A Comparative Analysis of Methodologies: Blockchain and Data Engineering in Pursuit of Provenance, Integrity, and Decentralization

Sripada H Ravindranath[1] and Sai Ajay Kumar Peddu[2]

[1] Manager, Department of Data Engineering and Analytics, Equinix, Bangalore. India
[2] Senior Manager, Department of Data Engineering and Analytics, Equinix, San Francisco, USA
[1]sravindranath@equinix.com [2]speddu@equinix.com

Abstract - The swift advancement of data engineering paradigms necessitates creative approaches to tackle the problems of guaranteeing data accuracy, source, and expandability. This paper conducts a comprehensive investigation at the intersection of data engineering and blockchain technology, concentrating on two crucial aspects. The decentralized paradigm for data processing and storage, and the provenance and integrity of the data. The study explores the shortcomings of conventional approaches in preserving an unquestionable and transparent record of the origin and history of data in the field of data provenance and integrity. Acknowledging the weaknesses of centralized systems, the paper explores blockchain's revolutionary possibilities. Blockchain provides a safe and tamper-evident method for determining and maintaining data provenance by utilizing its distributed ledger and cryptography concepts. Smart contracts get us essential and Significant gains in the dedicated efficiency, and security, and also trust are possible when blockchain is included into data engineering techniques. The transparent and auditable nature of blockchain transactions serves to strengthen trust. The technology's cryptographic foundations increase security by guaranteeing the integrity and immutability of data. Central points of failure are expected to be eliminated, and more automated and streamlined data operations will be made possible.

This paper sheds light on the practical uses of blockchain integration into data engineering in addition to elucidating its theoretical underpinnings. The study advances the existing conversation on developing data engineering techniques in an efficient manner by tackling the crucial issues of data provenance and integrity and promoting decentralized data architectures.

Key Words: Provenance, Integrity, Decentralization, Block-Chain, Data-Engineering, Comparison

1. Introduction

Data engineering, an integral component of contemporary information technology landscapes, faces multifaceted challenges that span trust, security, and the inherent limitations of centralized architectures. In the Indian context, where a burgeoning digital economy intertwines with diverse data sources, ensuring the reliability and security of information becomes paramount. The centralization of data, a prevalent paradigm in traditional systems, poses inherent vulnerabilities, often manifesting in concerns related to unauthorized access, data tampering, and single points of failure. These challenges underscore the critical nature need for innovative and planned solutions that can fortify the foundations of data engineering in the Indian technological ecosystem.

Enter blockchain technology, a disruptive force that has garnered global attention for its potential to redefine how data is managed, secured, and transacted. The fundamental principles of blockchain—decentralization, immutability, and cryptographic security—hold relevance in the Indian context, where building robust and trustworthy digital infrastructures is imperative. Blockchain, which is a technically decentralized and distributed-ledger technology, addresses the limitations of centralized architectures by providing a transparent and tamper-resistant framework for recording transactions. In the realm of data engineering, the application of blockchain is poised to revolutionize how data provenance is established and maintained, ensuring an auditable record of data origin and modifications.

This introduction sets the stage for a nuanced exploration of the symbiotic relationship between blockchain and data engineering in the Indian context. As we navigate through the intricacies of trust, security, and centralized architectures, the article aims to uncover how blockchain's principles can serve as a catalyst for transformative advancements in data handling and processing, aligning with the evolving needs of India's digital landscape. In conclusion, blockchain technology holds immense potential for revolutionizing data engineering practices by ensuring data integrity, security, and transparency. Its decentralized architecture, immutability, consensus situational mechanisms, and smart contracts offer robust solutions to the challenges faced in traditional data management systems. As organizations continue to explore the integrational capabilities of blockchain-technology into their data engineering workflows, they stand to benefit from enhanced data quality, security, and efficiency in the rapidly evolving digital landscape.

