming and resource-intensive due

to inefficient search methods. This emphasizes the need for improved solutions to streamline information retrieval in blockchain-based systems.

While researchers have made significant progress in leveraging blockchain's immutability to store and secure data, efficient data retrieval from blockchain has received far less attention. As blockchain continues to expand, the time and resources required for finding specific data increase substantially, making the process cumbersome in real-world applications where efficient and fast data verification is crucial. Addressing this gap is essential to ensure that blockchain-based systems remain scalable, efficient, and practical for long-term use.

In this thesis, our objective is to develop an efficient approach for finding verification information within a blockchain for audit trail purposes. This information is essential for verifying the integrity and authenticity of data stored in the blockchain. To address this challenge, the research focuses on introducing an optimized approach that not only ensures efficient verification but also integrates it into the design of a comprehensive audit trail framework. This framework will enhance the overall efficiency and scalability of blockchain-based audit systems, ensuring they remain practical for real-world applications while maintaining their core properties of immutability, security, and availability. By achieving this goal, this research will contribute to the advancement of blockchain technology and its adoption in domains that require reliable and efficient audit solutions.

This research focuses solely on verifying a specific piece of information within the blockchain, rather than locating or searching for it. The primary objective is to ensure that the information can be quickly and reliably verified as being part of the blockchain, addressing one of the critical aspects of blockchain audit systems. By narrowing our focus to verification rather than searchability, this work aims to develop an approach that is both efficient and practical for real-world applications.

Finding specific transactions is essential for auditing, fraud prevention, compliance, and financial crime investigations. Researchers have explored various transaction search approaches, such as using identifiers, relational databases and authenticated index structures (Tang et al., 2023). While these methods offer improvements over traditional linear transaction search, they fall short of achieving optimal performance, highlighting the need for further investigation and refinement.

To verify the presence of data on the blockchain, a typical approach involves searching for and locating the data. However, if the purpose of finding this data is verification, a more efficient solution can be employed using cryptographic accumulators.

These data structures are both time- and space-efficient, making them reliable and effective solution for this purpose. By implementing a cryptographic accumulator within a smart contract on the blockchain, the need for traditional transaction search methods is eliminated. Instead, we only use the verification information to verify the existence of data on the blockchain. This approach significantly reduces the system's complexity while ensuring secure, scalable and efficient data verification.

In this context, cryptographic accumulators, particularly the RSA-based cryptographic accumulator, emerge as a powerful solution for transaction verification. Unlike traditional methods, the RSA-based cryptographic accumulator allows transaction hashes to be securely accumulated into a compact data structure, enabling efficient membership verification without the need to process the entire blockchain. This not only enhances the scalability of blockchain systems but also ensures robust and tamper-resistant auditing.

This research conducts a comparative analysis of existing transaction search approaches and demonstrates the superiority of cryptographic accumulators in blockchain. Our study evaluates these apporaches, including the linear transaction search, database supported transaction search, and blockchain platform-specific solutions, while highlighting the performance advantages of the RSA-based cryptographic accumulator. The cryptographic accumulator achieves optimal efficiency and scalability, particularly in scenarios involving larger number of transactions or complex auditing requirements.

The remainder of this thesis is organized as follows: Chapter 2 explores the historical context and evolution of audit trails. It reviews existing blockchain applications in audit trails and examines the impact of blockchain on these systems by comparing traditional audit systems with blockchain-integrated systems. This chapter also identifies research gaps. Chapter 3 provides a comprehensive overview of blockchain technology, with a focus on Sawtooth and its architecture. It explores various transaction search methods within blockchain systems and introduces the Interplanetary File System (IPFS). The chapter then highlights the core component of this research: the Cryptographic Accumulator. Chapter 4 presents the methodology used in this research, detailing the system design and framework development. It explains the roles of the cryptographic accumulator, the on-and off-chain systems, and the IPFS within the proposed framework. Additionally, it outlines the evaluation metrics used to assess the system's performance. Chapter 5 presents the results of the study, encompassing both existing experiments and the proposed approach. It offers a detailed analysis and discussion of the test results, evaluating the system's performance. Additionally, the chapter compares the proposed system with traditional methods and explores the implications of the proposed approach for engineering

applications. Finally, Chapter 6 concludes the thesis by summarizing the findings, discussing the study's limitations, and identifying potential directions for future research.

2. LITERATURE REVIEW

2.1. Historical Context and Evolution of Audit Trails

Audit trail is a record of all the events and activities that occur within a particular system. It provides a chronological order of events that allows someone to trace the history of a transaction or an activity from its inception to its current state. These transactions or activities may be scientific, financial transactions or communication by individual people, system, or other entities. Audit trails are commonly used in finance, accounting, and other industries to ensure compliance with regulations, detect fraud, and maintain accountability. In information technology, an audit trail may consist of logs that record user activity, system events, or changes to data. The information in the audit trail can be used to troubleshoot issues, investigate security breaches, or analyze system performance. Audit trail is also used to demonstrate compliance with regulations or to provide evidence in legal proceedings. As per the definition of the National Institute of Standards and Technology (NIST), an audit trail is a time-ordered record of system activities that can be used to recreate and review the series of events and actions related to a security-related transaction. It covers the entire process from the start to the end and can provide insight into the steps leading up to the operation, procedure, or event (V. C. Hu et al., 2006).

There are two types of audit trails:

1. External Audit: It is an independent examination of an organization's or department's operational practices by a third-party auditor. It provides an unbiased assessment of the entity's internal controls and compliance with relevant standards and regulations (Pramukti, 2024).
2. Internal Audit: It is conducted by an organization's own internal audit team. It enables an organization/ department to examine its progress and implement measures to foster future expansion. It also identifies and mitigates risks and ensures compliance with internal policies and procedures (Noviani and Trisnaningsih, 2024).

All types of audits can be important for an organization's compliance with legal and regulatory requirements. External audits provide an objective assessment of an organization's practices, while internal audits can help identify and address internal control weaknesses and risks. Organizations should be prepared to cooperate and provide requested

documentation and information for any type of audit they may be subject to, to ensure a thorough and efficient audit process.

2.1.1. Traditional Methods of Creating and Maintaining Audit Trails

Centralized systems have traditionally been the foundation for managing audit trails (Talha, 2024). They provide unified control, allowing for consistent policy implementation and straightforward management of audit data. The consolidation of information in a single location simplifies access, and organizations have well-established practices for managing these systems. However, these systems have downsides. They are prone to failures due to their single point of failure (Garip et al., 2023), making them vulnerable to data breaches. Additionally, they can be tampered by individuals who can access and compromise the integrity of the audit trail. Limited transparency due to restrictive access and the potential for errors in data further reduces their reliability (Theodorakopoulos et al., 2024). Manual auditing, which is less common these days, offers simplicity and flexibility. It does not rely on complex technology and allows for the maintenance of records in various formats, including paper, which can be useful in specific scenarios. However, manual auditing is prone to human error, which can lead to inaccuracies. The process is also time-consuming and labor-intensive (Verevka and Shen, 2024), and paper documents are vulnerable to loss, theft, or unauthorized access, raising concerns about integrity and security.

2.1.2. Impact of Regulations and Technological Advancements

Regulatory changes have impacted audit trail practices. Increased regulations and compliance requirements now demand more secure and reliable auditing. These regulations aim to enhance audit trails and ensure they meet higher standards (Saleh et al., 2023). Privacy and data protection laws such as GDPR (Tikkinen-Piri et al., 2018) and California Consumer Privacy Act (CCPA, 2024) have introduced stricter rules for managing audit trails. These regulations define the scope and process for exercising audit authority, establish criteria for selecting individuals to audit, and protect consumers' personal information from disclosure unless a court order, warrant, or subpoena is issued (CCPA, 2024). The advancements in technology have significantly transformed audit trail management. The decentralization feature of blockchain and the shift from paper to digital records have improved reliability and accuracy by reducing errors. Such advancements have also enhanced the ability to manage and analyze audit trails more effectively. Additionally, blockchain technology has revolutionized audit trails by providing immutable and transparent records (Schulman, 2019). Its decentralized nature and inherent properties address many limitations of traditional systems, reducing the risk of tampering and fraud and improving data integrity (Doekhi, 2023).

Overall, an audit trail can provide valuable insights into an organization's activities and help ensure accountability and compliance (Pramukti, 2024). By implementing a blockchain-based audit trail, organizations can access a plethora of advantages, including transparency, exceptional availability, immutability, and non-repudiation. Hence, they can leverage the numerous benefits that blockchain technology provides.

2.2. Blockchain in Audit Trails

The literature review is structured around four key categories, each highlighting a crucial aspect of auditing and data management in contemporary contexts.

1. Auditing and Remote Audits (Governance): Auditing and Remote Audits play a critical role in maintaining transparency, accountability, and compliance across diverse domains.
2. Financial Audits: Financial Audits concentrate on the meticulous examination and evaluation of financial records, statements, and transactions to guarantee their accuracy and reliability.
3. Enterprise Applications: Enterprise Applications are the applications or solutions employed by corporate bodies to effectively manage a diverse range of business-related processes.
4. Data and Records: Data and Records focus on verifying the proper management and sharing of data and protection of records within and across organizations.

2.2.1. Auditing and Remote Audits (Governance)

Wang et al. (K. Wang et al., 2020) presented how blockchain technology could affect the auditing profession. A debate was raised by some believing that it could replace auditors, while others think it will add value to the industry. The authors suggested a conceptual model for an auditing information system that takes advantage of blockchain's features. The proposed system combines blockchain technology, smart contracts, automated auditing tools, and traditional auditing practices, which should enhance auditing efficiency and be cost effective by offering some features, such as real-time transaction recording, inventory monitoring, suspicious warnings, and analysis and reporting of auditing data.

Jayashreea et al. (Jayashree et al., 2021) proposed the utilization of blockchain technology for digital voting (I-Voting) to enhance the speed, cost-effectiveness, and convenience of public elections. The proposed I-Voting system depends on blockchain

technology and enables any registered voters to vote using any internet-connected devices. The authors highlighted the drawbacks of traditional voting methods and the advantages of online voting and noted the security issues that prevented many countries from adopting such systems. The proposed I-Voting system uses encryption, smart biometrics and real-time ID verification to ensure anonymity and security of the voting process. The authors also emphasized blockchain adoption to create tamperproof audit trails for voting and discussed the use of hash functions to secure user voting details.

Mccorry et al. (McCorry et al., 2021) discussed how blockchain technology can be utilized in electronic voting to ensure results are secure and verifiable. Traditional paper-based voting systems do not allow voters to confirm that their votes are accurately recorded, and digital voting systems that use Direct Recording Electronic (DRE) machines could record a different candidate without the voter realizing. A Voter Verifiable Paper Audit Trail (VVPAT) can help solve this problem of voters not being able to verify their votes are accurately recorded, despite the system still does not allowing voters to check their votes. The authors presented a thorough analysis of secure e-voting over blockchain technology and provided solutions for several types of voting scenarios.

2.2.2. Financial Audits

Wang et al. (Q. Wang et al., 2020) presented a solution to the difficulties of verifying energy bills in Internet of Things (IoT)-based systems. The challenges arise due to privacy issues and the need for bulk auditing requests. The proposed solution, named pAuditChain, utilizes homomorphic encryption and blockchain technology to ensure data auditing and privacy protection. Additionally, it uses a certificateless signature approach to enhance the efficiency of bulk auditing requests. Both individual users and authorities can use the system. The theoretical analysis confirms the privacy-preserving and auditing procedures. This is the first solution to secure bills generated by IoT smart meters while maintaining the auditing function. The approach can also be applied to other IoT devices.

Constantinides and Cartlidge (Constantinides and Cartlidge, 2021) proposed the utilization of blockchain technology to safeguard periodic double auctions for financial "dark pool" trading, while maintaining the privacy of pre-trade order details and verifying the actions of the auction operator. Dark pools are non-displayed financial trading venues designed for large volume traders to avoid market impact, but lack of transparency is a concern to regulators. In (Constantinides and Cartlidge, 2021), the authors introduced an innovative smart contract protocol for implementing periodic double auctions on a blockchain in a way that is verifiable and privacy-preserving. Traders first submit encrypted order commitments that the operator cannot read. During the auction matching phase, the

orders are revealed to the operator, and the results of the off-chain auction are made public. The protocol offers traders guarantees about their order execution. The smart contract protocol is adaptable and can be conveniently converted into a commercial system. The authors evaluated the protocol, which showed clear merits over the closest alternative protocol in the literature. Their conclusion was that a double auction that preserves privacy and has a publicly verifiable audit trail has advantages in dark pool financial trading and various application areas.

Handoko et al. (Handoko et al., 2022) discussed how blockchain technology can support auditors in remote audits, ensure data integrity, and detect suspicious financial reporting. Their study involved interviewing experienced auditors and examining relevant literature. It also showed that the characteristics of blockchain enable auditors to maintain audit trails and detect financial fraud. The success of blockchain adoption depends on factors such as human resources, IT control environment, and Certified Public Accountant (CPA) firm size. However, employing blockchain poses challenges such as cyber-attacks, IT risk, human error, and misuse. Thus, they suggested future research to determine how auditors can address these challenges.

2.2.3. Enterprise Applications

Ahmad et al. (Ahmad et al., 2019) first proposed BlockTrail as a blockchain design for auditing applications, addressing challenges such as transaction bulk and throughput limitations in existing blockchain systems. Their framework introduced a layered architecture that breaks down legacy blockchain systems into co-dependent hierarchies, reducing space and time complexity while improving scalability and efficiency. The authors emphasized the significance of audit trails in enterprise solutions, such as government agencies responsible for property appraisal and tax collection, and how current blockchain systems face significant challenges in terms of space and time complexity due to the volume and size of transactions. To address this issue, BlockTrail employs a multi-chain blockchain model that divides the network into many layers, each capable of processing transactions independently while utilizing the hierarchical structure of eGovernment applications to facilitate parallel transaction processing and reduce storage overhead.

In a later study, Ahmad et al. (Ahmad et al., 2021) refined BlockTrail to further reduce the complexity of conventional blockchain systems, particularly for enterprise business applications requiring audit trails. While maintaining the core layered architecture, the updated version enabled parallel transaction processing, which not only enhanced validation but also significantly reduced storage requirements and improved transaction throughput. This evolution made BlockTrail a more robust solution for secure data

processing without the need for intermediaries. In their later work, they discussed the security measures employed to increase defense capabilities and identify faulty replicas. They also aimed to expand the range of BlockTrail from auditing to the hierarchical model of central bank digital currency (CBDC), which has high scalability and throughput requirements.

Regueiro et al. (Regueiro et al., 2021) discussed the vulnerabilities of current audit trail implementations and proposed a blockchain-based mechanism to address these vulnerabilities. Audit trails are critical for ensuring good internal business control and detecting discrepancies and malicious activities. Despite the present implementations, which are vulnerable to attacks that compromise their integrity, most organizations maintain audit trails in relational databases that can be manipulated or deleted. The authors presented a comprehensive design and implementation of a blockchain-based system for creating protected audit trails by taking advantage of the security capabilities of blockchain technology. The proposed mechanism is general-purpose and can be used in any kind of ecosystem, such as banking and healthcare.

2.2.4. Data and Records

2.2.4.1. Data Management and Sharing

Shafagh et al. (Shafagh et al., 2017) discussed that the current IoT applications' cloud-centric architecture has led to isolated data silos, which restricts the potential of comprehensive data-driven analytics. To address this issue, the authors proposed a blockchain-based design for IoT as an auditable layer that provides distributed access control and data management. The proposed design aims to give users data ownership and enable secure data sharing while ensuring secure and resilient access control management.

Sanchez et al. (Sanchez et al., 2022) discussed the challenges of ensuring secure and trustworthy data sharing in decentralized environments where participants do not trust each other. Ensuring the accuracy and immutability of access logs is one of their significant challenges, as they hold a comprehensive record of data access for auditing purposes. The authors proposed AuditTrust, a secure audit trail for data sharing in a distributed environment that utilizes blockchain technology, including Hyperledger Besu, IPFS, Intel SGX TEE (Trusted Execution Environment), and Vault. Their system permits data sharing and requesting by data owners and consumers while maintaining the confidentiality and integrity of access logs. The evaluation of AuditTrust shows its effectiveness and low latency costs. Their system is designed as a framework for various application domains, where data needs to be exchanged between entities that do not trust each other.

Oakley et al. (Oakley et al., 2023) discussed the challenges associated with guaranteeing data reliability in clinical trials and proposed Scrybe, a permissioned blockchain, as a solution for recording evidence of the source of clinical trial data. The authors contend that the increasing use of context-aware smart devices and wearable IoT devices in clinical trials creates challenges that current electronic data capture and clinical data management systems are inadequate to address. Scrybe aimed to meet the audit criteria that guarantee the confidentiality, integrity and credibility of clinical trial data collected. Scrybe employs a lightweight mining algorithm and provides improved efficiency and cost-effectiveness. The authors also illustrated how Scrybe satisfies all of the relevant controls and offered a proof-of-concept (PoC) integration with Research Electronic Data Capture (REDCap) to demonstrate its resistance to tampering. They concluded that Scrybe has potential applications beyond clinical trial data provenance and that a forthcoming vision of Scrybe will feature smart-contract functionality to enhance its capabilities.

2.2.4.2. Record Keeping

Wang and Yang (H. Wang and Yang, 2021) discussed the utilization of blockchain technology in recordkeeping at the National Archives of Korea (NAK). The NAK built a blockchain recordkeeping platform to conduct research and development (R&D) in two areas of recordkeeping approaches. The first employs blockchain transaction audit trail technology to ensure the authenticity of audiovisual archives, whereas the second leverages blockchain technology to confirm that the datasets of various information systems developed by government agencies were handled without any form of forgery or tampering. They concluded that blockchain-based recordkeeping solutions can provide the record and archive management system with an elevated level of security and stability.

Ermolaev et al. (Ermolaev et al., 2020) addressed the difficulties of maintaining data accuracy and dependability nowadays, where the amount of information is increasing rapidly. To address this issue, the authors proposed the integration of a blockchain and a graph database, which can enhance auditing practices and offer a secure and immutable history record (BitFury Group, 2016). They employed two methods for combining these technologies: "apply-and-rollback" and "two-step". In the first approach, Neo4j database was created to execute queries of modification and change tracking before reporting them to the blockchain, while the latter approach was implemented on the Exonum framework and Neo4j database, creating an auditable trail of data integrity. The blockchain recorded changes to the nodes in the graph database including failed transactions. Direct read requests to Neo4j ensured that the system performs optimally, and Java programming language can be used to add business constraints. They noted that the solution has the potential for complete data recovery, which can be used as a reliable solution to ensure data integrity.

Adlam and Haskins (Adlam and Haskins, 2020) discussed data breaches in electronic health records (EHRs). The traditional approach cannot ensure patient privacy and medical records security (Kshetri, 2018), and audit logs can identify who is responsible for those breaches or manipulation. The authors proposed Hyperledger Fabric, a blockchain system that creates a tamper-proof audit trail for EHRs. Their experiment demonstrated that blockchain-based audit trail could replace the traditional audit logs used by EHRs. They proposed the system as a diagram that shows how data flows, and used smart contracts to investigate how private blockchain technology could enhance the EHR audit log process. Blockchain based audit trail improves the security and privacy of EHRs by offering transparency, immutability, auditability, and decentralization (Emmadi et al., 2019).

Sarode et al. (Sarode et al., 2022) highlighted the significance of maintaining accurate health records, specifically Electronic Health Records (EHR), and stresses the importance of maintaining the reliability of audit trails in these records. To address this issue, the authors proposed a blockchain-based audit management system that allows patients or authorized medical professionals to access a chronological record of all the tests, reports, and medical procedures performed during the patient's treatment. This system can address some issues, such as fragmented medical data across various hospitals and potential data loss due to system failures. The authors recommended using smart contracts to store essential information and maintain data integrity.

Blockchain technology provides an efficient solution for maintaining reliable audit trails by securely storing data in a tamper-resistant manner. This ensures the integrity and immutability of the information, making it a trusted choice for various applications. Its decentralized nature further enhances security and verifiability, addressing concerns related to tampering and data loss. By leveraging these capabilities, blockchain demonstrates significant value in creating robust and trustworthy systems for data management and verification.

Furthermore, blockchain offers substantial advantages in streamlining audit processes. The integration of smart contracts further enhances the efficiency of audit trails by automating key actions such as compliance checks, data reconciliation, and reporting. This automation reduces the need for manual interventions, minimizes errors, and improves overall operational efficiency, making audit processes more robust and reliable.

Investment and development in blockchain-based audit trail systems have the potential to introduce innovative frameworks that enhance accountability, regulatory compliance, and trust across various industries. As organizations increasingly seek reliable and efficient methods to track and verify records, blockchain-based audit trail systems can

be further refined to provide transformative solutions for practical, real-world engineering applications. This is particularly valuable in scenarios where efficient data verification is crucial.