2. Literature survey

Since Blockchain and data engineering methodologies have garnered attention for their roles in enhancing data provenance, integrity, and decentralization across various domains. Blockchain technology, initially developed for cryptocurrencies, has expanded its applications to domains like some of the supply-chain-management and healthcare, ensuring transparency and traceability. Studies by Swan et al. (2015) and Zheng et al. (2018) exemplify blockchain's potential in supply chain transparency and secure health data management, respectively. Blockchain's inherent features, including immutability and cryptographic security, contribute to preserving data integrity and provenance, as outlined by Yli-Huumo et al. (2016) and Zhang et al. (2018). In contrast, traditional data engineering methodologies, like relational databases and big data technologies, focus on efficient data processing and analysis. Research by Shah et al. (2019) and Kim et al. (2017) addresses data provenance and integrity challenges in big data environments, emphasizing techniques for capturing and assuring data lineage. Both blockchain and data engineering methodologies share the goal of ensuring trustworthy and decentralized data management, albeit through different approaches.

Blockchain offers transparency and decentralization but may face scalability limitations, whereas traditional data engineering provides robust processing capabilities but lacks the security of blockchain. A comparative analysis of these methodologies reveals the need for integrating blockchain's security features with traditional data engineering's efficiency for comprehensive data management solutions. Future research should focus on synergizing blockchain and data engineering techniques to address challenges and harness the benefits of both methodologies. This literature survey underscores the significance of understanding and leveraging the strengths of blockchain and data engineering for effective data governance and management. By bridging the gap between blockchain's security and data engineering's efficiency, organizations can enhance data integrity, provenance, and decentralization in their operations. The integrations of blockchain & data-engineering methodologies holds promise for revolutionizing data management practices and driving innovation in various industries. However, further research is needed to overcome challenges such as like scalability, application-interoperability, and regulatory-compliance in implementing integrated solutions. In conclusion, a holistic approach that combines the strengths of blockchain and data engineering methodologies is essential for addressing the evolving data management needs of the modern digital landscape.

3. Methodology

3.1 Blockchain and Data Provenance:

Data provenance refers to the documentation of the origin, lineage, and transformation history of data throughout its lifecycle. It encompasses information such as data sources, processing steps, transformations, and data quality metrics. Understanding data provenance provides insights into the reliable-nature, integrity, and trust-worthiness of data, enabling stakeholders to make informed decisions and ensure regulatory compliance.

3.1.1: Definition and Importance of Data Provenance:

Data provenance serves as a foundational concept in data engineering, providing much needed transparency and in terms of accountability in data management processes. Its importance can be outlined as follows:

A. Trust and Reliability: Data provenance enhances trust and reliability by enabling stakeholders towards trace the origin and evolution of data. It provides assurance regarding the authenticity of the content and accuracy of data-related items, fostering the confidence in decision-making processes.

B. Data Quality Assurance: By documenting data lineage and transformations, data provenance facilitates the identification and resolution of data quality issues. It enables data engineers to track errors, anomalies, and inconsistencies, improving data quality and reliability.

C. Regulatory Compliance: In regulated industries such as healthcare, finance, and pharmaceuticals, maintaining data provenance is essential for compliance with data governance and regulatory requirements. It ensures traceability, auditability, and accountability in data handling practices, mitigating the risk of non-compliance and penalties.

D. Auditing and Accountability: Data provenance supports auditing and accountability by providing a comprehensive record of data activities and access. It enables stakeholders to track data usage, modifications, and access permissions, enhancing transparency and accountability in data management processes.

E. Data Lineage Analysis: Data provenance facilitates lineage analysis, enabling stakeholders to get and understand how the data-flows through various systems, processes, and transformations. It helps identify dependencies, bottlenecks, and optimization opportunities in data pipelines, improving efficiency and performance.

3.2 Blockchain Integration for Enhancing Data Provenance:

Blockchain technology offers unique capabilities for enhancing data provenance in data engineering processes. Its decentralized, immutable, and transparent nature aligns well with the requirements of data provenance, providing a tamper-proof and auditable record of data transactions.

By integrating blockchain into data engineering workflows, organizations can:

A. Immutability: Blockchain's immutable ledger ensures that once data is recorded, it cannot be altered or deleted retro-actively. This feature enhances the integrity and trustworthiness of data provenance, providing a reliable record of data lineage and transformations.

B. Transparency: Blockchain enables real-time access to data transactions for all participants in the network, fostering transparency and accountability in data management processes. Stakeholders can verify the authenticity and integrity of data provenance, reducing the risk of manipulation or fraud.

C. Decentralization: Blockchain's decentralized architecture eliminates the need for central authorities or intermediaries, ensuring resilience and censorship resistance in data provenance. Data provenance records stored on the blockchain are distributed across the multiple nodes, and reducing the risk of single-points-of-failure or tampering.

D. Smart Contracts: Smart contracts can automate and enforce data provenance rules and policies, ensuring compliance with data governance and regulatory requirements. They can define conditions for data access, usage, and sharing, enhancing security and accountability in data engineering processes.

In conclusion, data provenance plays a critical role in ensuring the reliability, integrity, and trustworthiness of data in data engineering. Its documentation of data lineage and transformations provides transparency, accountability, and compliance in data management processes. By integrating blockchain technology into data engineering workflows, organizations can enhance data provenance by leveraging its immutable, transparent, and decentralized nature. This integration holds the potential to revolutionize data management practices, ensuring data integrity, security, and trust in an increasingly digital and interconnected world.

Now let's try to understand the Comparison of processing and complication involved in Both Traditional (Without Blockchain) and data engineering with Blockchain mechanism. Please refer the Fig. 1