2.3. Blockchain's Impact on Audit Trails

Blockchain offers a promising solution to many traditional audit trail issues. Traditional audit trails are often vulnerable to tampering and unauthorized access. Blockchain addresses that using its decentralized nature; the data is distributed across multiple nodes rather than being stored in a single central repository. This makes it much more difficult for malicious actors to alter or access the data without being detected. Furthermore, each block on the chain is linked to the previous one, creating a chain of blocks that is nearly impossible to alter. In financial and real estate industries, JPMorgan Chase uses blockchain to enhance transaction process to be completed in hours as compared to their original process that takes days (Shevchenko and Lunsford, 2023), and companies like Propy are using it to facilitate real estate transactions, making the process more efficient and cost-effective (Coinmagazine, 2023).

2.3.1. Comparative Analysis: Pre- and Post-Blockchain Integration

Audit trails are essential for maintaining a reliable record of specific events within organizations. Traditional audit trails that rely on centralized systems often encounter challenges related to transparency, immutability, data integrity and accessibility. The introduction of blockchain has significantly transformed these aspects, offering enhanced capabilities and benefits (Usul and Alpay, 2024). Table 2.1 compares audit trails before and after utilizing blockchain technology, illustrating the improvements brought by this adoption.

Blockchain technology has emerged as a promising solution for enhancing audit trails in various domains, including auditing and remote audits, financial audits, enterprise applications, and data management and sharing and record keeping (Peter and Vukovljak, 2023). By utilizing blockchain technology for audit trails, the immutability and integrity of financial records are guaranteed, resilience auditing processes in enterprise applications are ensured, and proper management and sharing of data and protection of records are also verified. This innovative approach to audit trails holds the potential to improve transparency, efficiency, reliability and integrity of audit trail practices across multiple domains, offering a sturdy basis for maintaining accurate and reliable audit trails.

Table 2.1. Comparing audit trails: before and after blockchain

Aspect	**Before Blockchain**	**After Blockchain**
Transparency	Limited transparency due to restricted access and centralized systems	Inherent transparency with decentralized ledger
Immutability	Prone to tampering; changes or deletions can occur	Immutable records that cannot be tampered
Data Integrity	Using a centralized system can be prone to errors and malicious activities	The decentralized nature helps verify data integrity
Cost	Manual processes; time-consuming and costly	Automated processes; save time and reduce operational costs
Accessibility and Availability	Centralized access; limited remote or external audit capabilities, and limited availability	Decentralized access; cross-organizational audits, remote accessibility, and high availability

2.4. Research Gaps

One critical gap in existing research on blockchain-based audit trail systems is the overwhelming focus on the immutability of information, with little attention given to the retrieval of relevant verification data. While immutability ensures data integrity, the inability to efficiently retrieve verification information undermines the practicality of these systems, especially in real-world applications where prompt and reliable information retrieval is crucial. Without efficient retrieval mechanisms, blockchain-based audit trails risk becoming impractical, limiting their adoption. This thesis addresses this overlooked issue by designing an efficient audit trail framework that not only capitalizes on blockchain's immutability but also facilitates instant access to verification information. By addressing this challenge, our study contributes to making blockchain-based audit trail systems efficient, practical and scalable for real-world applications.

3. BACKGROUND

3.1. Blockchain Technology

Blockchain is a decentralized and distributed ledger that records transactions in a sequential chain of data blocks. These blocks are linked and maintained across a peer-to-peer (P2P) network of nodes, ensuring security, availability, and immutability. Blockchain enables trustless and tamper-resistant data sharing and transaction processing without the need for a central authority. This technology has gained widespread attention in the IT industry and is now implemented in various real-world applications. Blockchain offers several advantages; it ensures secure and direct transactions between parties, reducing the risk of fraud. Additionally, users have access and control over their data and transactions, and the information stored on the blockchain is tamper-proof, ensuring transparency, availability and security. Blockchain also verifies data integrity, and transactions are permanent and cannot be altered or deleted (Rodeck and Curry, 2022). Blockchain accelerates transaction speeds, particularly within the banking sector, while simultaneously reducing costs through the elimination of intermediaries (Burghart and Wagner, 2024).

Blockchain can be classified into the following categories:

1. Public blockchain: Anyone can participate in the network, read, and write data (Rodeck and Curry, 2022). Examples of public blockchains include Bitcoin and Ethereum (Guegan, 2017).

2. Private blockchain: Access to the network is restricted to an organization or a group. A permissioned blockchain necessitates that the governing authority grants access privileges to individuals or entities to participate in the network (Rodeck and Curry, 2022). Private blockchains are often used by businesses for internal operations and supply chain management.

Blockchain serves as a robust tool for creating an audit trail. In blockchain, as illustrated in Figure 3.1, users submit transactions, which are then grouped together to create a block. In an audit trail context, each log or entry, such as an audit log or a service record, can be represented as a transaction. These transactions form a block when they are combined. Nodes within the blockchain network actively participate in mining—a

competitive process aimed at adding new blocks to the blockchain. To ensure the block's validity, all nodes must approve and accept its proper generation before it is added to the chain. Once a block is created and verified, it is connected to the existing chain, with the collaboration of the miners. The leader is chosen randomly among the miners to provide a secure environment since tampering with the system would require fooling all nodes simultaneously. Selecting the leader can be achieved with the help of consensus algorithms.

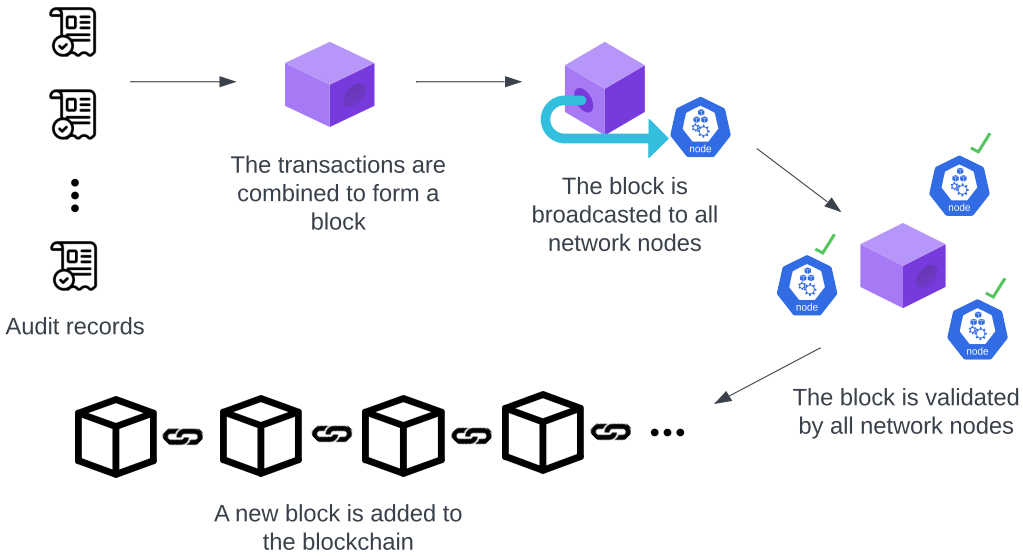


Figure 3.1. Blockchain construction

A block includes information from the previous block, including its hash value, creating a connection between blocks in the chain. This feature strengthens security, as modifying a single transaction would necessitate updating all subsequent information in the chain, which is a challenging task. Additionally, a task can be executed within the blockchain without the need for third-party intervention using automated scripts called smart contracts. Smart contracts operate autonomously on the blockchain, allowing for seamless integration of these procedures during the design phase of the blockchain.

3.1.1. The Advent of Blockchain Technology

Blockchain technology is characterized by some fundamental aspects. Blockchain is decentralized; it is a distributing authority and has control across a network of nodes rather than relying on a single central entity. This approach enhances resilience and reduces the risk of the single point of failure (SPOF). Blockchain is also an immutable system, so once data

is recorded on blockchain, it cannot be altered or erased (Z. Liu et al., 2023). This feature ensures its integrity and prevents tampering. Moreover, all participants in the blockchain network can see and verify transactions, shaping a transparent system that enhances trust and facilitates the detection of fraudulent activities. These inherent features enable the creation of a secure and reliable system that can facilitate auditing practices.

Blockchain technology can be broadly categorized into public and private blockchains, each with distinct characteristics and applications. Public blockchain is open to anyone who wants to read, write or participate, and they can remain anonymous. It provides high transparency since all transactions are visible to every participant in the network. This openness is essential for creating immutable audit trails, as each transaction can be verified and traced. Public blockchain employ consensus mechanisms such as Proof of Work (PoW) or Proof of Stake (PoS) to validate transactions (Rosa-Bilbao and Boubeta-Puig, 2023), ensuring data security and preventing tampering. Public blockchain is suitable for decentralized applications, cryptocurrencies, and public record-keeping. In some sectors such as financial institutions, where regulatory compliance is crucial, public blockchain offers a transparent and verifiable audit trail that can be reviewed by regulatory bodies to verify its integrity.

On the other hand, private blockchains have restricted access; only verified participants can join the network (Zahir et al., 2024), while their identity is revealed. It is suitable for internal use within an organization and restricted data management, such as in legal services. Private blockchain can be used for specific internal compliance and by internal auditors to provide verifiable and auditable records.

Public blockchains are suitable for applications that require open and decentralized verification (Zahir et al., 2024), such as public financial transactions, where audit trails need to be transparent. However, private blockchains are ideal where controlled access and data privacy are essential, such as internal audits or compliance tracking within organizations, where audit trails can be only accessed by authorized participants. Choosing between public and private blockchain depends on the specific use case and the organization or application's requirements.

3.1.2. Sawtooth & Its Architecture

Sawtooth is an enterprise blockchain platform designed for creating distributed ledger applications with a modular and highly scalable architecture. It enables developers to create applications and focus on defining business rules in their preferred programming language without needing deep knowledge of the underlying system. This makes Sawtooth

an excellent choice for businesses requiring customized transaction processing rules, permissions, or consensus mechanisms. Its design facilitates seamless integration into diverse enterprise ecosystems, ensuring both functionality and adaptability (Contributors to Sawtooth, 2024). Sawtooth supports multiple programming languages and provides some features such as verifying integrity, transparency, privacy, and scalability, and it can secure the IPFS (Khan et al., 2023).

In Sawtooth applications, transactions undergo a lifecycle that begins with creation by the client, proceeds through validation and processing by the transaction processor, before being added to the blockchain. Sawtooth introduces the concept of a transaction family, which encompasses transaction creation by the client, a data model (state), and business logic implemented through the transaction processor, which accepts and processes transactions.

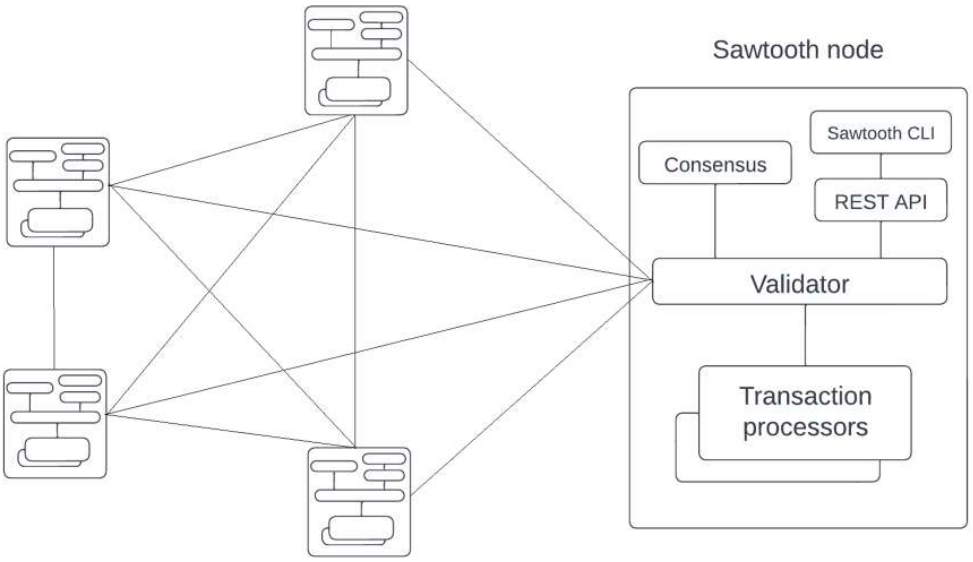


Figure 3.2. Sawtooth network

Within the Blockchain Sawtooth network (Khan et al., 2023), as illustrated in Figure 3.2 and Figure 3.3, each node consists of five components, and these nodes collectively form the Sawtooth network. These essential components are the validator operating on TCP port 4004, and one or more transaction processors responsible for handling transactions received from the validator. There is also a client component that initiates transactions and sends them through the Rest API on TCP port 8008 for processing by the validator. The Sawtooth node incorporates a consensus engine, it employs a

consensus algorithm that determines which node is authorized to publish new blocks within the blockchain. Sawtooth operates as a decentralized network of these nodes, all of which run identical transaction processor(s) and maintain a consistent copy of the blockchain ledger.

In the Sawtooth transaction lifecycle, as shown in Figure 3.3, a client initiates a transaction by submitting it to the validator. The transaction is then propagated throughout the entire validator network, where the validator conducts the necessary processes. Following this, a sawtooth node in the network is elected as the leader, and it publishes a candidate block. The candidate block is then propagated across the Sawtooth network. Validators validate the candidate block; transaction processors validate and execute each individual transaction contained within the candidate block. Once validated, the validators place the verified block on their blockchain and update the overall state (Capocasale et al., 2023).

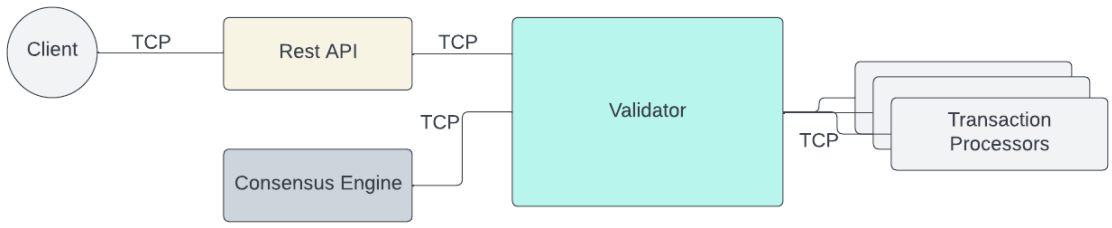


Figure 3.3. A Sawtooth node

Sawtooth stands out for supporting multiple languages for transaction processors, including Python, Rust, JavaScript, C/C++, Java, Go, and .Net. The client, on the other hand, can be developed in different languages, and they are not required to match the language used for transaction processors. Transaction families have an important role in Sawtooth framework; they serve as validators between the blockchain and the blockchain client, and they have the responsibility of determining which transactions are included on the blockchain and which are not. What sets Sawtooth apart from other blockchain frameworks is the freedom it grants to developers; they can define their own transaction families based on their specific requirements.

Sawtooth also supports multiple consensus mechanisms, including Proof of Elapsed Time (PoET), Practical Byzantine Fault Tolerance (PBFT), Raft and Devmode. Smart contracts can be written in popular languages such as C++, Go, Java, JavaScript, Python,

Rust, and even Solidity through the Seth integration. For state storage, Sawtooth utilizes a central LMDB database, offering an efficient solution for storing data in the network. Sawtooth's features and characteristics make it a highly adaptable platform and a powerful tool, making it well-suited for decentralized applications (dApps) (Nikolić et al., 2023).

3.1.2.1. Sawtooth Transaction Process

Submitting a transaction to the Sawtooth involves a structured process that ensures the integrity, security and validity of transaction. The steps are as follows:

- Encode Data: First, the document or data is transformed into an encoded transaction payload. Encoding not only optimizes the data for secure and efficient transmission but also ensures compatibility with the blockchain's data-handling requirements.
- Generate Address: Every piece of data on the blockchain requires an identifier. This address can be created using a combination of a predefined namespace and specific details tied to the account, such as the transaction family and public key. This ensures a globally unique and tamper-proof address that represents the account in the blockchain's state.
- Create Transaction: The encoded payload, along with necessary metadata and signer details, is bundled into a transaction. This encapsulated transaction represents the action to be performed, such as creating an account or storing a hash value, as in our proposed system.
  - Create Batch of Transactions: For multiple related operations, such as adding multiple hash values, transactions are grouped into a batch, which is submitted to the validator network.
- Transaction Processor Validation: The transaction processor, which implements the business logic of the transaction family, validates the transaction. If the transaction is valid, the transaction processor prepares to update the blockchain's global state.
- Submit to Validator: The batch of transactions is submitted to the blockchain validator. This can be done via REST APIs, an SDK like the Python SDK, or directly through ZeroMQ (Contributors to Sawtooth, 2024). The validator checks the integrity and validity of the transaction and appends it to the blockchain.
- Consensus Engine Approval: To maintain the blockchain's integrity, all participating nodes must agree on the validity of transactions. The consensus engine executes an algorithm to achieve this, such as PBFT or PoET. Once consensus is reached, the transaction is allowed for inclusion in the blockchain.

- Update Blockchain State: After the consensus and validation processes, the transaction processor updates the blockchain's global state. This involves storing the document metadata or derived information in the Merkle-Radix tree at the designated account address. The result is a tamper-resistant and traceable record within the blockchain.

This structured process is depicted in Figure 3.4. The transaction process ensures the generation of a tamper-proof, traceable, and securely distributed record, thereby enhancing data integrity and security in distributed ledger applications (Contributors to Sawtooth, 2024). It also emphasizes Sawtooth foundational concepts, such as data encoding, transaction validation, consensus mechanisms and state updates.

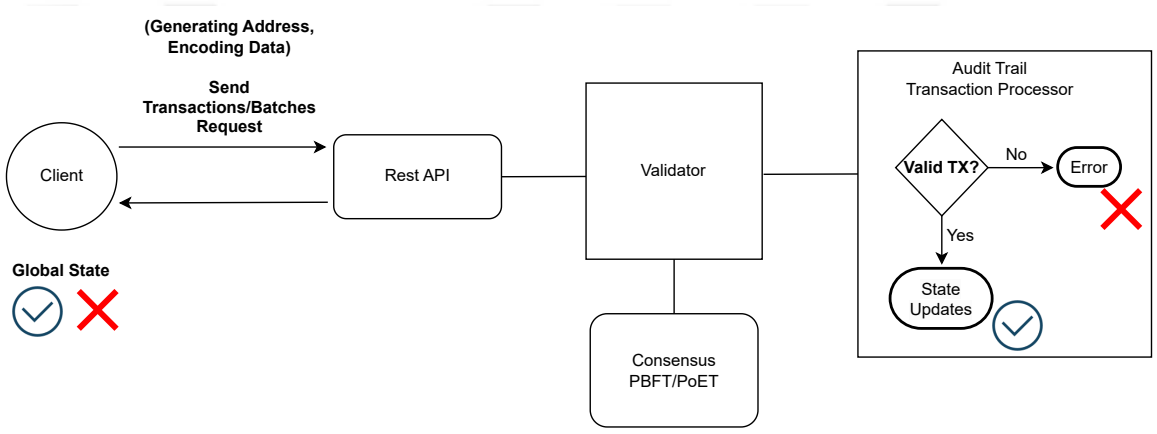


Figure 3.4. Validating transactions and updating the global state in Sawtooth.

3.1.3. Searching in Blockchain

Searching for transactions within blockchain includes different methods, such as searching using identifiers, database structures, and authenticated index structures that offer enhanced efficiency and scalability over the traditional linear search.

3.1.3.1. Searching Using Identifiers

Documents are stored in a distributed database rather than on the blockchain. These documents are linked to their unique identifier, which could be keywords, tags, or fingerprints that are stored on the blockchain. When searching for a specific document using specific keyword(s), the one containing one or more of the searched keywords will be retrieved. This system functions similarly to an explorer, facilitating efficient retrieval of relevant documents. While this method simplifies the search process, its effectiveness relies on the accuracy and relevance of the identifiers linked to each document.

Searchable Symmetric Encryption (SSE) has been extensively studied as a means to enable secure data outsourcing and search functionalities while preserving database and query privacy. However, traditional SSE solutions often rely on the assumption of an honest-but-curious server, which may not align with real-world scenarios, leaving potential security vulnerabilities unaddressed. These limitations highlight the need for more robust and decentralized approaches to ensure secure and efficient search operations. In this context, Hu et al. have made significant contributions to decentralized privacy-preserving search schemes by leveraging blockchain technologies. Their earlier work, Hu et al. (S. Hu et al., 2018), proposed a decentralized search scheme built on the Ethereum blockchain. This scheme eliminates the need for data provider verification by utilizing smart contracts to ensure accurate and immutable search results. Additionally, they introduced a fair mechanism to treat all participants equally and encourage correct computations. However, this approach faces challenges related to high overhead and scalability, primarily due to the substantial costs associated with executing complex search tasks on smart contracts.

Building on their previous work, Hu et al. (S. Hu et al., 2019) presented an enhanced decentralized privacy-preserving search scheme designed for large datasets. Unlike their earlier proposal, this scheme operates within a private blockchain network, restricting participation to authenticated entities. While it shares a similar encrypted index structure with the previous proposal, the use of a faster consensus algorithm in the private blockchain significantly improves efficiency. Nevertheless, both public and private blockchain implementations incur increased costs due to the computational overhead of smart contract operations for searching, updating, and computing.

These decentralized SSE schemes represent a significant departure from traditional SSE approaches, which rely on central servers. By leveraging blockchain technologies, Hu et al.'s proposals ensure correct and immutable search results without requiring verification from data owners, thereby enhancing security against malicious adversaries. Despite their advancements, challenges such as high operational costs and scalability issues remain critical barriers to widespread adoption. Addressing these challenges will be essential for the future development and practical implementation of decentralized privacy-preserving search schemes.

Ali et al. (Ali et al., 2022) addressed the critical need for security in patient health records (PHR) by proposing an approach that combines blockchain technology, homomorphic encryption, and deep learning. Traditional keyword searches for PHR stored in databases pose security risks, prompting the development of a secure searchable blockchain system. Unlike traditional techniques that focus solely on data storage, their proposed approach utilizes homomorphic encryption to enable secure data access via

keyword searches. The study presented an extended approach leveraging blockchain technology and deep learning to ensure secure keyword searches in digital healthcare systems. By training models to detect and monitor attacks like DoS and DDoS, the system enhances security and reduces breaches in healthcare data. The proposed method allows users to encrypt data locally and upload it to the blockchain ledger securely, enabling searches without decryption through homomorphic SSE. Evaluation of the suggested approach against benchmark models demonstrates its effectiveness in improving security, anonymity, and user behavior monitoring. The proposed system supports flexible access control policies, distributed data, redundancy, and fault tolerance. Simulation results conducted on the Hyperledger Fabric tool validate the efficacy of the proposed method compared to existing models like Medrec, Medchain, and Medbichain.

Nie et al. (Nie et al., 2022) proposed a new blockchain-based system designed to securely share electronic health records (EHRs) through IoT devices. This system employs searchable encryption and hidden data structures to ensure both data confidentiality and integrity during the sharing process. EHR ciphertexts are stored in IPFS, allowing users with appropriate access permissions to search for and retrieve relevant data. The system maintains time-bound authorization and authenticates search results to ensure secure access. The scheme utilizes smart contracts to implement secure search, time control, verifiable keyword search, fast search, and forward privacy in IoT environments. Performance analysis and proofs demonstrated that the proposed protocol successfully meets its design goals, offering high scalability and feasibility. Additionally, performance evaluations highlighted the scheme's efficiency, showing lower communication overhead and computational costs compared to existing alternatives.

Tahir et al. (Tahir and Rajarajan, 2018) introduced a framework aimed at preserving data privacy while improving the efficiency of keyword searches on encrypted data within a permissioned, distributed ledger blockchain network. The framework utilizes trapdoors generated from a master key, an index table, and smart contracts to defend against distinguishability attacks, thereby enhancing search security. By utilizing these trapdoors, only encrypted transaction identifiers are disclosed, necessitating decryption for transaction ID retrieval and subsequent data block retrieval. However, if multiple correlated transaction IDs are returned, traversing correlated data blocks sequentially can lead to time-consuming processes. Their study lacks evaluation and comparison, leaving the search performance unknown. Their framework aims to enable keyword search directly on blockchain-stored data, prioritizing both data utility and privacy preservation through the integration of traditional Searchable Encryption techniques with permissioned blockchain technology, particularly Hyperledger Fabric.

Jiang et al. (S. Jiang et al., 2019) present a cooperative search scheme in blockchain-based data markets to address the cost implications of multiple queries on the client side. Their scheme utilizes smart contracts and the gas system to allocate pricing costs for each query to both data owners and miners. By aggregating multiple queries into groups, costs can be reduced, provided that the resulting integrated query size does not significantly exceed that of the original queries in terms of keyword count. However, the proposed strategy focuses solely on keyword-based query conditions and group division, neglecting other important factors such as spatial-temporal constraints, which may lead to additional computation, communication, and gas costs. Their study discusses the potential of blockchain-based decentralized storage services to address privacy and security concerns in cloud-based data markets. It introduces an Ethereum-based data market with a new pricing model, where queries involve compensating both data owners and miners. A cooperative search scheme is proposed to decrease client-side costs by grouping queries based on keyword count, while a testbed analyzes the impact of pricing on grouping outcomes. Various incentive mechanisms and cost-sharing strategies are explored to promote fair distribution of costs among users, with experimental results demonstrating cost reductions for both groups and individuals.

Cai et al. (Cai et al., 2018) proposed a dynamic and efficient scheme for performing encrypted keyword searches on decentralized storage platforms, designed to support ecosystem sustainability. The scheme improves search efficiency by optimizing document and index locality, ensuring that each document and its associated index are stored on the same service peer. This approach significantly reduces communication overhead during search queries. Additionally, the scheme leverages smart contracts to log encrypted search activities on the blockchain, facilitating a fair dispute resolution protocol. However, this approach assumes the honesty of the dispute resolution committee and does not account for potential joint fraud or hacking attempts.

Zhang et al. (C. Zhang et al., 2021) introduced a scheme for enabling authenticated keyword searches in hybrid-storage blockchain systems. This scheme is built around two core components: the Suppressed Merkle Inverted Index, which ensures efficient query processing and result verification with logarithmic maintenance costs, and the Chameleon Inverted Index, which achieves a constant maintenance cost by leveraging chameleon vector commitment. Authenticated query algorithms are specifically developed for these indexes. The study includes formal security proofs and evaluates the scheme's performance using real Ethereum datasets. Experimental results show that their proposed approach delivers high query performance while significantly reducing on-chain maintenance costs. However, given the untrusted nature of storage service providers, clients must verify the soundness and completeness of the search results.

Liu et al. (X. Liu et al., 2024) introduced a trustworthy and reliable multi-keyword search (TRMS) system in a blockchain-assisted cloud-edge storage environment. This system allows users to perform faster searches using edge servers or broader searches using cloud servers, depending on their needs. To ensure the reliability of search results and identify trustworthy servers, the system leverages blockchain-based smart contracts. These smart contracts execute search algorithms and update a score-based trust management model, which records search results and trust scores on the blockchain. This enables users to verify the accuracy of the results and select more reliable servers based on the recorded trust scores. Their proposed system was implemented on the Hyperledger Fabric platform, utilizing the score-based trust management model that is updated via smart contracts. Both search results and trust scores are published and stored on the Fabric ledger. Additionally, the system incorporates multi-threading technology to improve the efficiency of index construction and search processes. A threshold-based depth-first search algorithm is also designed to minimize the number of nodes accessed in the index tree, further optimizing performance. Test results demonstrate significant reductions in time cost for index construction and search, validating the security and efficiency of their proposed scheme. However, the authors acknowledged limitations and challenges such as the inability to evaluate the relevance of content stored in multiple servers and the need to optimize data processing for devices with poor performance.

Morishima et al. (Morishima and Matsutani, 2018) discussed the challenges faced by blockchain systems due to the high workload of full nodes, which verify all transactions in the network. Most user nodes in blockchain systems do not act as full nodes because the workload is too heavy for personal mobile devices. Therefore, search queries directed at full nodes—such as confirming balances and transaction histories—can create bottlenecks due to the limited number of full nodes compared to the number of users. To address this issue, the authors propose an acceleration method for blockchain search using GPUs. They introduce an array-based Patricia tree structure optimized for GPU processing, leveraging the fact that blockchain queries involve no update or delete operations. This structure enables efficient key searches without storing all existing keys in GPU memory, allowing management of many keys within a single GPU. Experimental evaluations compare the proposed method with existing GPU-based key-value search and conventional CPU-based search in terms of blockchain key search throughput. Results show that their proposed method achieves significantly higher throughput. Even when all keys cannot be stored in a single GPU, the proposed method still outperforms CPU-based search by 4.8 times.

3.1.3.2. Searching Based on Database Structures

This approach involves utilizing databases, such as a relational database, where transactions are stored within the blockchain, while identifiers are stored in the database. In this case, the search is performed in the database easily. This method raises verification concerns, prompting several authors to propose solutions to address the issue.

Peng et al. (Z. Peng et al., 2019) introduced a Verifiable Query Layer (VQL) designed to address the inefficiency and authenticity concerns associated with data queries in blockchain-based systems. VQL serves as a middleware layer, extracting transactions from the underlying blockchain and reorganizing them in databases to offer various query services for public users. To prevent tampering with data stored in the middleware, each constructed database is assigned a cryptographic hash value. This database fingerprint, along with relevant properties, undergoes verification by miners and is then stored in the blockchain. The authors implement VQL and conduct experiments using Ethereum. Test results show that VQL effectively supports different data query services while ensuring the authenticity of query results within the blockchain.

Xie et al. (Xie et al., 2019) addressed the privacy concerns of lightweight clients in the Bitcoin network, which currently rely heavily on full blockchain nodes for computational and storage tasks, sacrificing privacy in the process. Traditional solutions fail to retrieve relevant transactions efficiently and securely from full nodes without revealing sensitive information. To address this, the authors proposed a new approach utilizing private information retrieval (PIR) to enable lightweight clients to fetch transactions from untrusted full nodes without disclosing specific transaction details. They introduced an optimized protocol employing two-step PIR to achieve efficient transaction retrieval while safeguarding query privacy. A prototype implementation validates the feasibility of their proposed scheme. By employing two-step PIR, the lightweight client can enhance user privacy without significantly increasing communication overhead. Compared to traditional schemes, their approach offers notable improvements in user privacy. Their solution enables lightweight clients to query recent transactions without compromising privacy.

Pratama and Mutijarsa (Pratama and Mutijarsa, 2018) addresses the usability challenges of blockchain from both developer and user perspectives by proposing query functionalities within a query layer system. This system aims to simplify access to blockchain data. Three main query functionalities are discussed: retrieval queries, which enable users to locate blockchain data using multiple search parameters; aggregate queries, which provide statistical analysis of blockchain data; and ranking queries, which sort blockchain data based on specific components. The implementation leverages Ethereum as

the blockchain platform, MongoDB as the cloud storage service, and REST API for web services. To evaluate performance, throughput and response time metrics were analyzed. The results indicate that the query layer system's throughput is lower than Ethereum's native service for blockchain data retrieval. Additionally, the system's response time is influenced by factors such as the number of threads and the volume of data stored in the cloud.

Muzammal et al. (Muzammal et al., 2019) introduced ChainSQL, an innovative open-source system that bridges blockchain and traditional databases to overcome the limitations of both. Blockchain excels in resisting data modifications but lacks efficient search query capabilities due to its suboptimal data formatting. Conversely, distributed databases offer rapid query processing and well-structured data organization but struggle with data reliability. ChainSQL merges the strengths of these technologies, combining blockchain's decentralization, distribution, and auditability with the speed and data organization of distributed databases. The system delivers several key features: a tamper-resistant and consistent multi-active database, a reliable and cost-effective disaster recovery solution at the data level, and an auditable transaction log mechanism. These capabilities are demonstrated through three use cases: a tamper-resistant multi-active database, a disaster recovery backup at the data level, and an audit middleware solution. Therefore, ChainSQL provides a secure, efficient and versatile approach to data management. Experimental evaluations demonstrate its effectiveness, showcasing ChainSQL as the first system of its kind to combine blockchain's tamper-resistance with the fast query processing of distributed databases. Its utility is demonstrated across various business domains, such as finance and supply chain, highlighting promising applications for the future.

3.1.3.3. Searching Based on Authenticated Index Structures

The authenticated index structure involves creating a specialized data structure. These authenticated data structures can be integrated into relational databases to enhance security, improve efficiency, and accelerate data retrieval processes.

Zhu et al. (Zhu et al., 2019) introduced SEBDB, a novel blockchain database designed to address the usability and scalability challenges inherent in existing blockchain platforms. While blockchain has found applications in various fields to establish trust among multiple participants, such as supply chain management and digital asset transfer, current platforms lack the convenience and efficiency of traditional databases, especially when handling complex tasks involving both on-chain and off-chain data. SEBDB integrates relational data semantics into blockchain by representing each transaction as a

tuple with multiple attributes in a predefined table. It employs a SQL-like language as the general interface, facilitating convenient application development without requiring code-level APIs. Additionally, SEBDB treats traditional relational database management systems (RDBMS) as an important component, leveraging their decades of optimization and success. Their study outlines a mini benchmark to evaluate the performance of the blockchain database, demonstrating its effectiveness and efficiency through extensive experiments. SEBDB is positioned as the first platform to consider both usability and scalability, making several key innovations such as adding relational semantics to the blockchain system, utilizing SQL-like language as an interface, redefining operations for blockchain compatibility, and proposing a benchmark for evaluating blockchain databases.

Peng et al. (Y. Peng et al., 2020) introduced a shared database system designed to address the challenges of collaboration among different entities over a shared database, called FalconDB. Traditional solutions often require trust in others, demand high hardware resources, or exhibit relatively low performance, creating a trilemma between security, compatibility, and efficiency. FalconDB aims to overcome these challenges by enabling efficient and secure collaboration on a database among parties with limited hardware resources. FalconDB utilizes database servers equipped with verification interfaces that clients can access, while query and update digests are stored on a blockchain for authentication purposes. By leveraging blockchain as a consensus platform and distributed ledger, FalconDB eliminates the need for mutual trust among participants. Their system is designed to minimize client storage costs while enabling real-time, concurrent access to the database from anywhere. With blockchain-level security featues, FalconDB facilitates efficient and secure collaboration among users while maintaining low storage requirements. Their platform integrates a blockchain framework with authenticated data structures (ADS) to establish a public outsourced database. In the setup, server nodes handle the storage of both the database and the blockchain, whereas client nodes only need to store block headers, ensuring lightweight and scalable client participation. This architecture allows FalconDB to withstand a certain degree of malicious behavior among nodes, ensuring robustness and reliability in collaborative database environments.

Lai and Wang (Lai and Wang, 2022) discussed the challenge of long query response times in blockchain-based data traceability systems and proposed an optimized query scheme to address this issue. By leveraging blockchain technology, their proposed scheme aims to enhance data traceability, ensuring the security and efficiency of query services. Key features of the scheme include an index structure independent of the Merkle tree to support efficient intra-block queries and the introduction of new fields in the block header to optimize inter-block queries. Experimental results demonstrated that their proposed scheme achieves higher

query efficiency and lower overhead compared to traditional approaches, thereby improving the effectiveness of product quality assurance through enhanced data traceability.

3.1.3.4. Searching Using Merkle Trees

A Merkle tree is a cryptographic data structure designed to efficiently and securely verify the integrity of transactions or data. It works by organizing data into a hierarchical, tree-like structure through hashing. In this structure, each leaf node represents the hash of an individual transaction, while each non-leaf node is the hash of its child nodes. The root of the tree, called the Merkle root, provides a single, compact fingerprint that encapsulates all the transactions, enabling quick and reliable verification of data integrity.

In transaction verification, the Merkle tree enables quick and secure validation of individual transactions without requiring the processing of the entire dataset. For instance, in blockchain systems, users can verify whether a specific transaction is included in a block by examining a small set of hashes, known as a Merkle proof, rather than the entire block. This approach ensures data integrity and security while maintaining high efficiency, as only a minimal portion of the tree is needed for verification. The Merkle tree's fixed and consistent verification time makes it a reliable tool for systems that demand scalable and tamper-proof transaction validation.

In IIoT systems with heterogeneous devices, the standard Merkle tree storage structure falls short in optimizing verification efficiency for transactions that require frequent checks. To address this limitation, Wang et al. (J. Wang et al., 2021) proposed an enhanced Merkle tree structure tailored for efficient transaction verification in trustworthy blockchain-enabled IIoT systems. Their research analyzed the traditional construction and verification mechanisms of the Merkle tree and introduced improvements to enhance its effectiveness as a cornerstone for ensuring data integrity in blockchain-enabled IIoT environments. However, it identifies a limitation in current Merkle tree structures, particularly in scenarios where transactions require frequent verification. To address this, the authors proposed an optimized Merkle tree structure designed to enhance transaction verification efficiency in heterogeneous IIoT environments. Their proposed structure aims to reduce computational and network transmission requirements while improving verification speed, especially for frequently verified transactions. The effectiveness of the optimized Merkle tree structure is validated through experiments, demonstrating its potential to reduce network load and enhance block verification efficiency in blockchain-based IIoT systems.

3.1.3.5. Other Search Approaches

Tran et al. (Tran et al., 2019) proposed a new search engine correlated with smart contracts in the blockchain, addressing the growing interest in blockchain technology and its distributed peer-to-peer network. Their study highlighted the utility of smart contracts, numeric protocols defining promises between parties, which have led to millions of contracts being created and stored by different users on blockchain platforms. Despite variations in wording, these contracts often exhibit similar functionalities. The system assists users in verifying their smart contracts by referencing existing similar contracts stored on the blockchain. This search engine offers an adaptive set of similar smart contracts tailored to the user's specific similarity needs. The system, developed on the Ethereum blockchain platform, demonstrates its potential through experiments conducted across various scenarios. However, their study acknowledges a limitation: the lack of an extensive performance evaluation for their proposed system.

Malik et al. (Malik et al., 2018) addressed the increasing consumer demand for transparency in food production and handling due to rising incidents of food mislabeling and mishandling. The authors proposed a permissioned blockchain framework called ProductChain. They advocated for a sharded, three-tiered architecture to ensure data availability for consumers, restrict access to competitive partners, and provide scalability for handling transaction load. Additionally, they introduced a transaction vocabulary and access rights management system to regulate read and write privileges on the blockchain. ProductChain safeguards trade flows confidentiality when retrieving provenance information by consumers and stakeholders. Simulation results demonstrated acceptable query times for product ledger retrieval, even when collated from multiple shards. The framework was designed to benefit supply chain participants, consumers and audit control bodies, with security analysis showing resilience against various attacks.

Zhang and Poslad (X. Zhang and Poslad, 2018) introduced GAA-FQ (Granular Access Authorisation supporting Flexible Queries), an architecture designed for Blockchain-based Electronic Medical Records (EMRs). GAA-FQ comprises an access model and an access authorization scheme, offering access control while remaining compatible with the underlying blockchain data structure. GAA-FQ allows authorization at different levels of granularity and improves computation performance by dispensing with the need for a public key infrastructure PKI. The access authorization scheme supports flexible queries without requiring PKI for authorization or encryption/decryption of queried EMRs. Their proposed scheme enhances control over data access, enabling secure EMR information management. Their paper validates the computation performance and transmission efficiency of GAA-FQ through simulations, comparing it with a baseline

scheme called ESPAC, which lacks Blockchain integration. The system provides finer control over access authorization and efficiently responds to queries while protecting unauthorized private data. This makes it particularly suitable for resource-constrained devices in IoT eHealth systems, offering enhanced security and flexibility in managing EMR information on Blockchain platforms.

Ruan et al. (Ruan et al., 2021) proposed LineageChain, a secure and efficient provenance system designed for blockchains. LineageChain addresses the need for capturing and managing provenance information in blockchain systems, particularly for applications that rely on provenance data at runtime. Unlike traditional approaches that require replaying all transactions to query data history, LineageChain provides fine-grained provenance support through simple and elegant interfaces to smart contracts. LineageChain captures provenance during contract execution and stores it in a secure manner using a Merkle tree. It offers efficient provenance queries facilitated by a novel skip list index. The system is implemented on top of the Fabric blockchain platform, and a blockchain-optimized storage system called ForkBase. Their study presented extensive evaluations of LineageChain, demonstrating its benefits for supporting provenance-dependent applications. Test results indicated high query performance and minimal storage overhead, showcasing LineageChain's potential for enabling a new class of blockchain applications that rely on provenance information.

Rahman et al. (Rahman et al., 2022) proposed a hierarchical blockchain-based platform to ensure the integrity of data in smart city IoT systems and address interoperability challenges. It highlighted issues with traditional centralized cloud-based systems and suggested using the blockchain technology to enhance data integrity and transparency. Their proposed solution, called Blockchain-of-Blockchains (BoBs), utilizes a decentralized hierarchy of blockchains. This concept, along with a new data structure called "blockchain tree," aims to manage the complexities of heterogeneous organizations with a hierarchical structure. Their study discusses the benefits of blockchain, such as immutability and security, in ensuring trustable platforms for sensitive data communication. Their proposed solution is evaluated using Hyperledger Fabric and Ethermint as a PoC, demonstrating high efficiency and security. Overall, the authors concluded that the BoBs concept can address data integrity and interoperability challenges not only in smart city IoT systems but also in other scenarios such as supply chains, smart homes, and education platforms.