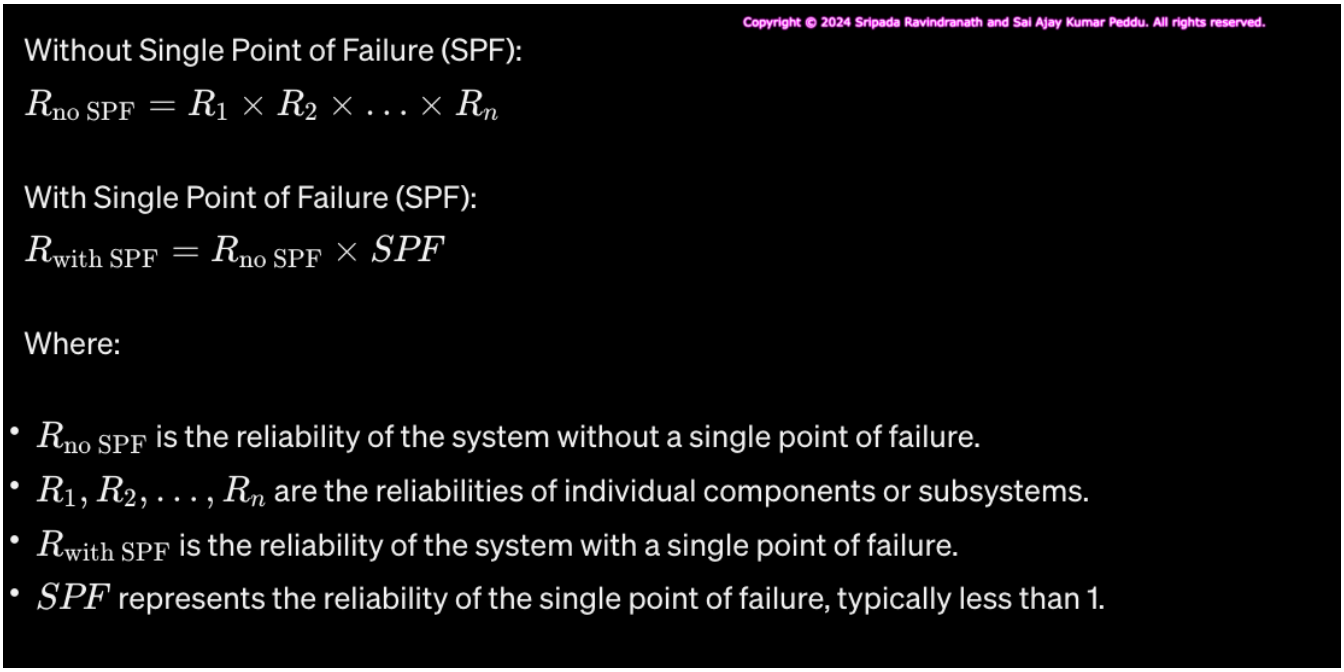

Fig. 1 Comparison of Single Point of Failure

Now let's compare the Two items about **Data Processing time** and **Resource Utilization as** shown in Fig. 2 below