In blockchain networks, nodes validate transactions and achieve consensus on their order. However, as transaction volumes grow, the demand for processing power also increases. To address this issue, Yoo et al. (Yoo et al., 2018) introduced the domain-based

static sharding, allowing for parallel processing of multiple transactions. In their approach, the blockchain is divided into multiple shards based on specific domains. Each shard operates independently, with nodes validating multiple transactions concurrently and maintaining separate ledgers for their respective shards. This design aims to enable low-latency transaction processing. Additionally, smart contracts are utilized to facilitate fast querying of local transactions and cross-shard transactions. This approach aims to improve the scalability and efficiency of blockchain networks while maintaining transaction security and integrity.

While searching using identifiers, including keyword-based search, offers a straightforward method for locating information on blockchain, this thesis focuses on a more specialized aspect of blockchain transaction retrieval. Existing transaction search approaches, such as searching using keywords that focus on retrieving documents using keywords, can be effective for broader inquiries but lack the precision needed for efficient verification in auditing scenarios. Similarly, searching based on a database structure or an authenticated index structure introduce hierarchical and indexed methodologies that improve efficiency but may still fall short in scalability and performance as the blockchain grows.

Advanced search methods, such as searching using Merkle Trees, provide a cryptographic foundation for verifying data integrity but primarily address the validation of existing records rather than the rapid identification of specific transactions. Existing transaction search methods attempt to balance efficiency and complexity but often introduce overheads that reduce practicality for real-world audit trail systems.

In contrast, our primary objective is to identify and verify whether a specific, unique identifier exists within the blockchain without processing the entire blockchain. This process, also known as transaction search, is critical for verifying the presence of a particular transaction or piece of data without relying on broader or less precise search methods. This type of transaction search bypasses the challenges associated with keyword-based or hierarchical methods, offering a more direct and efficient approach. By narrowing our focus to this specialized search, we aim to address a fundamental challenge in blockchain auditing and set the stage for designing a more efficient and practical verification framework.

3.1.4. Transaction Search Approaches

In the literature, searching for a transaction on the blockchain is addressed through various approaches, each designed to meet specific use cases and efficiency requirements

within blockchain systems. Some blockchain platforms offer search functions that can be utilized, including Sawtooth. Alternatively, one can download transactions from blockchain and perform a linear search. Another approach is to verify and add a transaction ID to a local database with proof. Assuming the system administrator provides a reliable method, such as a smart contract, to update and input transaction information into the database. This allows querying the database instead of the blockchain.

Another option involves using the Merkle hash of a block to verify the existence of a transaction within it. To perform this, a Merkle proof is required. One can query each block in the blockchain, utilize the Merkle proof with the transaction ID to determine if it indeed exists in the blockchain.

3.1.4.1. Linear Search

In linear search, all transactions are considered as a sequential list of items. During the search, we traverse the blockchain one transaction at a time, checking and comparing each transaction to determine if it exists within the blockchain. This method is the most straightforward approach to performing a search within the blockchain, as shown in Figure 3.5.

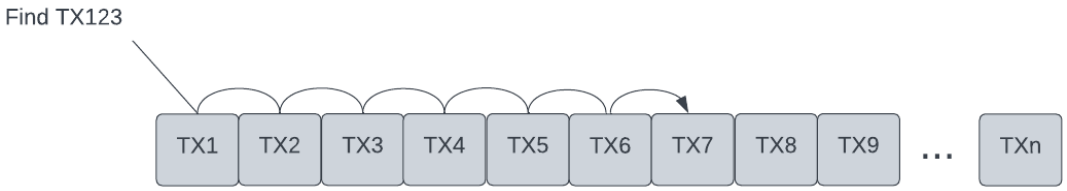


Figure 3.5. Linear search

3.1.4.2. Platform Specific Solutions

Sawtooth utilizes a searching approach to efficiently query and retrieve blocks and transactions within its blockchain infrastructure. This approach can be used to find detailed information about the blocks and transactions, and it can show the other nodes in the network (Contributors to Sawtooth, 2024). This tool can be used from any node's shell containers by specifying the URL of the other node's REST API.

3.1.4.3. Searching Using Databases

This approach involves utilizing a database. Transactions are stored within the blockchain, while their identifiers are maintained in a database. In this case, the search is performed in the database rather then the blockchain, as depicted in Figure 3.6. This approach raises verification concerns, prompting researchers, such as (Yue et al., 2023), to develop solutions that address these issues while enhancing the system's reliability and efficiency. In our experiments, we used SQLite that uses B-tree (Burghart and Wagner, 2024) for organizing indexes, as illustrated in Figure 3.7.

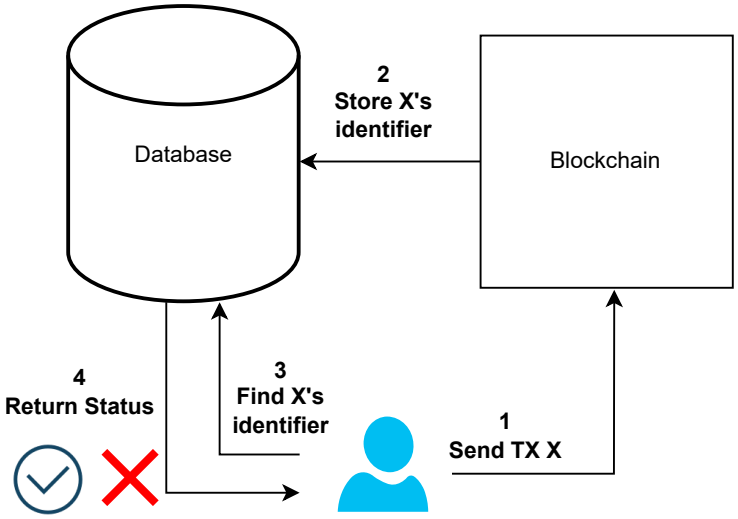


Figure 3.6. Database-based search

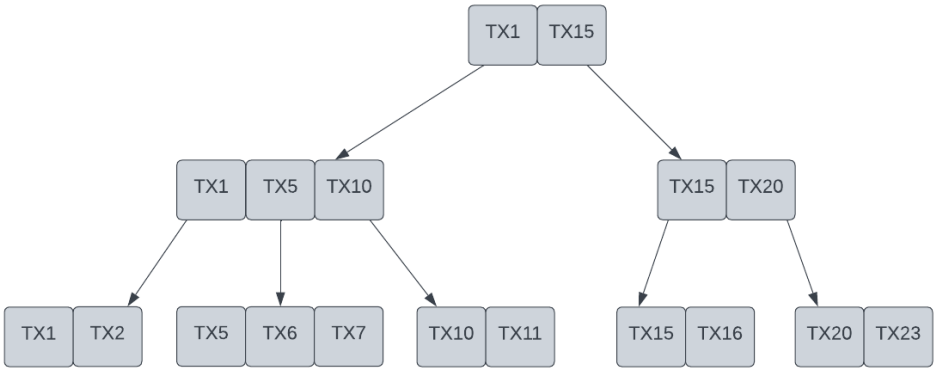


Figure 3.7. B-tree structure

3.1.4.4. Searching Using a Merkle Hash

Blockchain transactions leverage Merkle trees for efficient searching. Merkle trees help verify transactions without downloading all transactions. Each transaction gets a unique hash, and these hashes are paired and re-hashed, constructing a hierarchical structure that converges into a single Merkle root at its highest level. To find a specific transaction, one only needs the transaction itself and Merkle proof (Chen et al., 2023). By following the proof and recalculating hashes at each step, they can verify if the transaction belongs to the block. This system saves bandwidth and processing power, making blockchain searches significantly faster and more efficient.

In our experiments, we utilized the Splinter Community's Sawtooth, previously known as Hyperledger Sawtooth, for our testing. During our experiments, we utilized the RSA-based cryptographic accumulator to verify the inclusion of a transaction, which will be discussed in detail at the end of this chapter. We performed a linear transaction search and assessed its performance. We also retrieved transactions from the blockchain to evaluate the platform-specific solutions. Finally, we designed and implemented a database-based system, storing all transaction details—along with their corresponding hashes—in a relational database, and conducted transaction queries to evaluate the system's performance. However, the Merkle hash approach could not be applied in our experiments as it was unavailable within the Sawtooth framework, limiting our ability to test and evaluate its performance. Nonetheless, this functionality remains accessible and can be effectively utilized on other blockchain platforms.

3.2. InterPlanetary File System (IPFS)

Storing large data directly on the blockchain is inefficient due to the high computational costs associated with it. However, IPFS presents an effective solution. IPFS is a P2P distributed database designed to efficiently store and share data. Rather than relying on a centralized storage location, IPFS operates by fragmenting large files into smaller chunks and distributing them across multiple computers or nodes in a network (Salunke et al., 2023). When data is added to IPFS, it undergoes a structured process to ensure both integrity and high availability. In this thesis, the IPFS is utilized to store company documents. It provides an outstanding solution for maintaining reliable records within a distributed system, making it an effective and efficient way to share and access company documents for auditing purposes.

Storing a document on the IPFS network involves several key steps that ensure the document is distributed, accessible, and persistent in a decentralized environment.

Figure 3.8 illustrates the process, which begins with sending the document to the network. This document is then hashed using a cryptographic algorithm, producing a unique Content Identifie (CID) that represents the document. The CID is immutable and guarantees that any tampering with the document will result in a completely different hash, providing data integrity.

Once hashed, the document is split into smaller chunks, if necessary, and stored locally on the user's node. These chunks are distributed across the IPFS network. During this phase, the local node connects with other peers in the network and makes the document or its chunks available for retrieval. The CID acts as the key for accessing the document; no central server is involved in storing or locating the data.

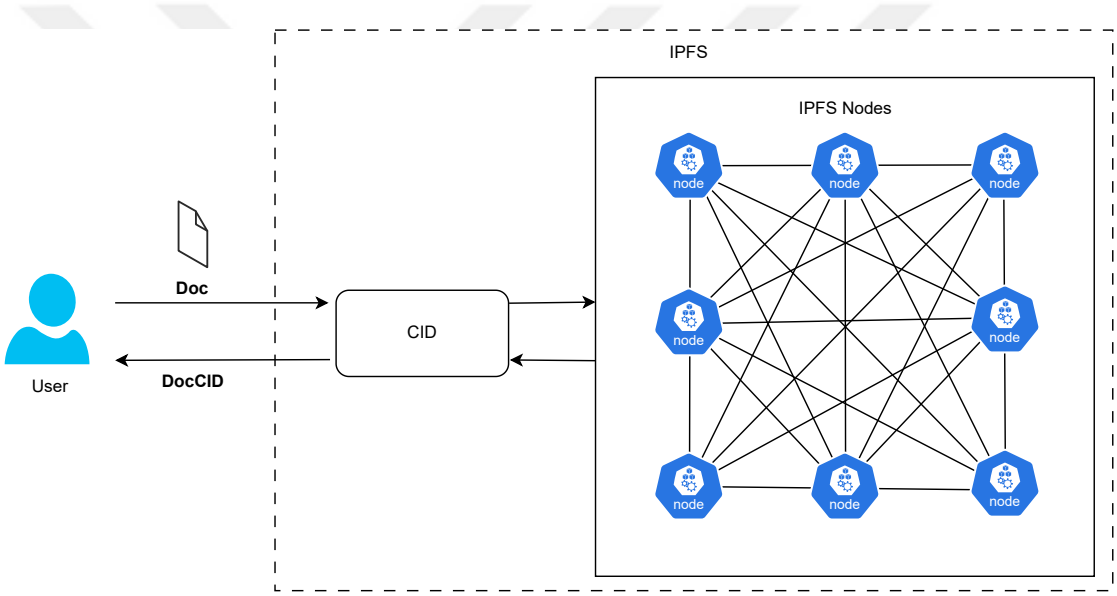


Figure 3.8. Storing documents in IPFS

To ensure data availability, IPFS employs a distributed hash table (DHT)(Confais et al., 2023), where nodes share information about the availability of document chunks. When a user requests a document using its CID, IPFS locates the nodes hosting the chunks, retrieves them, and reconstructs the original document. To share the document with others, the user only needs to share its hash. When they wish to access the document, they can retrieve it from IPFS by using the hash, obtaining a copy of it. This process is illustrated in Figure 3.9.

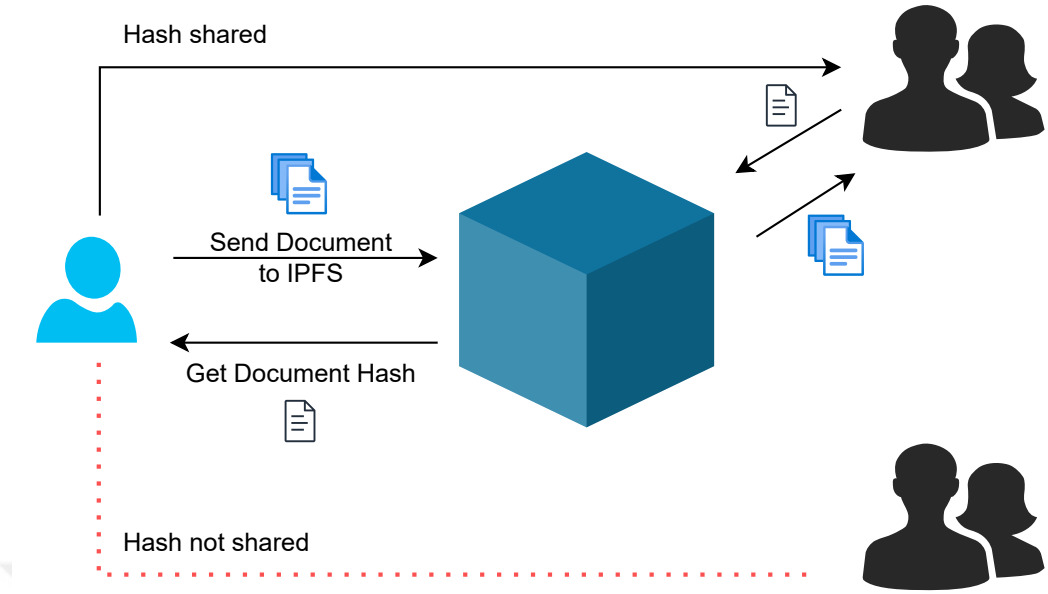


Figure 3.9. Document retrieval and sharing in IPFS

In centralized systems, SPOF can happen and cause congestion and even disable internet access for an entire country. IPFS addresses this issue by directly pointing to resources rather than specific locations. It ensures that data is accessed from the nearest node, reducing the centralization problems and enhancing network availability (Vimal and Srivatsa, 2019).

IPFS utilizes several fundamental concepts that support its decentralized storage and file sharing. These concepts include content addressing, distributed system architecture, MerkleDag, and chunking.

- Content Addressing: all content is uniquely identified, where every document is assigned a cryptographic hash. This content addressing approach ensures that documents remain tamper-proof because any changes that happen to the content will result in changing its hash.
- Distributed sytem: documents are stored and distributed across serveral nodes in a P2P network. This architecture reduces the need for a central authority and enables fault tolerance as well as resilience.
- MerkleDag: IPFS organizes locating and retrieving data using MerkleDag (Merkle Directed Acyclic Graph). MerkleDag is built upon DHT and BitSwap. The DHT enables IPFS to form a large-scale network, where nodes can quickly locate and

retrieve data. BitSwap facilitates data exchange between nodes. MerkleDag provides IPFS with some features including content addressing, tamper resistance by verifying content integrity, and deduplication where identical content is stored only once within IPFS (Benet, 2014).

- Chunks: IPFS divides documents into smaller segments (i.e. chunks), each segment holding a maximum of 256 KBs of data and/or links to other segments. Each of them is stored on a distinct node within the IPFS network (Makina et al., 2022). The CID uniquely identifies each segment, and it is computed from the content of the segment (Steichen et al., 2018).

IPFS eliminates the need for centralized authority, enhancing resilience against censorship and data loss by removing SPOF. Its content addressing method uniquely identifies documents by their content rather than location, ensuring documents remain accessible through the same CID even if moved or replicated across nodes. This decentralized approach facilitates efficient data sharing and retrieval, allowing users to access document parts from multiple nodes simultaneously. Additionally, IPFS supports version control, enabling easy access to previous document versions, and its CID mechanism ensures data integrity by preventing unauthorized modifications. These features make IPFS a robust solution for improving dApps while addressing concerns about centralization, data privacy, and reliability.

Researchers have explored the foundations and concepts of IPFS that signifies a paradigm shift in data storage and access, transitioning from traditional centralized systems to a P2P architecture. For example, Regueiro et al. (Regueiro et al., 2021) introduced a system that improves user interaction with existing audit trail solutions by offering a graphical interface through the blockchain monitor client. This client provides a REST API, enabling users to easily interact with the blockchain and perform essential audit-related functions. Their system generates Blockchain monitor events via smart contracts, which notify users of updates within the audit trail system. By leveraging blockchain as a verification mechanism, auditors can validate audit trail records transparently, while the underlying blockchain remains abstracted from end users.

Their system integrates Kibana, a robust graphical interface, to visualize data from the blockchain monitor. Kibana enables real-time analytics through customizable dashboards, presenting information processed by the blockchain monitor. To ensure security, access to these dashboards is authentication-protected, and roles can be assigned to control the retrieval of specific information. All audit data is securely recorded on a decentralized blockchain, ensuring availability and resistance to tampering. Furthermore,

smart contracts embedded in the blockchain enhance security by providing direct auditing functionalities.

In another development, BlockIPFS (Nyaletey et al., 2019) integrates IPFS with Hyperledger Fabric to ensure data traceability, integrity, and protection of authorship. In their system, only the hash of the data is stored on the blockchain, while the actual documents are encrypted and stored on IPFS. Decryption keys are securely managed within dedicated blockchain channels, ensuring that only authorized users can access the document metadata. Users can query the blockchain to verify document ownership and metadata without compromising confidentiality. Planned enhancements to BlockIPFS include tracking all user activities, such as uploads, downloads, and access logs, as well as improving update handling in IPFS and providing access control for multiple document copies through IPFS.

Kalis (Kalis, 2018) explored implementing automatic audit trails using Apache Isis and blockchain technology. Two methods were evaluated:

1. IPFS and Ethereum Hashing: Audit trail data was stored in IPFS, with corresponding hashes recorded on Ethereum to verify data integrity. While this method confirmed integrity, it could not prevent tampering.

2. Encrypted Audit Trails on Ethereum: Entire audit trail data was stored encrypted on Ethereum, addressing tampering concerns but at the cost of scalability.

The first method was implemented using Apache Isis's AuditerService and PublisherService SPIs, connected to Ethereum via Web3j. Smart contracts linked transaction identifiers with data hashes, enabling validation against the Ethereum blockchain. Despite its innovative approach, this method faced challenges, including application interruptions, difficulties retrieving audit entries, and scalability issues when multiple blockchain transactions were pending simultaneously.

Sanchez et al. (Sanchez et al., 2022) proposed AuditTrust, a system that enables a user to act as either Data Owner (DO) or Data Consumer (DC) within a web application. DOs manage policies, while DCs handle roles, with both utilizing the audit tool for partial or comprehensive audits. The system uses PostgreSQL for database management, supporting concurrent writes without locking the entire database. Communication between the front-end and back-end is facilitated by a RESTful API implemented with FastAPI, while an Nginx server serves as a reverse proxy and web server.

Their prototype was designed to enable secure data sharing between DOs and DCs. During development, it utilized a private, permissioned Hyperledger Besu blockchain, with plans to transition to the Ethereum Mainnet for production. Local PostgreSQL databases are used to replicate blockchain data, allowing participants to effectively monitor network activity. To overcome blockchain limitations, IPFS is employed for storing images, videos, and access logs, ensuring data integrity through CIDs. Confidentiality is maintained through a combination of symmetric and asymmetric encryption, with secure key management facilitated by Intel SGX and Vault.

AuditTrust incorporates robust verification protocols, including authentication and challenge-response mechanisms, ensuring secure interaction. While hashes of shared data and access logs are stored on the blockchain, the actual data resides off-chain, achieving a balance between efficiency and security.

3.3. Cryptographic Accumulators

A cryptographic accumulator is a compact, space- and time-efficient data structure designed for set membership verification. The cryptographic accumulator, as shown in Figure 3.10, allows for the representation of a set of elements in a compact manner, enabling verification of whether a specific element belongs to the set without revealing the entire set. This structure is fundamental in computer science and cryptographic systems due to its space and time efficiency (Ozcelik et al., 2021). In this study, the cryptographic accumulator is employed to accumulate the verification information of transactions, and is stored within the blockchain's global variable. The adoption of this data structure significantly enhances the audit trail system performance by enabling efficient verification of transactions.

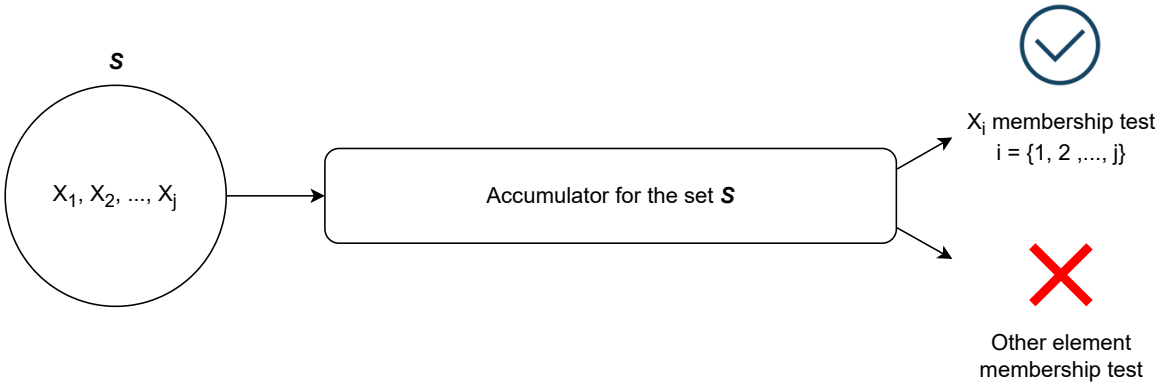


Figure 3.10. Key concept of cryptographic accumulators

Cryptographic accumulators offer an efficient solution for verifying transaction membership in blockchain systems. As blockchain grows, traditional search algorithms become increasingly resource-intensive. However, the cryptographic accumulators overcome this limitation by condensing a set of elements into a compact value. This structure enables fast and verifiable membership proof without the need to search the entire blockchain. By generating and updating witness values, the cryptographic accumulator facilitates efficient verification of transaction inclusion in the blockchain, thereby ensuring both integrity and reliability (Ozcelik and Skjellum, 2021). This approach is crucial for maintaining trust in decentralized systems while minimizing computational overhead. Cryptographic accumulators address several challenges in managing large datasets. Table 3.1 highlights the key featues of using the cryptographic accumulators.

Table 3.1. Key features of cryptographic accumulators

Feature	**Description**	**Benefits**
Space Efficiency	Compact representation of large sets.	Suitable for memory-constrained environments.
Time Efficiency	Sublinear time complexity for membership verification.	Faster than linear search.
Security	Soundness, completeness, undeniability and indistinguishability (Derler et al., 2015).	Critical for authentication and validation systems.
Scalability	Handles dynamic changes to the set (additions/deletions).	Supports dynamic use cases, e.g., blockchain or access control.

Cryptographic accumulators are designed to efficiently manage and verify set memberships while ensuring security. To better understand the classifications of cryptographic accumulators and their use cases, it is essential to examine how they differ in terms of the operations they support and the types of proofs they generate. Tables 3.2 and 3.3 provide an overview of the various types of cryptographic accumulators, categorized based on their capabilities for set modification and membership testing (Ozcelik et al., 2021).

Table 3.2. Types of cryptographic accumulators based on set modification

Type	**Description**	**Use Case**
Dynamic	Allows modifications (additions/deletions) to the set.	Applications requiring frequent updates, e.g., access control lists.
Static	Immutable sets after initial creation.	Immutable databases.
Additive	Only allows additions to the set.	Time-stamped document accumulators or permission grants.
Subtractive	Only allows deletions from the set.	Revocation lists, e.g., certificate revocation in PKI systems (Akkaya and Cebe, 2018).

Table 3.3. Types of cryptographic accumulators based on membership testing

Type	**Description**	**Use Case**
Positive	Proves membership of an element.	Verifying authorized credentials.
Negative	Proves non-membership of an element.	Detecting fraudulent entries or anomalies.
Universal	Handles both membership and non-membership proofs.	Comprehensive access control or decentralized identity systems.

Cryptographic accumulators have become integral to modern cryptographic systems due to their ability to address critical challenges such as scalability, privacy preservation and computational efficiency. Their flexibility and efficiency in handling set-membership proofs make them invaluable across various applications. From blockchain technology and identity management to privacy-enhancing credential systems, accumulators play a vital role in enabling secure and efficient solutions. Table 3.4 summarizes key application areas of cryptographic accumulators, showcasing their broad impact in advancing secure and scalable systems.

Table 3.4. Applications of cryptographic accumulators

Application Area	**Examples**
Decentralized Systems (e.g. Blockchain)	Membership checks, privacy-preserving proofs.
Centralized Systems (e.g. Database Management)	Efficient querying of large datasets.
Authentication Systems	Access control, credential verification.
Digital Signatures	Group signatures, fail-stop signature schemes.
Privacy-Preserving Systems	Anonymous credentials, unlinkable proofs.

3.3.1. RSA-Based Cryptographic Accumulator

The demand for efficient and scalable methods of membership verification has driven the development of cryptographic accumulators. These structures enable the compact representation of large sets of elements while allowing for efficient proofs of membership, making them important for systems requiring secure and efficient validation. However, traditional cryptographic accumulators are inherently static, meaning they lack the flexibility to accommodate changes in the set, such as the dynamic addition or removal of elements. This limitation poses a significant challenge in practical applications where memberships are frequently updated.

Dynamic cryptographic accumulators address this challenge by extending the functionality of traditional accumulators to support efficient updates while preserving their compact representation. This method ensures that the computational cost of adding or removing elements remains constant, regardless of the size of the accumulated set. The Strong RSA Assumption is key to this construction, ensuring security against potential threats. This advancement marks a significant step forward in cryptographic research, particularly in scenarios that require secure and scalable membership verification (Li et al., 2024).

A cryptographic accumulator is a mathematical function that aggregates a set of elements into a single, fixed-size value (i.e. accumulator) while supporting efficient membership proof generation. These proofs confirm whether a specific element is part of the accumulated set (Khedr et al., 2019). The following steps outline the key operations

involved in constructing and utilizing dynamic accumulators (Camenisch and Lysyanskaya, 2002; Li et al., 2024).

1. Setting up the accumulator: The initialization phase involves generating the parameters that define the accumulator:

   - Generate an RSA modulus: $n = p \cdot q$, where p and q are large prime numbers.

     The modulus n defines the multiplicative group $\mathbb{Z}_n^*$, consisting of integers coprime with n under modular multiplication.
   - Select a generator: Choose a generator $g \in \mathbb{Z}_n^*$. This generator is used as the basis for building the accumulator value and must be securely chosen to work with modular exponentiation within $\mathbb{Z}_n^*$.

2. Accumulation Function: To compute the accumulator A for a set $S = \{x_1, x_2, \ldots, x_k\}$:

   - Representation of the set: Each element x_i in S must belong to $\mathbb{Z}$ (integers) and be coprime with n.
   - Accumulate the set: The accumulator A is calculated as:

     $$A = g^{\prod_{i=1}^{k} x_i} \mod n$$

     This involves taking the product of all elements (x_i) in S, raising g to the power of this product, and reducing modulo n. The result is the accumulator value A, which is a compact value that uniquely represents the entire set S. This value remains a fixed size, regardless of the number of elements in S.
   - Security Basis: The Strong RSA Assumption ensures that it is computationally infeasible to extract any individual element x_i from A without knowledge of the entire product $\prod_{i=1}^{k} x_i$.

3. Membership Proof: Membership verification allows proving that an element x belongs to the set S.

   - Witness Generation: For any $x \in S$, the witness w is computed as:

     $$w = g^{\prod_{y \in S, y \neq x} y} \mod n$$

     This witness is the accumulated value of all elements in S except x. The witness w is compact and can be efficiently computed and quickly verified.
   - Verification: Using the witness w, membership is verified by checking:

     $$A = w^x \mod n$$

     If this equation holds, x is confirmed as part of S.

4. Dynamic Updates: The dynamic accumulator supports efficient updates to the set S when elements are added or removed:

   - To add a new element x' to the set S:

     $$A' = A^{x'} \mod n$$

     The updated accumulator A' incorporates x' into the set.

   - To remove an element x' from S:

     $$A' = A^{1/x'} \mod n$$

     This operation assumes the existence of the modular inverse $x'^{-1} \mod n$, which requires x' to be coprime with n.

     Adding or removing elements requires only modular exponentiation, making updates computationally efficient.

Dynamic accumulators have addressed the limitations of static constructions, transforming the design of cryptographic systems. They serve as a cornerstone of modern cryptography, supporting practical applications in credential systems, blockchain, and identity management. With their efficiency, scalability, and strong security guarantees, dynamic accumulators have become an essential tool for systems that require secure and flexible membership verification.

4. METHODOLOGY

4.1. System Design

To design an efficient system that retrieves verification information with minimal overhead, we propose the Blockchain-Based Efficient Audit Trail System (BEATS). This system is a hybrid approach that combines on-chain and off-chain components to enhance its security and performance. BEATS is designed to meet the needs of two distinct user groups: the owner, who maintains document ownership, and the auditor, who verifies document integrity. Figure 4.1 provides an overview of BEATS.

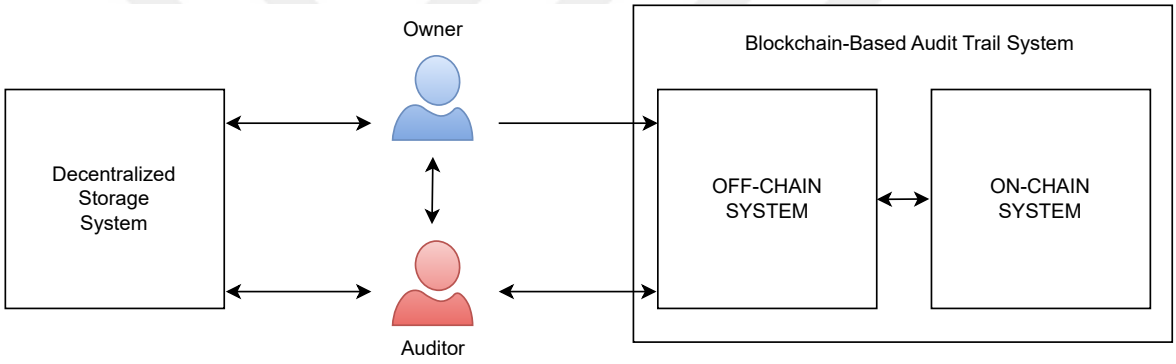


Figure 4.1. Overview of BEATS, illustrating the user-system interactions.

The owner, typically a company, securely stores its documents in a decentralized storage system (e.g. IPFS) and creates an account in the off-chain system to manage the metadata of these documents. The metadata is then recorded in the on-chain system, ensuring that the blockchain serves as a tamper-proof ledger for metadata while the document information remains securely off-chain. The blockchain can only be accessed through the off-chain system, which implements access control mechanisms to enforce proper permissions.

Auditors, on the other hand, interact with the system when verifying data integrity. They send a request to the company, which in turn provides the CIDs and grants access to its IPFS network to retrieve the relevant documents. Auditors can then independently utilize the blockchain-based verification system using the hash values, verifying the authenticity and integrity of the documents.

This dual-system architecture addresses critical challenges, including scalability, access control and computational efficiency. By separating storage and verification functionalities, the proposed system enables companies and auditors to efficiently prove document integrity while ensuring the transparency and security required for robust audit trails.

Building on this framework, the BEATS architecture can be explored in greater detail to illustrate its operational workflow and core components. As shown in Figure 4.2, the system enables secure and efficient document management by leveraging IPFS for decentralized storage and the blockchain for tamper-proof metadata recording.

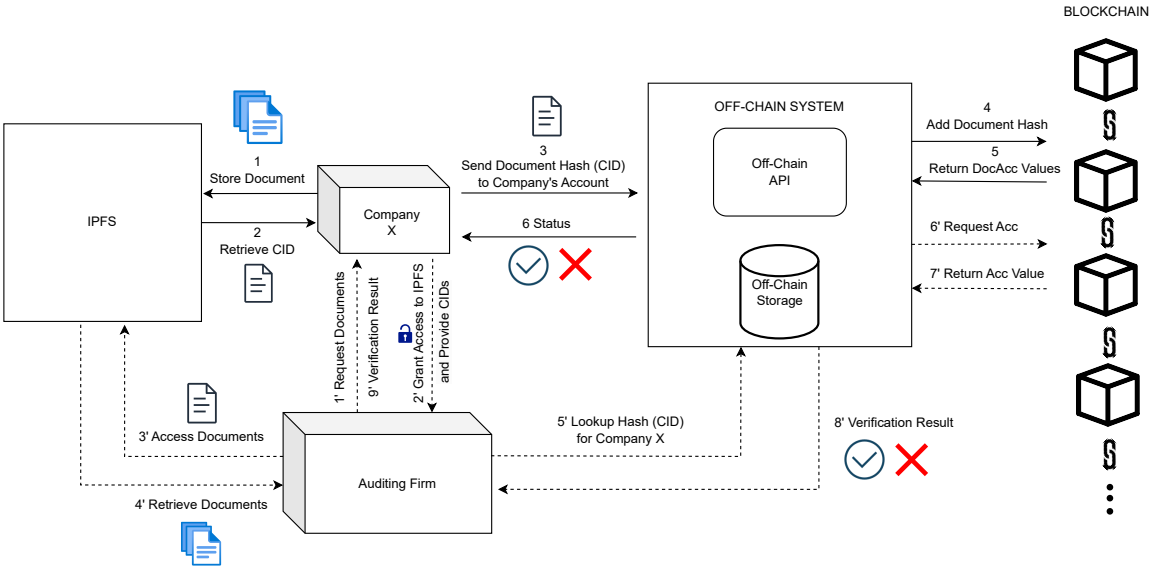


Figure 4.2. Detailed workflow of BEATS

The process begins with companies uploading their documents to IPFS while recording their corresponding hash values in the blockchain. This setup serves as an audit trail, enabling companies to maintain reliable records within a distributed system, and providing an efficient method for accessing and sharing documents for auditing purposes. Companies can upload and modify documents in IPFS but cannot directly interact with the blockchain, preserving the system's integrity. Any modifications to documents in IPFS will result in a different hash, ensuring that changes are detectable.

By integrating IPFS with blockchain, the system provides a robust solution for storing general or sensitive information, whether encrypted or not. This combination of decentralized environments ensures high data availability and supports reliable auditing

processes. However, relying solely on IPFS and blockchain introduces an auditing challenge: for auditors to verify the integrity of a document, they need to retrieve its hash value from the blockchain. This retrieval process typically involves searching through all transactions, as the hash value is stored as part of blockchain transactions. Auditors can address this challenge by either writing a smart contract to retrieve the desired document hash from a specific transaction or downloading and parsing the entire blockchain. However, these methods are time-intensive and could be vulnerable to misuse. For instance, an auditor with access to random transactions might retrieve a document without determining whether it has been tampered with.

To overcome this limitation, the system incorporates an RSA-based cryptographic accumulator that accumulates all hash values. Using this technology, auditors can verify the integrity of a document without going through individual transactions. This approach adds an extra layer of security and enhances the verification process.

With every document hash stored, the accumulator value is updated and recorded in the blockchain's global variable. When an auditor needs to verify the integrity of a document, they can use our verification system to verify its record (accumulator variables). This mechanism allows the auditor to confirm whether a document is listed in the system or has been tampered with. By leveraging the cryptographic accumulator, auditors can bypass transaction searches, significantly reducing verification time. This innovation enhances the system's efficiency and scalability, enabling more frequent and efficient integrity verification processes.

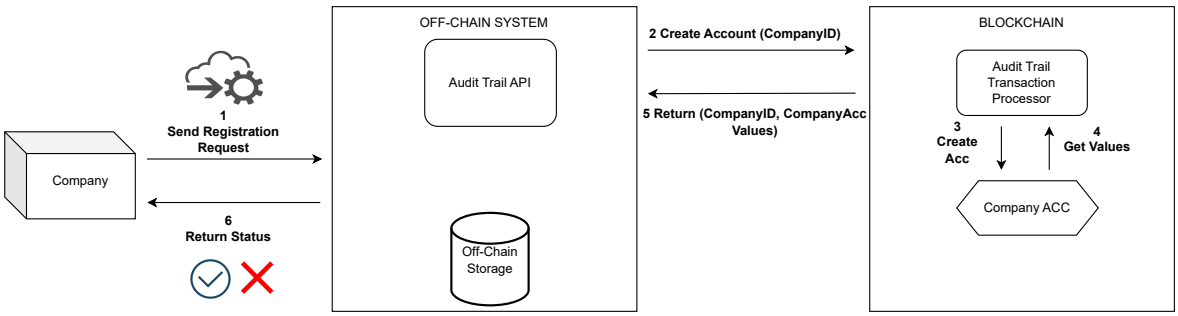


Figure 4.3. Registering in the system

A company can join the system by creating an account, a process outlined in Figure 4.3. It starts with the company submitting a registration request to the off-chain system. In response, the off-chain system sets up dedicated storage to manage its data.

Simultaneously, it sends a request to the blockchain to create a new account. This account is uniquely identified by an address generated specifically for the company using a combination of the company's public key and a custom Sawtooth transaction family.

The transaction processor is the key to this process. It generates an accumulator value and store it in the blockchain's global variable. This value, along with the document record, are then retrieved and sent back to the off-chain storage. Finally, the company is notified about the outcome of the account creation, confirming whether this process has been successfully completed. This ensures a seamless integration between the off-chain system and the blockchain, laying the foundation for secure and efficient data management.

For a company to add a document to its account, the company starts by sending its hash value to the off-chain system. The off-chain system then leverages the company's identifier to send this hash value to the blockchain. On the blockchain side, the transaction processor takes over, validating and executing the operation to incorporate the document hash into the company's accumulator, which is stored in the blockchain's global variable. Once the transaction is processed, the transaction processor retrieves the updated document record, including the prime representative, nonce, witness and the new accumulator value.

With these updated values successfully retrieved by the off-chain system, the system creates two tables and generates an updated accumulator version. The first table maps the retrieved document records to their corresponding document hash values for record-keeping, while the second table logs only the accumulator version and the prime representative to facilitate updating the document logs. This update process is crucial for witness updates, which play a vital role in the verification process.

Finally, the company is notified whether the document hash was successfully added or if any issues occurred during the process. This workflow is illustrated in Figure 4.4, offering a step-by-step understanding of how documents are securely added to the system.

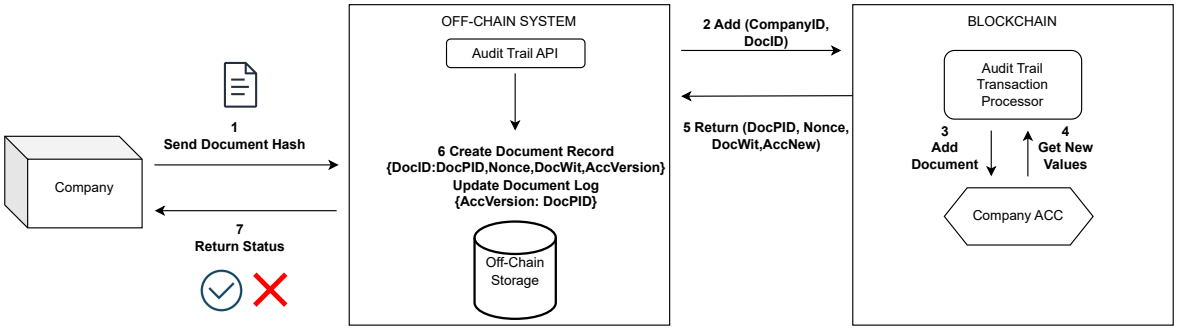


Figure 4.4. Adding documents to accounts

When a company decides to audit its documents, it engages an auditing firm for the task. Upon the auditor's request, the company grants access to its IPFS network and provides the CIDs, enabling the auditor to access relevant digital documents and retrieve them.

As shown in Figure 4.5, the verification process begins with the auditor referencing the document logs table to update the witness value, and subsequently retrieving the accumulator value from the blockchain's global variable. The auditor then utilizes our verification system using the document hash value, which maps the hash value to the corresponding document record stored in the off-chain system for verification purposes. Finally, the verification system leverages the document's prime representative and the updated witness value to prove whether the document is authentic, intact and valid, or if any tampering has occurred. This comprehensive procedure verifies the integrity and reliability of the audited documents, thereby reinforcing trust in the system.