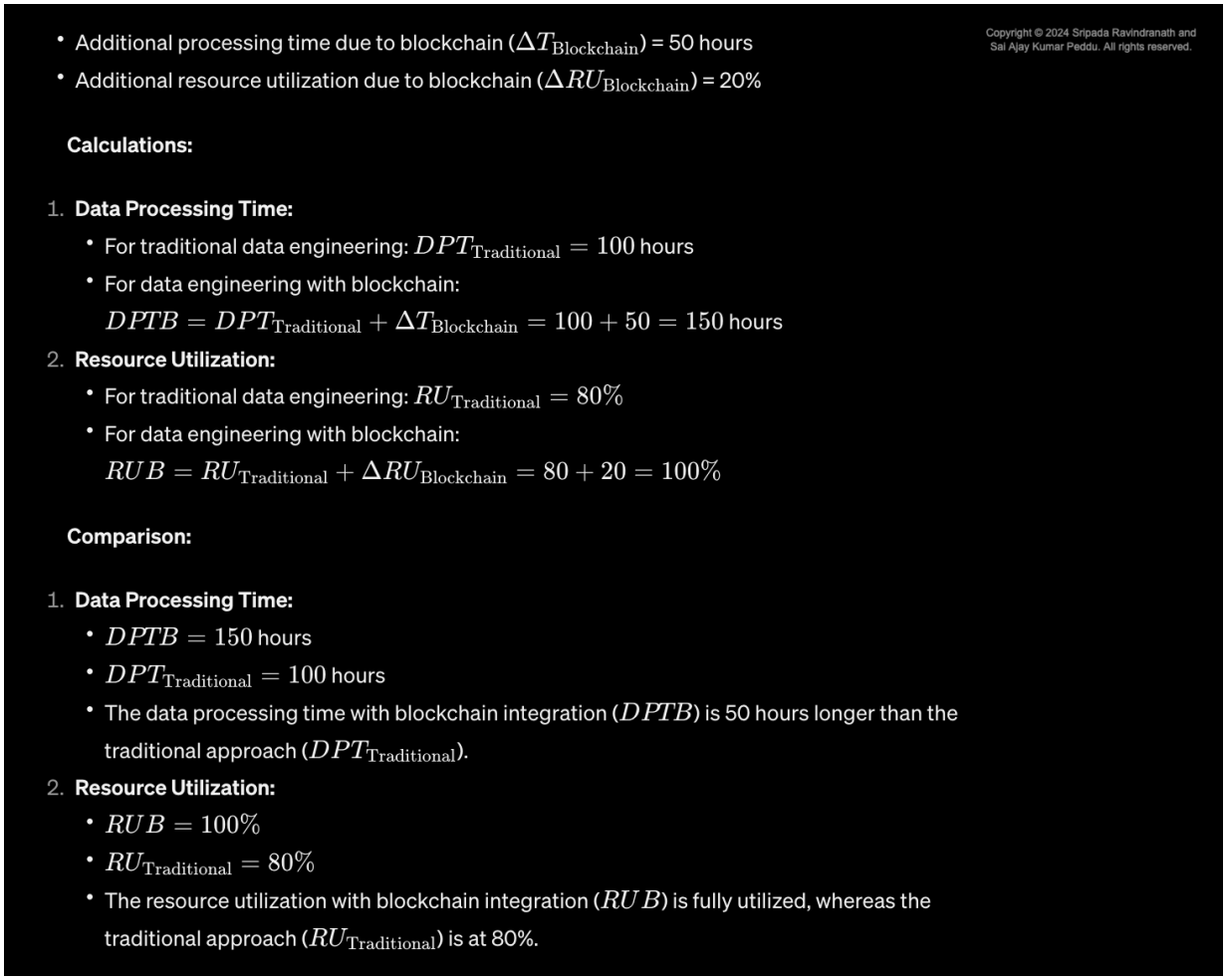

Fig. 2 Comparison of Data processing-time and Resource Utilization

3.2 Role of Blockchain in Data Provenance:

3.2.1 Overview of how blockchain ensures transparency and traceability in data transactions:

Let's delve into how blockchain technology contributes to ensuring transparency and traceability in data transactions:

A. Transparency:

Blockchain ensures transparency by providing an immutable and publicly accessible ledger of transactions. Every transaction that occurs on a blockchain is recorded in a block, which is cryptographically linked to the previous block, forming a chain of blocks (hence the name blockchain).

Each participant in the blockchain network has a copy of the entire transaction-history, allowing real-time verification and validation of transactions by any interested party. Because the blockchain is decentralized and distributed across multiple nodes, there is no central authority controlling the data. This decentralization enhances transparency as there is no single point-of-failure or any manipulation.

B. Traceability:

Blockchain facilitates traceability by providing a clear and auditable trail of data transactions from their origin to their current state. Each transaction is timestamped and cryptographically secured, making it tamper-proof and verifiable.

Participants can trace the lineage of data, tracking its movement and transformations throughout its lifecycle on the blockchain. This capability is particularly valuable in supply chain management, where the provenance of goods and raw materials can be crucial for ensuring quality, authenticity, and compliance.

Smart contracts, which are self-executing contracts with predefined-rules encoded in-code, can automate and enforce specific conditions and actions based on data transactions. This further enhances traceability by enabling automated, transparent, and auditable execution of business logic.

C. Smart Contract as tool:

Blockchain ensures transparency and traceability in data transactions, enabling organizations to track the origin, ownership, and movement of data with a high degree of accuracy and security. Smart contracts, which are self-executing those underlying contracts with the dedicated terms of the agreement directly written into code, can serve as a mainstream-powerful tool for enforcing data provenance rules from the before mentioned blockchain. Here's how smart contracts facilitate this

C.1. Automated Execution:

Smart contracts automatically execute predefined rules and conditions encoded within their code when certain triggers or conditions are met. These rules can include data provenance requirements such as verifying the authenticity of data sources, validating data integrity, or enforcing access control policies.

C.2 Tamper-proof Execution:

Once deployed on the blockchain, smart contracts operate in a tamper-proof environment, ensuring that the execution of data provenance rules is transparent, auditable, and resistant to tampering or manipulation. This enhances the trustworthiness and reliability of data transactions.

C.3 Decentralized Enforcement:

Smart contracts designed to work on designated mode of decentralized network/networks of nodes, eliminating need for intermediaries/or centralized authorities to enforce data provenance rules. This decentralized enforcement ensures that data transactions adhere to predefined rules without relying on trusted third parties.

C.4 Immutable Recordkeeping:

The underlying execution of smart-contracts and their outcomes are recorded on the blockchain, providing an immutable record of data transactions and the enforcement of data provenance rules. This enables organizations to maintain a transparent & failry-auditable trail of data provenance activities over time.