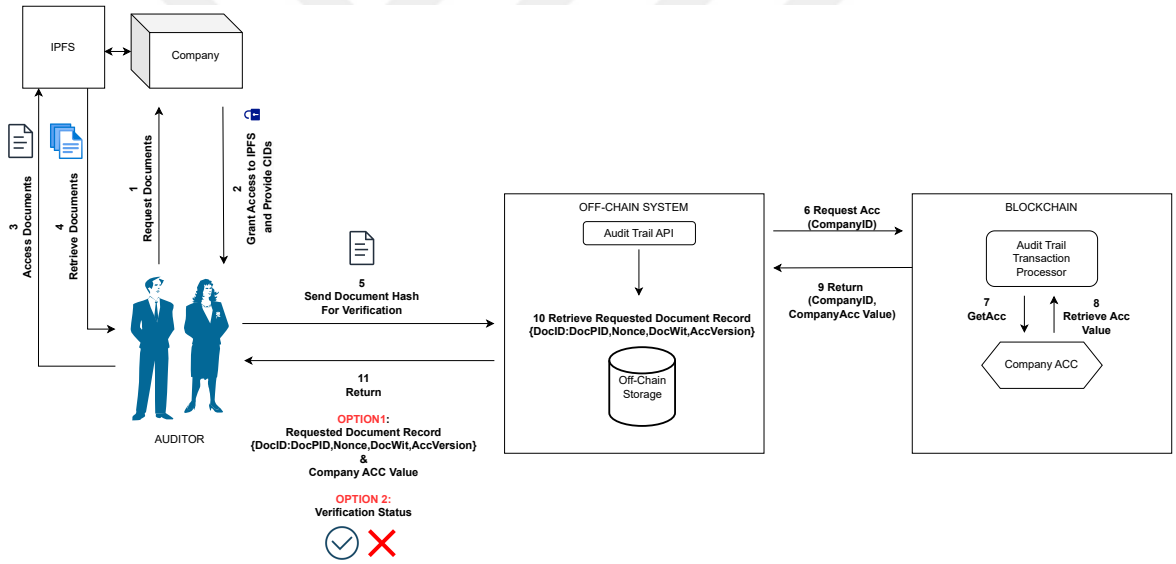


Figure 4.5. The audit process

4.2. Framework Development

For the development of the on-chain audit trail system, we utilized the Splinter Community's Sawtooth 1.2 (Chime). The system architecture is designed to include custom users who interact solely with the off-chain system. These users do not have direct access to or interaction with the blockchain itself. Instead, all communication with the on-chain system is managed exclusively by the off-chain system, which acts as an intermediary. This design simplifies user interaction and ensures that the blockchain remains a tamper-proof and immutable ledger for audit purposes.

We also developed a custom client API and transaction processor API to support system development and facilitate seamless interaction with the Sawtooth framework.

During our experiments, we employed the RSA-based cryptographic accumulator to retrieve verification information. This approach ensures the integrity and security of the audit trail process.

4.3. Evaluation Metrics

In this study, we focused on evaluating the efficiency of retrieving verification information from the blockchain for audit trail verification purposes. To assess the system's performance, we conducted experiments that involved submitting document hash values to our blockchain-based audit trail system. We simulated worst-case scenarios by querying for non-existent hash values to assess the system's behavior under such conditions, providing a comprehensive evaluation of verification efficiency.

Throughout the testing, each verification query was repeated multiple times to ensure statistical reliability. Average measurements were then calculated to facilitate precise evaluation and comparison. Key performance metrics were measured and recorded to evaluate system efficiency, focusing on CPU time spent on accumulator value retrieval and verification, as well as memory consumption associated with these processes.

5. RESULTS AND DISCUSSION

5.1. Experiments (Existing & Proposed)

This section outlines the experimental setup used to evaluate blockchain search algorithms, detailing the platform, experiment scenarios, and performance metrics. The testing environment operated on a computer equipped with an AMD Ryzen 3 processor and 6+ GB of RAM, running the Linux Ubuntu operating system. The blockchain framework used was Splinter Community's Sawtooth 1.2 (Chime), configured with a network of five nodes employing the PBFT consensus algorithm for existing approaches. The proposed system, on the other hand, uses Kubernetes (K8s) for a single Sawtooth node.

In existing transaction search approaches, the experiments involve generating and processing a large number of transactions. We conducted search queries for non-existent transactions to simulate worst-case scenarios, providing insight into the efficiency of the transaction search algorithms. We used a client container to submit transactions to the blockchain, and repeated each search query multiple times to ensure statistical reliability. We then calculated average measurements to ensure precise evaluation and comparison. Key performance metrics, including transaction retrieval time, CPU utilization, and memory consumption, were measured and recorded, with the results presented in the following section.

In contrast, our proposed system focuses on efficiently retrieving verification information from the blockchain to support audit trail verification. The experiments involve submitting document hash values to the audit trail system and retrieving their records to evaluate its performance. To simulate the worst-case, we queried non-existent hash values, assessing the system's behavior and verification efficiency under such scenarios. For comparative analysis, we evaluated the same performance metrics as those measured in the existing transaction search approaches.

Test results indicate that the existing transaction search approaches are inefficient; as the number of transactions in the blockchain increases, their performance gets worse over time. However, in the proposed system, communication complexity, computational complexity and memory usage remain constant, operating in O(1) time and space complexity. This efficiency is achieved through the use of a time- and space-efficient cryptographic accumulator. Regardless of the number of documents accumulated, the

accumulator size, witness size, and verification time remain unchanged. These properties demonstrate that the proposed system significantly outperforms the existing approaches in terms of scalability and efficiency.

5.2. Results (Existing & Proposed)

In this section, we present the results of our experiments, which also show the performance of each transaction search method. Test results show that our proposed system outperforms other transaction search approaches and proves the system's ability to facilitate efficient and reliable verification processes. Figure 5.1 presents the transaction download time required for the linear search. In this experiment, downloading the transaction list is essential for conducting the linear search, which adds additional overhead to the system. The figure also illustrates that the volume of transactions downloaded to the client node responsible for performing the linear search, scales linearly as the number of transactions increases. While real observations may vary due to network conditions and other factors, the consistent upward trend in transaction downloads is observable, and indicated by the dotted line.

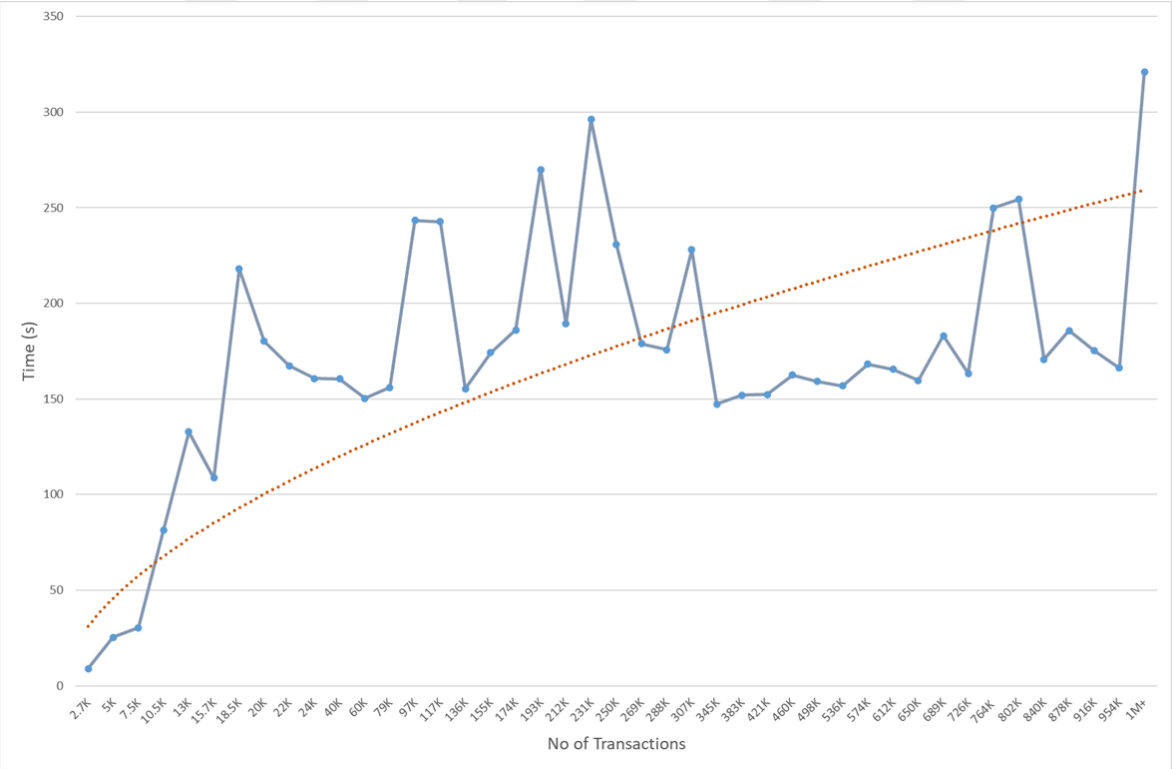


Figure 5.1. Average transaction download time required for linear search.

After downloading the transactions, the client node proceeds to perform a search. Figure 5.2 demonstrates a clear linear increase in search time corresponding to the number of transactions. The cumulative effect of adding the transaction download time to the search time results in a notable increase in total transaction retrieval time.

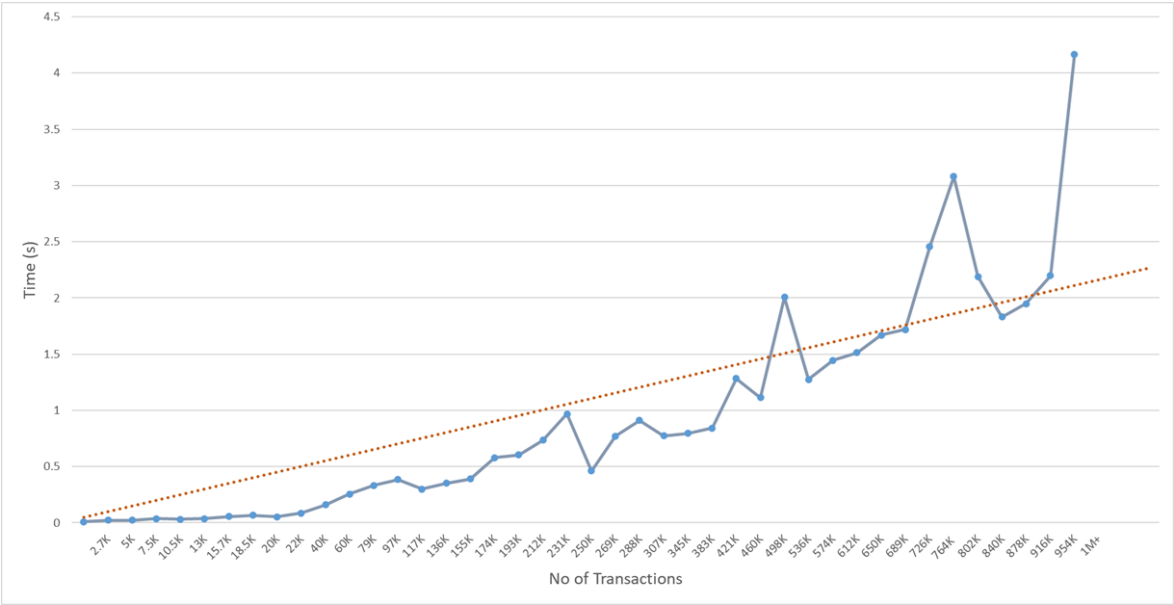


Figure 5.2. Average CPU time required for linear transaction search.

Sawtooth incorporates a transaction search solution. Figure 5.3 illustrates the retrieval time for this solution. Analysis of the average retrieval time reveals consistent variations throughout the experiments.

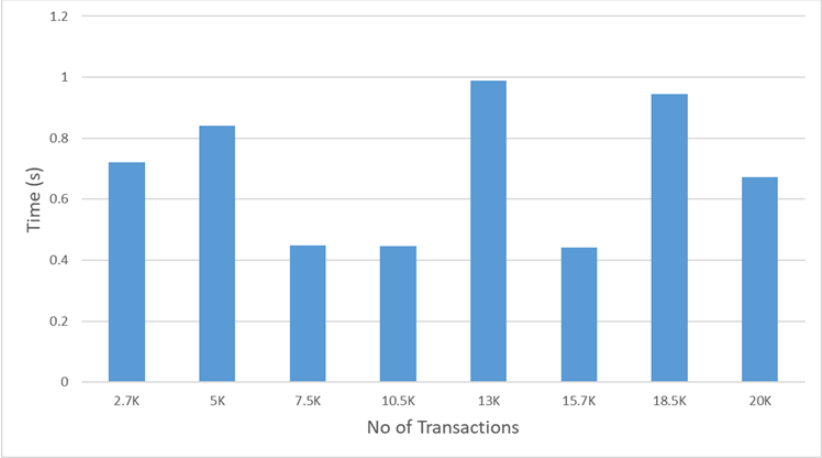


Figure 5.3. Average transaction retrieval time for platform-specific solutions.

Figure 5.4 displays the transaction retrieval time within a database-supported system. Upon analyzing the average time, a logarithmic increase becomes apparent as the number of transactions increases.

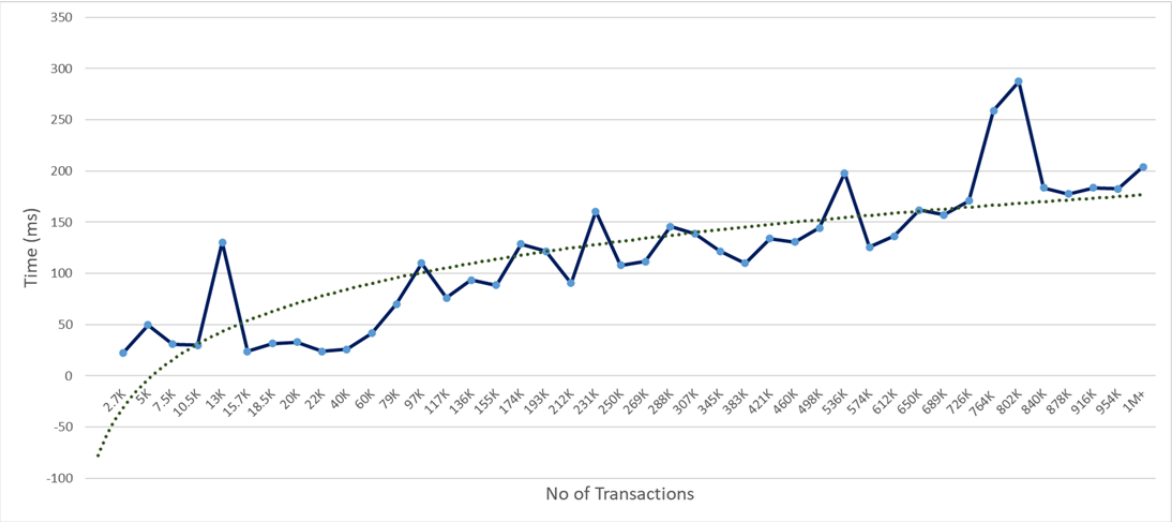


Figure 5.4. Average CPU time required for transaction retrieval in a database-supported blockchain system.

In comparison, Figure 5.5 illustrates the relationship between the number of documents and the time required for two essential processes in our proposed system: retrieving the accumulator value from the blockchain and performing verification. In contrast to the results of existing approaches, our proposed system demonstrates that both retrieving the accumulator value from the blockchain and performing verification exhibit an O(1) growth pattern, indicating that their execution time remains constant regardless of the increasing number of documents. Figure 5.6 focuses specifically on the verification process, highlighting minor fluctuations in its performance.

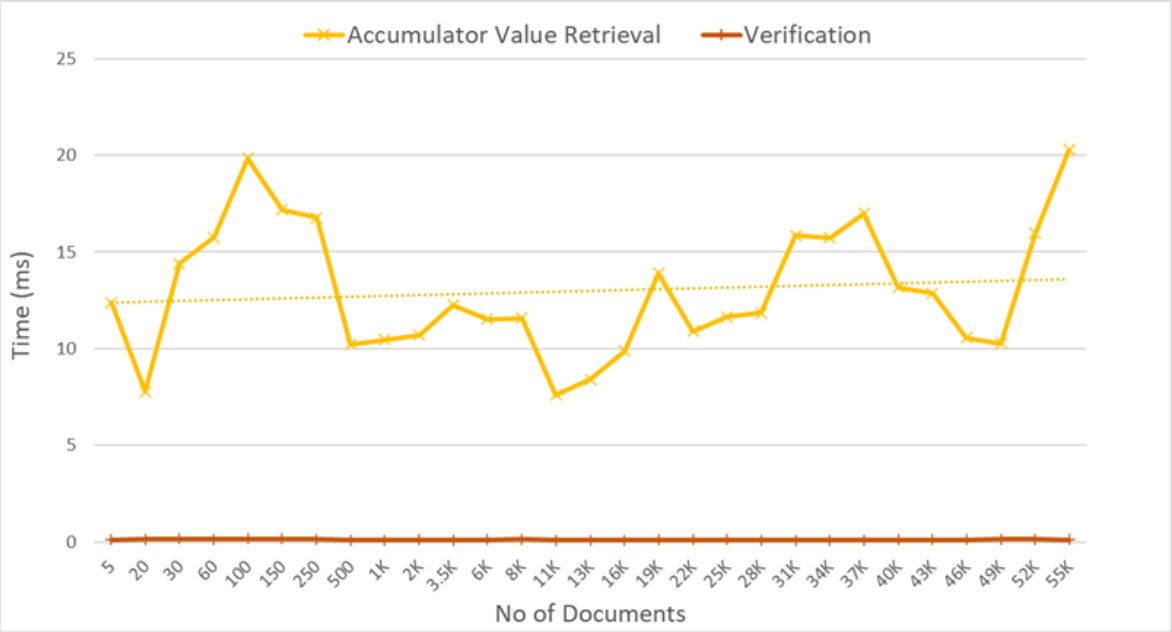


Figure 5.5. Average CPU time required for retrieving the accumulator value from the blockchain and performing verification.

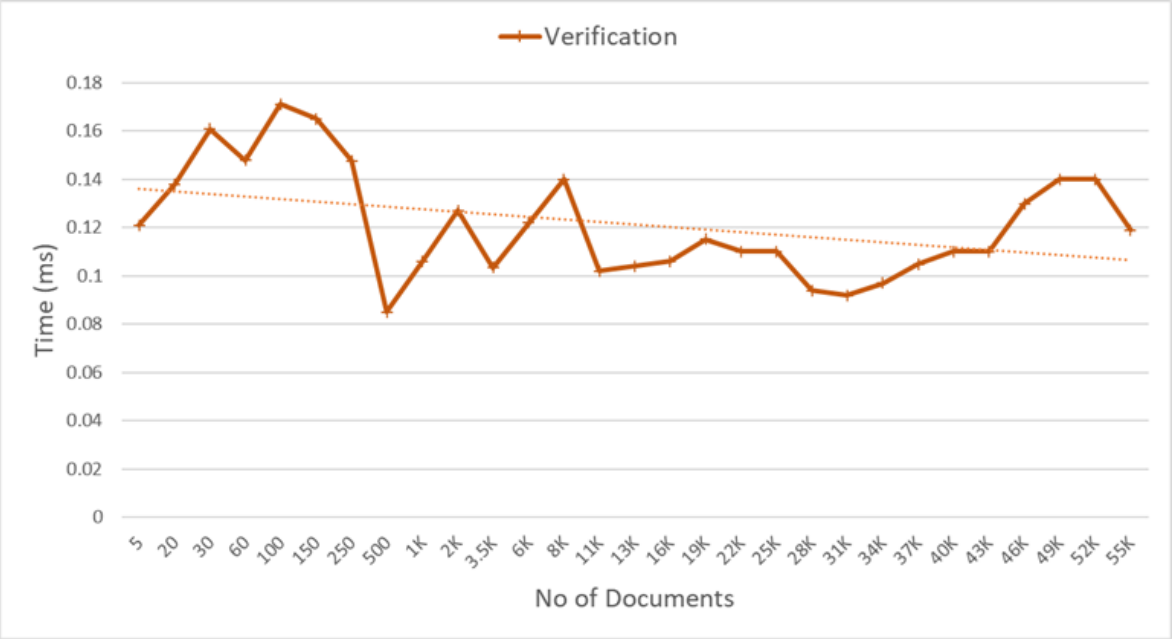


Figure 5.6. Average CPU time required for verification.

Figure 5.7 showcases the system's efficiency in memory utilization. The results emphasize minimal overhead, demonstrating consistent performance regardless of the number of documents involved.

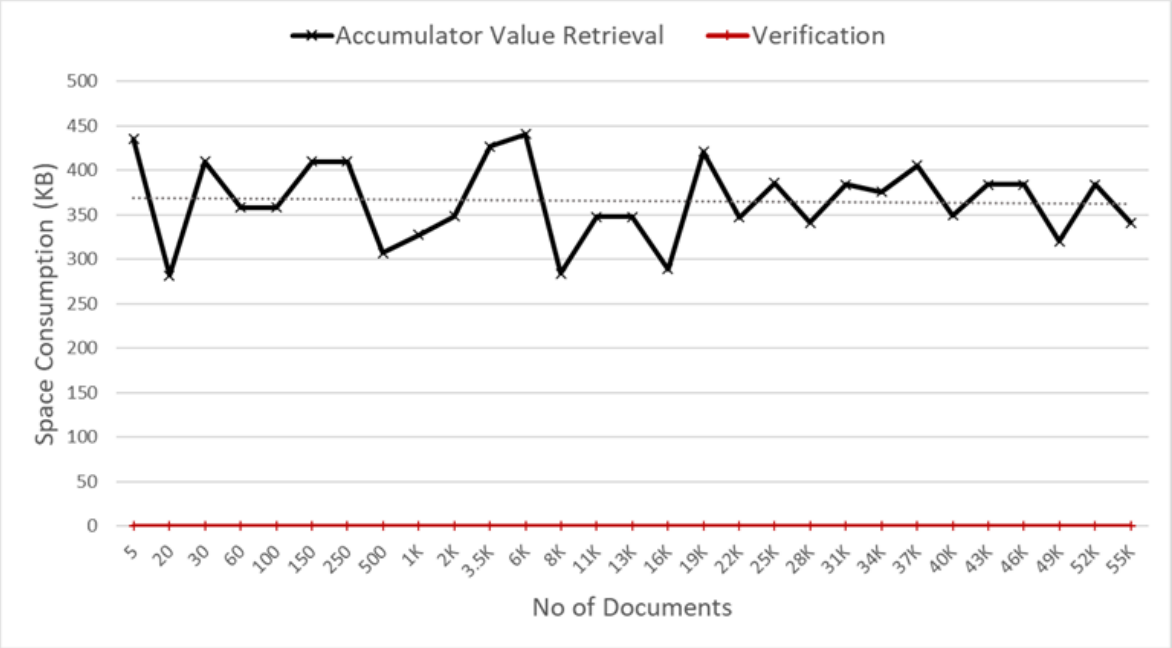


Figure 5.7. Average RAM consumption required for retrieving the accumulator value from the blockchain and performing verification.

5.3. Discussion

We thoroughly examined the performance metrics, limitations, and practical implications derived from our findings. As shown in Figure 5.2, there is a clear linear increase in transaction retrieval time as the number of transactions rises, starting at approximately 9.8 ms and continuing to scale proportionally. Notably, periodic spikes occur, particularly around every quarter million transactions. These spikes may be attributed to factors such as non-random data distribution or fluctuations in CPU cache utilization. Occasional peaks may also occur in comparison operations due to the absence of a cached required transaction during a search that can lead to delays when retrieving it from the memory.

During our experiments, we also monitored network conditions that affect transaction retrieval times. Downloading all transactions before conducting searches took a few minutes, as illustrated in Figure 5.1. The data retrieval over TCP sockets showed variability, reflecting the impact of the TCP protocol on download times.

To manage heavy memory workloads, traditioanl transaction search methods employed paging and swapping mechanisms. As memory became insufficient, swap space on the hard disk was utilized, effectively expanding memory capacity but also increasing input-output operations, thereby elevating CPU system time. Notably, our experiments indicated swap space activation after approximately 22K transactions.

The search methodology in Sawtooth had minimal impact on the client side, as all search operations were conducted within blockchain nodes, and the results were delivered to the client as a service. However, increased data downloads triggered a backoff system, initially accelerating data retrieval but causing slowdowns. Moreover, as transaction volumes exceeded 20K, network nodes encountered overwhelming requests, leading to sequential disconnections and network downtime. The results of this approach are depicted in Figure 5.3.

Figure 5.4 demonstrates that the database-based search outperforms other transaction search methods when handling substantial transaction volumes. However, database-based search operations heavily rely on disk access for data retrieval. If the underlying hardware experiences slow operation or becomes overloaded, it creates bottlenecks that significantly impair search performance.

As evidenced in Figures 5.2 and 5.4, the computational time for search algorithms escalates with transaction volume, highlighting potential scalability challenges. With transaction numbers approaching infinity, theoretical limitations in search time become apparent, emphasizing scalability concerns.

During our experiments, we also noticed that the linear search performance is worse than O(n), where 'n' is the total number of transactions, due to potential swap space usage and increased CPU system time demands. Hence, scalability considerations should encompass enhanced memory and disk performance when dealing with large transaction volumes.

On the other hand, our proposed system effectively addresses and outperforms existing challenges related to transaction search efficiency. As illustrated in Figure 5.5, the time required to retrieve the accumulator value from the blockchain remains relatively constant and minimal, regardless of the number of documents. The accumulator value retrieval time varies between 7ms and 20 ms, which can be attributed to factors such as network conditions or processing requirements. This demonstrates the system's ability to manage increasing data volumes without experiencing delays. This achieves consistent time

complexity, indicating that the accumulator value retrieval process is independent of the dataset size, ensuring consistent performance as the system scales.

The verification process, also depicted in Figure 5.5, demonstrates even greater time efficiency. Across all document counts, the verification time remains consistently close to 0 milliseconds, with negligible variability. Upon closer examination of Figure 5.6, the process takes less than 0.5 milliseconds to complete, showcasing near-instantaneous performance. This remarkable speed underscores the system's ability to verify data integrity on the blockchain with minimal latency. The observed O(1) time complexity for verification further emphasizes its scalability, as the execution time remains unaffected by the number of documents. This feature is particularly beneficial in blockchain environments requiring frequent and real-time verifications, ensuring high-speed operations without compromising accuracy.

Turning our attention to space utilization, Figure 5.7 highlights the memory consumption associated with retrieving the accumulator value from the blockchain. The memory usage remains within a bounded range, typically fluctuating between 280 KB and 450 KB. These fluctuations can be attributed to factors such as network latency or synchronization issues, which may temporarily increase memory usage during data retrieval operations. Additionally, the use of specific data structures, such as the cryptographic accumulator employed in this system, introduces some memory overhead during the retrieval process. Despite these fluctuations, memory consumption does not exhibit a significant increase with the growing number of documents, demonstrating the efficiency of the accumulator value retrieval process in terms of memory utilization. This illustrates constant space complexity, indicating that the accumulator value retrieval process is highly efficient and independent of dataset size. This performance proves that memory requirements remain within practical limits, even in large-scale blockchain-based audit trail applications, making the system both scalable and resource-efficient.

The verification process, also illustrated in Figure 5.7, exhibits exceptional consistency in RAM consumption. Although the graph might initially suggest zero memory usage, closer examination reveals a small, constant amount of memory being utilized. This consumption remains unchanged as the number of documents increases, highlighting that the accumulator's size does not grow with additional documents. Specifically, the space required for the accumulator in the verification process is fixed at around 256 bits, which is the size of the accumulator. Unlike the space required for the accumulator value retrieval, the verification process maintains a static RAM consumption, forming a horizontal line across the graph. This behavior aligns with an O(1) space complexity, emphasizing the system's efficiency in managing large datasets.

Overall, both accumulator value retrieval from blockchain and verification processes demonstrate O(1) time and space complexity, indicating high efficiency and scalability even as the number of documents increases. The system's ability to maintain low and stable resource requirements addresses the limitations of traditional blockchain transaction search methods, where scalability is constrained, and transaction retrieval times often increase with data size. These findings position our approach as an effective and scalable solution for blockchain-based audit trail systems, offering reliable performance for real-world applications that demand speed, scalability, and resource efficiency.

In this thesis, we first explored three traditional methods for transaction retrieval from the blockchain. Following this, we implemented our proposed approach, which focuses on efficiently finding the verification information. Test results demonstrate that our system significantly outperforms existing transaction search methods in both search time and space consumption, highlighting its potential as a scalable and efficient solution for blockchain-based audit trail applications.

5.3.1. Comparison with Existing Systems

The advantages of our proposed system become increasingly apparent as the number of documents to be audited grows. Unlike existing systems, which performance gets worse with an increase in transaction size, our system maintains unwavering efficiency due to its constant complexity for communication, computation and space utilization. In comparison, existing systems operate with best-case complexities approximately between O(n) and O(log n), which performance significantly diminishes as the number of transactions in the system increases. This difference underscores the scalability and robustness of our approach. Regardless of the volume of documents, the communication time, space requirements, and verification processes remain constant. This performance makes our system exceptionally suitable for large-scale applications, where efficient auditing of extensive datasets is crucial.

The experimental results validate these claims, showing that our proposed system outperforms existing systems in terms of efficiency and scalability. By utilizing the RSA-based cryptographic accumulator, the proposed system ensures that the accumulator size and verification time remain unchanged, even as the number of documents grows. This highlights the superiority of our system, offering a reliable, scalable and efficient solution for audit trail verification compared to existing methods.

5.3.2. Implications for Engineering Applications

Cryptographic accumulators are powerful tools with the potential to enhance security, privacy and efficiency across diverse applications and systems. In blockchain and distributed ledger technologies, they enable membership proof in a set of transactions or blocks, as in (Sun et al., 2023) study, improving the scalability and efficiency of blockchain operations. Additionally, cryptographic accumulators support privacy-preserving smart contracts, enhancing the security and confidentiality of blockchain-based systems.

Beyond blockchain, cryptographic accumulators excel in secure and privacy-preserving data aggregation and analysis. For example, in IoT networks and applications, as examined by (Huang and Shen, 2024), devices can contribute data without revealing sensitive information, facilitating efficient and secure data analysis. In identity management (Sun et al., 2023), accumulators allows multiple wallet accounts to be aggregated into a single proof, ensuring that the user's identity verification can be proven without revealing sensitive details such as wallet balances or transaction records. Cryptographic accumulators can also be used for integrity verification, as in (Y. Zhang et al., 2024), where it is utilized to protect encryption keys from unauthorized access or corruption, providing high-security assurance.

Cryptographic accumulators are valuable for secure authentication and access control. They enable device authentication in networks, establish secure communication channels, and restrict access to authorized entities only. This maintains security and integrity in sensitive environments. For example, (W. Jiang et al., 2024) utilized a cryptographic accumulator to enhance communication security by enabling message traceability, synchronization, and batch signature authentication. This cryptographic accumulator based system outperform traditional methods like IPsec in both security and efficiency, reducing computational overhead significantly. Cryptographic accumulators ensure data integrity, non-repudiation, and secure transmission while supporting scalable and resource-efficient operations.

Another notable application lies in privacy-preserving data sharing across various domains. In healthcare, cryptographic accumulators enable researchers to analyze medical data without compromising patient privacy. Similarly, in finance, they allow for secure analysis of financial data while protecting individual transaction details.

In summary, cryptographic accumulators can address key challenges in modern engineering systems. Their ability to enhance security, preserve privacy, and optimize

efficiency makes them essential in advancing technology. As practical solutions, they can shape the future of secure and efficient systems across industries.

6. CONCLUSION AND RECOMMENDATIONS

Blockchain technology is set to revolutionize a range of industries, including audit trails. As blockchain networks evolve, they will become more capable of handling a large number of transactions and data without affecting performance. Potential advances in scalability and privacy are subjected to enhance blockchain's capabilities, leading to more comprehensive and interconnected audit trail systems. This will enable organizations to maintain detailed audit trails while protecting sensitive information. This will be valuable in industries where data privacy is a critical concern, such as finance and healthcare. The integration of artificial intelligence (AI) (Vani et al., 2024), machine learning (ML) (Adelakun et al., 2024) and smart contracts is expected to enhance the functionality of audit trails. AI-powered analytics can identify anomalies and predict potential fraud, while smart contracts can allow for more complex and automated audit processes. Additionally, the integration of blockchain with IoT devices will create new opportunities for real-time audit trails (Ratnayake et al., 2024). IoT devices generate vast amounts of data that can be securely recorded and verified on a blockchain. This integration will provide immutable audit trails for processes and transactions involving IoT devices, enhancing transparency and accountability in industries including manufacturing and supply chain management.

Blockchain's impact on governance structures is significant. Decentralized governance models, increased transparency, and compliance processes are among the potential benefits. Blockchain technology has the potential to enhance compliance processes by automating and recording compliance activities. With smart contracts, organizations can automate the enforcement of regulatory requirements. This will reduce the administrative tasks associated with compliance and minimize human error, leading to more accurate reporting. Audit trails on a blockchain will be updated in real-time, providing auditors and regulators with immediate access to accurate and verified data, which will enable more proactive audits. Blockchain's global nature will facilitate cross-border compliance by providing a unified and transparent system for verifying international transactions. This will simplify compliance with global regulations and standards, making it easier for multinational organizations to manage and report on their activities across different countries. By automating compliance processes and reducing the need for intermediaries, blockchain technology has the potential to significantly lower compliance costs. Organizations will benefit from reduced administrative overhead and fewer resources dedicated to managing compliance, allowing them to allocate resources more efficiently. The future of blockchain holds tremendous potential for transforming

audit trails and governance practices. These innovations will enhance the effectiveness of audit trails and transform governance structures and compliance protocols by increasing transparency, accountability and privacy, all while mitigating risks.

An audit trail serves as a comprehensive record of activities within a system, essential for tracking changes, particularly in accounting and financial systems. Audit failures have underscored the pressing need for more reliable audit mechanisms. Traditional auditing methods reveal significant vulnerabilities, relying on outdated systems that are prone to security breaches and unauthorized access, which can compromise audit integrity. Additionally, these methods are often time-consuming, expensive, and susceptible to delays and errors due to manual processes, with the lack of real-time verification increasing the risk of undetected fraud. The significance of maintaining a trustworthy audit trail has been emphasized by numerous instances of fraud and data mismanagement in recent years. The scams and revenue loss that businesses encounter in recent years have turned the world's eyes towards the importance of having a robust audit trail system that can detect and prevent such incidents.

Blockchain addresses many of the shortcomings of traditional audit methods by providing a tamper-proof, real-time ledger of transactions that is accessible to all parties involved. This technological advancement offers enhanced transparency, reduced fraud risk, and increased efficiency in auditing processes. By automating transaction recording and eliminating many manual procedures, blockchain can significantly lower the time and costs associated with traditional auditing practices.

The growing significance of blockchain technology in the field of audit trails is evident from many surveys conducted. Blockchain technology offers important features such as immutability, transparency and availability, making it exceptionally well-suited for audit trails. Through the implementation of blockchain-based audit trail, organizations and governmental entities can enhance their audits by taking advantage of the features offered by the blockchain technology, which enables all stakeholders track and verify audit results without third-party dependency.

Implementing blockchain for audit trails enables organizations to create immutable and transparent records of all system activities, including data changes and access attempts. Smart contracts can be employed to set specific conditions for recording events on the blockchain, ensuring accurate documentation and reducing the likelihood of fraudulent activities. Unlike traditional centralized systems, such as databases, which often fall short in providing comprehensive and tamper-proof records, blockchain-based systems introduce essential features of decentralization and immutability. Centralized systems rely on a single

point of control, making them vulnerable to unauthorized access, data manipulation, or even catastrophic failures. In contrast, blockchain operates on a decentralized network where data is distributed across multiple nodes, ensuring no single entity has complete control. This decentralization not only enhances security but also ensures transparency, as all participants in the network can verify the authenticity of the records.

Existing studies often overlook the challenge of efficiently retrieving specific information from the blockchain, which is a critical requirement for practical audit trail systems. As the volume of transactions on the blockchain grows, locating and verifying specific pieces of information becomes increasingly time-consuming and resource-intensive. This is because traditional transaction search approaches require searching through large amounts of data in a linear or indexed manner, which is not scalable for real-world applications where quick access to specific records is essential.

This thesis addresses this significant limitation by proposing a novel blockchain-based efficient audit trail system. The proposed system introduces an innovative solution to optimize data verification using a cryptographic accumulator, a time- and space-efficient data structure that enables efficient data verification without the need to process the entire blockchain. The cryptographic accumulator is implemented on the blockchain, allowing for reliable and efficient verification. To further enhance the system's performance, the blockchain is divided into two parts: on-chain and off-chain systems. The generation of cryptographic accumulator values, including verification proofs, is implemented on-chain to ensure security. These values are then stored off-chain alongside the hashes of the documents. This hybrid approach reduces the load on the blockchain network, optimizing storage efficiency and improving overall system performance while preserving the integrity and verifiability of the data.

The proposed system significantly improves the speed and efficiency of audit trail verification processes. It surpasses the performance of state-of-the-art approaches, offering a scalable and practical solution for real-world applications. This study lays the foundation for future research and development, enabling the creation of more practical, efficient, and scalable solutions with broad applicability across industries.

6.1. Summary of Findings

Researchers explore various application areas to utilize blockchain in audit trails to take advantage of transparency through decentralized and immutable transaction recording. Blockchain technology ensures data integrity and fosters trust in the auditing process.

Furthermore, by eliminating intermediaries and minimizing manual tasks, blockchain reduces costs associated with auditing.

This research evaluates the available approaches of finding transactions within blockchain systems. Initially, a linear search approach to search for a specific transaction by its unique transaction ID was applied. This approach entails downloading all transactions from the blockchain and then sequentially scanning transactions starting from the genesis block until the target transaction is found, resulting in a time complexity of O(n).

The second approach involves using platform specific solutions to retrieve transactions from the blockchain. This approach offers improved performance compared to linear transaction search as it delegates the task to the blockchain node, which executes it on behalf of the client. However, as the workload on the nodes increases with the number of requests, it can lead to connectivity issues and eventual loss of connection. The cost of this approach looks like O(1), but this is not accurate. In practice, the search tests were executed on the miners rather than on the client, and thus the results observed on the client node merely reflect the performance of the miners.

Another approach involves using a database-supported system, where transactions are stored in the blockchain while their identifiers are saved in the database. By utilizing tree data structures such as B tree, the search complexity is optimized to c*log(n), where c is a constant coefficient, significantly more efficient than the previous approaches. However, this approach needs to be running in parallel with the blockchain, which requires additional space and computational power compared to the linear transaction search and platform-specific solutions approaches. Additionally, implementing this approach requires verifying the integrity, and addressing this issue adds additional complexity to the system.

Many blockchain platforms incorporate a Merkle hash in each block to facilitate efficient and secure transaction verification. Since the complexity of Merkle tree search is log(n), traversing each block and verifying its hash incurs a theoretical complexity of m*log(k), where ’m’ is the number of blocks, and ’k’ represents the average transactions per block. This approach outperforms linear transaction search, and it does not have the integrity verification concerns that the database-based search has. This approach also exhibits exceptional performance particularly when dealing with smaller or fewer blocks.

Unlike these approaches, where performance degrades as transaction size increases, our proposed system maintains consistent efficiency. Test results show that it operates with O(1) time and space complexity, ensuring that both communication time, computation time and space utilization remain constant regardless of the transaction size. This unique

characteristic allows our system to scale effectively, handling a larger number of documents without compromising performance, making it highly efficient and reliable even as the volume of documents grows.

In comparison, existing systems demonstrate that the database-supported approach provide the most efficient performance and scalability. On the other hand, platform-specific solutions, followed by linear transaction search, are the least efficient ones.

Our proposed system overcomes the limitations of these existing solutions, addressing the speed, communication, and scalability issues. It offers a faster and more effective approach, ensuring improved performance across a range of scenarios.

6.2. Limitations

A key limitation of this study lies in its focus on the existence of documents rather than providing a mechanism for their specific retrieval from the blockchain. While the system verifies adding document information to the blockchain, it does not inherently offer functionalities for companies to locate and access them directly. This implies that companies may encounter difficulties retrieving documents for investigation or other purposes.

Another limitation arises from the off-chain system design. As the number of participating companies increases, the system could become susceptible to SPOF and security vulnerabilities, potentially exploited by hackers.

By addressing these challenges, the study can provide a more comprehensive understanding of the practical implications and potential future enhancements of the proposed blockchain-based audit trail framework. This necessitates the development of robust search and retrieval mechanisms, potentially involving indexing techniques or advanced query languages, to enable users to efficiently locate and access specific document information within the blockchain.

6.3. Future Research Directions

The application of blockchain-based audit trail systems remains a field that requires further exploration. Future research could focus on developing efficient and secure retrieval mechanisms for blockchain-based audit trails. This includes evaluating and implementing advanced indexing techniques, such as Bloom filters or Merkle trees, to locate and retrieve

specific documents efficiently. Researchers could also explore specialized methods enabling companies to search and retrieve information effectively based on various criteria, such as timestamps or document types. Additionally, investigating decentralized search protocols could provide solutions for secure and efficient information retrieval without relying on centralized systems, mitigating information security risks.

These research questions underscore the opportunities to enhance blockchain-based audit trail systems, offering promising directions for future exploration. Addressing these challenges could unlock the full potential of blockchain in audit applications, paving the way for innovative advancements in secure and efficient audit practices.

This thesis provides a comprehensive exploration of blockchain and cryptographic accumulator's role in audit trail systems. By examining recent literature and technological frameworks, it underscores how cryptographic accumulators can enhance audit practices through improved security and efficiency. This study identifies current challenges and emphasizes the need for continued research to address its limitations. Finally, this work contributes to a deeper understanding of blockchain and cryptographic accumulators' potential in auditing and serves as a foundation for future innovations in this dynamic and impactful field.

REFERENCES

ACFE. (2022). Press release: 2022 rttn launch [Accessed: 2024-09-27]. https://www.acfe.com/about-the-acfe/newsroom-for-media/press-releases/press-release-detail?s=2022-RTTN-launch

Adelakun, B. O., Fatogun, D. T., Majekodunmi, T. G., & Adediran, G. A. (2024). Integrating machine learning algorithms into audit processes: Benefits and challenges. *Finance & Accounting Research Journal*, *6*(6), 1000–1016.