Now lets' consider the data from sampling information and plot the graph to visualise and realise the threshold and single-point-of-Failure between the items under comparison as shown in below Fig. 3

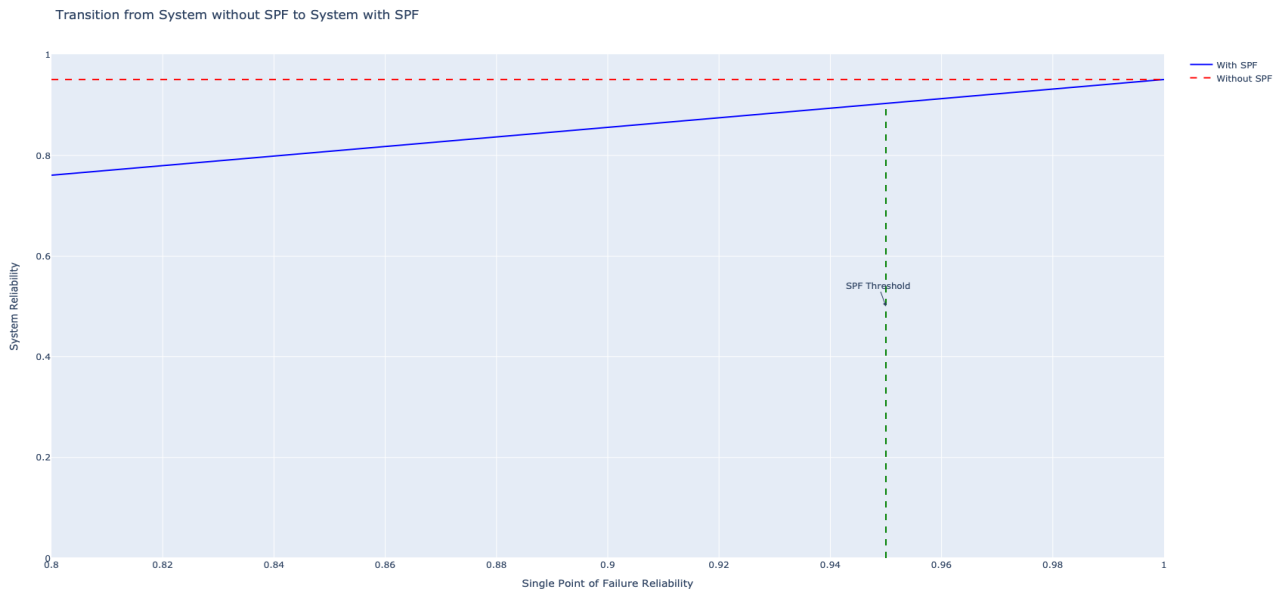

Fig. 3 Impact of Single-Point-of-Failure

Now let's analyse the Graph in details. The graph visually demonstrates the impact of a single-point-of-failure on the reliability of a system. The red dashed line represents the ideal scenario where there is no single-point-of-failure, resulting in maximum system reliability. By comparing the blue line (with SPF) to the red dashed line

(without SPF), viewers can easily grasp the significance of eliminating single-points-of-failure to improve overall system reliability.

In summary, smart-contracts serve as powerful tool for enforcing data provenance rules on blockchain, leveraging automation, tamper-proof execution, decentralization, and immutable recordkeeping to ensure transparency, integrity, and traceability in data transactions.

4. Ensuring the Data-Integrity with Blockchain:

Ensuring data integrity is system-crucial for maintaining the accuracy, reliability, and trustworthiness of data in various applications. Blockchain technology offers innovative-solutions to address many challenges through its decentralized and tamper-evident nature.

4.1 Traditional Approaches to Data Integrity:

Traditional approaches to data integrity involve numerous methods and fair-techniques used in centralized systems before the advent of blockchain technology. These methods typically include:

A. **Checksums and Hashing**: Checksums and cryptographic hashing algorithms such as MD5, SHA-256, and SHA-3 are commonly used to verify the integrity of data by generating a unique fixed-size hash value based on the input data. Any alteration to the data will result in a different hash value, indicating potential tampering.
B. **Digital Signatures:** Digital signatures are used to authenticate origin-of-the-data and integrity of data by applying cryptographic techniques. A digital signature is generated using the sender's private key and can be verified by anyone using the sender's public key, ensuring data integrity and non-repudiation.
C. **Access Controls and Encryption:** Access controls and encryption mechanisms are employed to restrict unauthorized access to data and protect it from tampering or modification. Role-based access controls (RBAC) and encryption algorithms help safeguard data integrity by ensuring that only-authorized-users can modify or access sensitive data.