Adlam, R., & Haskins, B. (2020). A permissioned blockchain approach to electronic health record audit logs. *Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications*, 1–7.

Ahmad, A., Saad, M., Al Ghamdi, M., Nyang, D., & Mohaisen, D. (2021). Blocktrail: A service for secure and transparent blockchain-driven audit trails. *IEEE Systems Journal*, *16*(1), 1367–1378.

Ahmad, A., Saad, M., Njilla, L., Kamhoua, C., Bassiouni, M., & Mohaisen, A. (2019). Blocktrail: A scalable multichain solution for blockchain-based audit trails. *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, 1–6.

Akkaya, K., & Cebe, M. Efficient public-key revocation management for secure smart meter communications using one-way cryptographic accumulators. In: Florida International Univ., Miami, FL (United States). 2018. https://www.osti.gov/biblio/1485600

Ali, A., Pasha, M. F., Ali, J., Fang, O. H., Masud, M., Jurcut, A. D., & Alzain, M. A. (2022). Deep learning based homomorphic secure search-able encryption for keyword search in blockchain healthcare system: A novel approach to cryptography. *Sensors*, *22*(2), 528.

Aßländer, M. S., & Burkatzki, E. (2023). The financial fraud of the german fintech company wirecard: Structural causes and failures of the supervisory authorities. In *Sustainable finance and financial crime* (pp. 291–306). Springer.

Baranidharan, S., Ajekwe, C. C. M., & Nakitende, M. G. (2023). Accounting fraud and bankruptcy: The case of wirecard ag. In *Theory and practice of illegitimate finance* (pp. 222–244). IGI Global.

Benet, J. (2014). Ipfs-content addressed, versioned, p2p file system. *arXiv preprint arXiv:1407.3561*.

BitFury Group. (2016). On blockchain auditability.

Burghart, F., & Wagner, S. (2024). A bijection for the evolution of *B*-trees. *arXiv preprint arXiv:2406.06359*.

Cai, C., Weng, J., Yuan, X., & Wang, C. (2018). Enabling reliable keyword search in encrypted decentralized storage with fairness. *IEEE Transactions on Dependable and Secure Computing*, *18*(1), 131–144.

Camenisch, J., & Lysyanskaya, A. (2002). Dynamic accumulators and application to efficient revocation of anonymous credentials. In M. Yung (Ed.), *Advances in cryptology — crypto 2002* (pp. 61–76). Springer Berlin Heidelberg.

Capocasale, V., Gotta, D., & Perboli, G. (2023). Comparative analysis of permissioned blockchain frameworks for industrial applications. *Blockchain: Research and Applications*, *4*(1), 100113.

CCPA. (2024). California consumer privacy act of 2018 contents [Accessed: November 30, 2024; CALIFORNIA CONSUMER PRIVACY ACT OF 2018 - effective 7/15/2024 - AB 3286 update, posted to cppa.ca.gov August 2024]. https://cppa.ca.gov/regulations/pdf/cppa_act.pdf

Chen, E., Liang, J., Huang, R., Hung, P., Chen, D., Hsu, A., Chalkias, K., & Pleros, S. (2023). Building random, fair, and verifiable games on blockchain. raffle smart contract designs on sui network. *arXiv preprint arXiv:2310.12305*.

Coinmagazine. (2023). Propy: The blockchain-powered real estate revolution [Accessed: November 30, 2024]. https://coinmagazine.medium.com/propy-the-blockchain-powered-real-estate-revolution-7698c47d687c

Confais, B., Parrein, B., Lacan, J., & Marques, F. (2023). Characterization of the ipfs public network from dht requests. In A. Hameurlain & A. M. Tjoa (Eds.), *Transactions on*

large-scale data- and knowledge-centered systems lv (pp. 87–108). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-68100-8_4

Constantinides, T., & Cartlidge, J. (2021). Block auction: A general blockchain protocol for privacy-preserving and verifiable periodic double auctions. *2021 IEEE International Conference on Blockchain (Blockchain)*, 513–520.

Contributors to Sawtooth. (2024). Sawtooth [Copyright 2018-2024 contributors to Sawtooth. Accessed: 2024-12-09]. https://sawtooth.splinter.dev/

Deloitte. (2019). 2019 global blockchain survey [Deloitte Insights]. https://www2.deloitte.com/us/en/insights/topics/understanding-blockchain-potential/global-blockchain-survey-2019.html

Derler, D., Hanser, C., & Slamanig, D. (2015). Revisiting cryptographic accumulators, additional properties and relations to other primitives. In K. Nyberg (Ed.), *Topics in cryptology — ct-rsa 2015* (pp. 127–144). Springer International Publishing.

Doekhi, R. J. (2023). The intercompany settlement blockchain: Benefits, risks, and internal it-controls [Open Access under the terms of the Creative Commons Attribution 4.0 International License.]. *Advanced Digital Auditing*, *47*. https://doi.org/10.1007/978-3-031-11089-4

Edwards, D., et al. (2004). Personal information protection and electronic documents act.

Emmadi, N., Vigneswaran, R., Kanchanapalli, S., Maddali, L., & Narumanchi, H. (2019). Practical deployability of permissioned blockchains. *Business Information Systems Workshops: BIS 2018 International Workshops, Berlin, Germany, July 18–20, 2018, Revised Papers 21*, 229–243.

Ermolaev, V., Klangberg, I., Madhwal, Y., Vapper, S., Wels, S., & Yanovich, Y. (2020). Incorruptible auditing: Blockchain-powered graph database management. *2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 1–3.

Garip, S., Bilgen, M., Altin, N., Ozdemir, S., & Sefa, I. (2023). Reliability analysis of microgrids: Evaluation of centralized and decentralized control approaches. *Electric Power Components and Systems*, *51*(19), 2319–2338.

GDPR. (2019). Gdpr. 2019 [Regulation, General Data Protection].

Guegan, D. (2017). Public blockchain versus private blockhain.

Handoko, B. L., Arfianti, F., & Marlinda, S. (2022). The utilization of blockchain technology on remote audit to ensure audit data integrity in detecting potential fraudulent financial reporting. *Proceedings of the 2022 6th International Conference on Software and e-Business*, 104–112.

Hu, S., Cai, C., Wang, Q., Wang, C., Luo, X., & Ren, K. (2018). Searching an encrypted cloud meets blockchain: A decentralized, reliable and fair realization. *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, 792–800.

Hu, S., Cai, C., Wang, Q., Wang, C., Wang, Z., & Ye, D. (2019). Augmenting encrypted search: A decentralized service realization with enforced execution. *IEEE Transactions on Dependable and Secure Computing*, *18*(6), 2569–2581.

Hu, V. C., Ferraiolo, D., Kuhn, D. R., et al. (2006). *Assessment of access control systems*. US Department of Commerce, National Institute of Standards; Technology ...

Huang, C., & Shen, X. ((2024). Privacy-preserving cross-domain authentication in internet of things. In *Decentralized privacy preservation in smart cities* (pp. 61–90). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54075-2_4

Jayashree, D., Pandithurai, O., Yogeswari, S., Swetha, K., & Shalini, A. (2021). I-voting system based on block chain. In *Smart intelligent computing and communication technology* (pp. 81–86). IOS Press.

Jiang, S., Duan, Y., & Wu, J. (2019). A client-biased cooperative search scheme in blockchain-based data markets. *2019 28th International Conference on Computer Communication and Networks (ICCCN)*, 1–9.

Jiang, W., Wang, Y., & Ye, S. (2024). A memorable communication method based on cryptographic accumulator. *Electronics*, *13*(6), 1081. https://doi.org/10.3390/electronics13061081

Kalis, R. (2018). Using blockchain to validate audit trail data in private business applications. *University of Amsterdam*.

Khan, A. A., Shaikh, A. A., & Laghari, A. A. (2023). Iot with multimedia investigation: A secure process of digital forensics chain-of-custody using blockchain hyperledger sawtooth. *Arabian Journal for Science and Engineering*, *48*(8), 10173–10188.

Khedr, W. I., Khater, H. M., & Mohamed, E. R. (2019). Cryptographic accumulator-based scheme for critical data integrity verification in cloud storage. *IEEE Access*, *7*, 65635–65651. https://doi.org/10.1109/ACCESS.2019.2917628

Kshetri, N. (2018). Blockchain and electronic healthcare records [cybertrust]. *Computer*, *51*(12), 59–63.

Lai, C., & Wang, Y. (2022). Achieving efficient and secure query in blockchain-based traceability systems. *2022 19th Annual International Conference on Privacy, Security & Trust (PST)*, 1–5.

Li, Y., Cao, L., Zheng, G., Men, H., & Chen, L. (2024). Improved rsa dynamic cryptographic accumulator-based anonymous batch authentication scheme for internet of vehicles. *Computers and Electrical Engineering*, *117*, 109261. https://doi.org/https://doi.org/10.1016/j.compeleceng.2024.109261

Li, Y. (2023). Study on the reasons for the failure of the audit of luckin coffee and suggestions for countermeasures. *International Conference on Economic Management and Green Development*, 729–737.

Liu, X., Zhang, S., Huang, H., & Malekian, R. (2024). A trustworthy and reliable multi-keyword search in blockchain-assisted cloud-edge storage. *Peer-to-Peer Networking and Applications*, *17*(2), 985–1000.

Liu, Z., Ren, L., Feng, Y., Wang, S., & Wei, J. (2023). Data integrity audit scheme based on quad merkle tree and blockchain. *IEEE Access*, *11*, 59263–59273.

Makina, H., Letaifa, A. B., & Rachedi, A. (2022). Leveraging edge computing, blockchain and ipfs for addressing ehealth records challenges. *2022 15th International Conference on Security of Information and Networks (SIN)*, 01–04.

Malik, S., Kanhere, S. S., & Jurdak, R. (2018). Productchain: Scalable blockchain framework to support provenance in supply chains. *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, 1–10.

McCorry, P., Mehrnezhad, M., Toreini, E., Shahandashti, S. F., & Hao, F. (2021). On secure e-voting over blockchain. *Digital Threats: Research and Practice (DTRAP)*, *2*(4), 1–13.

Morishima, S., & Matsutani, H. (2018). Accelerating blockchain search of full nodes using gpus. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 244–248.

Muzammal, M., Qu, Q., & Nasrulin, B. (2019). Renovating blockchain with distributed databases: An open source system. *Future generation computer systems*, *90*, 105–117.

Newray, K. L. (2023). Improving public record management system at government ministries and agencies: Factors to consider and action steps to take [IJFMR230611060]. *International Journal for Multidisciplinary Research (IJFMR)*, *5*. https://www.ijfmr.com

Nie, X., Zhang, A., Chen, J., Qu, Y., & Yu, S. (2022). Time-enabled and verifiable secure search for blockchain-empowered electronic health record sharing in iot. *Security and Communication Networks*, *2022*(1), 1103863.

Nikolić, S., Zdravković, N., Franc, I., & Arivazhagan, N. (2023). A comparison on hyperledger consensus mechanism security and their applications.

Noviani, S. A., & Trisnaningsih, S. (2024). Effectiveness of the audit function as an internal control tool. *International Journal of Management Research and Economics*, *3*(2), 103–108. https://doi.org/10.54066/ijmre-itb.v2i3.210

Nyaletey, E., Parizi, R. M., Zhang, Q., & Choo, K.-K. R. (2019). Blockipfs-blockchain-enabled interplanetary file system for forensic and trusted data traceability. *2019 IEEE International Conference on Blockchain (Blockchain)*, 18–25.

Oakley, J., Worley, C., Yu, L., Brooks, R. R., Özçelik, İ., Skjellum, A., & Obeid, J. S. (2023). Scrybe: A secure audit trail for clinical trial data fusion. *Digital threats: research and practice*, *4*(2), 1–20.

Ozcelik, I., Medury, S., Broaddus, J., & Skjellum, A. (2021). An overview of cryptographic accumulators. *Proceedings of the 7th International Conference on Information Systems Security and Privacy*. https://doi.org/10.5220/0010337806610669

Ozcelik, I., & Skjellum, A. (2021). Cryptorevocate: A cryptographic accumulator based distributed certificate revocation list. *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, 0865–0872. https://doi.org/10.1109/CCWC51732.2021.9376112

Peng, Y., Du, M., Li, F., Cheng, R., & Song, D. (2020). Falcondb: Blockchain-based collaborative database. *Proceedings of the 2020 ACM SIGMOD international conference on management of data*, 637–652.

Peng, Z., Wu, H., Xiao, B., & Guo, S. (2019). Vql: Providing query efficiency and data authenticity in blockchain systems. *2019 IEEE 35th international conference on data engineering workshops (ICDEW)*, 1–6.

Peter, N., & Vukovljak, B. (2023). The impact of blockchain on audit profession. *MAP Social Sciences*, *3*(1), 11–24. https://doi.org/10.53880/2744-2454.2023.3.1.11

Pramukti, A. (2024). Internal audit versus external audit: A qualitative perspective. *Golden Ratio of Auditing Research*, *4*(2), 78–88. https://doi.org/10.52970/grar.v4i2.372

Pratama, F. A., & Mutijarsa, K. (2018). Query support for data processing and analysis on ethereum blockchain. *2018 International Symposium on Electronics and Smart Devices (ISESD)*, 1–5.

Rahman, M. S., Chamikara, M., Khalil, I., & Bouras, A. (2022). Blockchain-of-blockchains: An interoperable blockchain platform for ensuring iot data integrity in smart city. *Journal of Industrial Information Integration*, *30*, 100408.

Ratnayake, R., Liyanage, M., & Murphy, L. (2024). Machine learning for data trust evaluations in blockchain-enabled iot systems. *2024 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 1–2.

Regueiro, C., Seco, I., Gutiérrez-Agüero, I., Urquizu, B., & Mansell, J. (2021). A blockchain-based audit trail mechanism: Design and implementation. *Algorithms*, *14*(12), 341.

Ringelstein, C., & Staab, S. (2009). Dialog: Distributed auditing logs. *2009 IEEE International Conference on Web Services*, 429–436.

Rodeck, D., & Curry, B. (2022). What is blockchain. *Forbes*.

Rosa-Bilbao, J., & Boubeta-Puig, J. (2023). Ethereum blockchain platform. In *Distributed computing to blockchain* (pp. 267–282). Elsevier.

Ruan, P., Dinh, T. T. A., Lin, Q., Zhang, M., Chen, G., & Ooi, B. C. (2021). Lineagechain: A fine-grained, secure and efficient data provenance system for blockchains. *The VLDB Journal*, *30*, 3–24.

Saleh, M. A., Amanzholova, S. T., Sagymbekova, A. O., Zaurbek, A., & Almisreb, A. A. (2023). How can blockchain strengthen cybersecurity? unravelling the promises and challenges. *DTESI (workshops, short papers)*.

Salunke, N., Sonawane, S., & Motwani, D. (2023). Decentralized evidence storage system using blockchain and ipfs. *International Conference on Information, Communication and Computing Technology*, 259–280.

Sanchez, H., Tysebaert, S., Rath, A., & Rivière, E. (2022). AuditTrust: Blockchain-Based Audit Trail for Sharing Data in a Distributed Environment [EDCC 2022]. In S. Marrone et al. (Eds.), *Dependable Computing – EDCC 2022 Workshops* (Vol. 1656). Springer, Cham. https://doi.org/10.1007/978-3-031-16245-9_1

Sarode, R. P., Watanobe, Y., & Bhalla, S. (2022). A blockchain-based approach for audit management of electronic health records. *International Conference on Big Data Analytics*, 86–94.

Schulman, J. (2019). How blockchain technology will affect the audit [Accessed: 2025-01-23]. *RSM US LLP*. https://rsmus.com/insights/services/audit/how-blockchain-technology-will-affect-the-audit.html

Shafagh, H., Burkhalter, L., Hithnawi, A., & Duquennoy, S. (2017). Towards blockchain-based auditable storage and sharing of iot data. *Proceedings of the 2017 on cloud computing security workshop*, 45–50.

Shevchenko, E., & Lunsford, R. (2023). Blockchain disruption in finance: Jpmorgan chase's success story and the transfer of quorum to consensys. *Journal of Finance and Accountancy*, *32*.

Steichen, M., Fiz, B., Norvill, R., Shbair, W., & State, R. (2018). Blockchain-based, decentralized access control for ipfs. *2018 Ieee international conference on internet of things (iThings) and ieee green computing and communications (GreenCom) and ieee cyber, physical and social computing (CPSCom) and ieee smart data (SmartData)*, 1499–1506.

Sun, N., Zhang, Y., & Liu, Y. (2023). A universal privacy-preserving multi-blockchain aggregated identity scheme. *Applied Sciences*, *13*(6), 3806. https://doi.org/10.3390/app13063806

Tahir, S., & Rajarajan, M. (2018). Privacy-preserving searchable encryption framework for permissioned blockchain networks. *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 1628–1633.

Talha, M. (2024). Blockchain in accounting: Transforming transparency and security in financial records. *Dandao Xuebao/Journal of Ballistics*, *36*(1), 63–73.

Tang, J., Lu, X., Xiang, Y., Shi, C., & Gu, J. (2023). Blockchain search engine: Its current research status and future prospect in internet of things network. *Future Generation Computer Systems*, *138*, 120–141.

Theodorakopoulos, L., Theodoropoulou, A., & Halkiopoulos, C. (2024). Enhancing decentralized decision-making with big data and blockchain technology. a comprehensive review.

Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). Eu general data protection regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, *34*(1), 134–153.

Tran, H., Menouer, T., Darmon, P., Doucoure, A., & Binder, F. (2019). Smart contracts search engine in blockchain. *Proceedings of the 3rd International Conference on Future Networks and Distributed Systems*, 1–5.

REFERENCES

Usul, H., & Alpay, M. F. (2024). From traditional auditing to information technology auditing: A paradigm shift in practices. *European Journal of Digital Economy Research*, *5*(1), 3–9. https://doi.org/10.5281/zenodo.12819118

Vani, V., et al. (2024). Blockchain and ai for secure and sustainable healthcare development. In *Cybersecurity and data management innovations for revolutionizing healthcare* (pp. 308–329). IGI Global.

Verevka, T., & Shen, Z. (2024). Key factors in efficient auditing of investment projects in time of digitalization. In *Understanding the digital transformation of socio-economic-technological systems: Dedicated to the 120th anniversary of economic education at peter the great st. petersburg polytechnic university* (pp. 207–215). Springer.

Vimal, S., & Srivatsa, S. (2019). A new cluster p2p file sharing system based on ipfs and blockchain technology. *Journal of Ambient Intelligence and Humanized Computing*, 1–7.

Wang, H., & Yang, D. (2021). Research and development of blockchain recordkeeping at the national archives of korea. *Computers*, *10*(8), 90.

Wang, J., Wei, B., Zhang, J., Yu, X., & Sharma, P. K. (2021). An optimized transaction verification method for trustworthy blockchain-enabled iiot. *Ad Hoc Networks*, *119*, 102526.

Wang, K., Zhang, Y., & Chang, E. (2020). A conceptual model for blockchain-based auditing information system. *Proceedings of the 2nd International Electronics Communication Conference*, 101–107.

Wang, Q., Huang, L., Chen, S., & Xiang, Y. (2020). Blockchain enables your bill safer. *IEEE Internet of Things Journal*, *9*(16), 14162–14171.

Xie, Y., Zhang, C., Wei, L., Niu, Y., & Wang, F. (2019). Private transaction retrieval for lightweight bitcoin client. *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 440–446.

Yoo, H., Yim, J., & Kim, S. (2018). The blockchain for domain based static sharding. *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing*

And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 1689–1692.

Yue, C., Zhang, M., Zhu, C., Chen, G., Loghin, D., & Ooi, B. C. (2023). Veribench: Analyzing the performance of database systems with verifiability. *Proceedings of the VLDB Endowment*, *16*(9), 2145–2157.

Zahir, A., Groshev, M., Antevski, K., J. Bernardos, C., Ayimba, C., & De La Oliva, A. (2024). Performance evaluation of private and public blockchains for multi-cloud service federation. *Proceedings of the 25th International Conference on Distributed Computing and Networking*, 217–221.

Zhang, C., Xu, C., Wang, H., Xu, J., & Choi, B. (2021). Authenticated keyword search in scalable hybrid-storage blockchains. *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, 996–1007.

Zhang, X., & Poslad, S. (2018). Blockchain support for flexible queries with granular access control to electronic medical records (emr). *2018 IEEE International conference on communications (ICC)*, 1–6.

Zhang, Y., Zhang, Z., Ji, S., Wang, S., & Huang, S. (2024). Conditional proxy re-encryption-based key sharing mechanism for clustered federated learning. *Electronics*, *13*(5), 848. https://doi.org/10.3390/electronics13050848

Zhu, Y., Zhang, Z., Jin, C., Zhou, A., & Yan, Y. (2019). Sebdb: Semantics empowered blockchain database. *2019 IEEE 35th international conference on data engineering (ICDE)*, 1820–1831.