4.2 Challenges in ensuring data integrity using traditional methods:

Despite their wide-spread adoptions, traditional methods of ensuring data integrity have several limitations and challenges:

A. **Centralization:** Many traditional data integrity mechanisms rely on centralized systems and trusted third parties to verify and enforce data integrity, making them vulnerable to single points of failure, manipulation, and collusion.
B. **Complexity and Cost:** Implementing and managing traditional data integrity solutions can be complex and expensive, requiring specialized skills, infrastructure, and ongoing maintenance.
C. **Limited Transparency and Auditability:** Traditional approaches to data integrity may lack transparency and auditability, making it difficult to verify the integrity of data transactions and detect unauthorized modifications or tampering.

4.3 Vulnerabilities in centralized systems leading to data tampering risks:

Centralized systems, where data is stored and managed by a single authority/ or organization, are prone/susceptible to various vulnerabilities that can compromise data integrity:

A. Single Point of Failure:

Centralized systems have a single point of failure, such as a central server or database, which, if compromised, can lead to widespread data tampering or manipulation.

B. Insider Threats:

Authorized insiders with privileged access to centralized systems may abuse their privileges to tamper with data or manipulate records for personal gain or malicious purposes.

C. External Attacks:

Centralized systems are vulnerable to external cyber-attacks, such as hacking, malware, or denial-of-service (DoS) attacks, which can result in data breaches, unauthorized access, or data manipulation.

D. Lack of Transparency:

Centralized systems may lack transparency and accountability, making it challenging to detect and prevent data tampering or unauthorized modifications.

Now let's Figure out the mathematical relations related to vulnerabilities in the centralised systems by making use of the contents like lack of Transparency.

5. Decentralized Data Storage and Processing

Decentralized data storage and processing have emerged as crucial components in modern data engineering, offering solutions that address the challenges of provenance, integrity, and decentralization. This section presents a comparative analysis of centralized and decentralized architectures, focusing on their impact on scalability, reliability, and security.

5.1 Centralized vs. Decentralized Architectures

5.1.1 Centralized Architectures:

Centralized data architectures are characterized by a single point of control, where data is stored, processed, and managed within a centralized infrastructure. In a centralized model, data is typically stored in a single location or a limited number of servers, with access and control managed by a central authority or organization.

A. Advantages of Centralized Architectures:

1. **Simplicity**: Centralized architectures are often simpler to implement and manage, requiring fewer resources and expertise.
2. **Centralized Control:** Centralized architectures offer centralized control over data access, security policies, and management, providing a clear hierarchy of authority and accountability.

B. Disadvantages of Centralized Architectures:

1. **Single Point of Failure:** Centralized architectures are vulnerable to single points of failure, where the failure of a central server or system can disrupt access to data and services.
2. **Lack of Scalability:** Centralized architectures may struggle to scale effectively to accommodate growing data volumes or user demands, leading to performance bottlenecks and reduced reliability.
3. **Security Risks:** Centralized architectures present security risks such as data-breaches, un-authorized access, mostly an insider threats, as they dependent/rely on a single entity to safeguard sensitive information.

5.1.2 Decentralized Architectures:

Decentralized data architectures distribute data storage and processing across a network of nodes, eliminating the need for a central authority or intermediary. In a decentralized model, data is replicated and synchronized across multiple nodes, with consensus mechanisms ensuring agreement on the state of the network.

A. Advantages of Decentralized Architectures:

1. **Fault Tolerance:** Decentralized architectures are resilient to single points of failure, as data is mainly distributed across multiple-nodes. This enhances reliability and availability, as the failure-of-a-single-node does not disrupt the entire network.
2. **Scalability:** Decentralized architectures can scale more effectively to accommodate growing data volumes and user demands, as additional nodes can be added to the network without centralized coordination.
3. **Data Sovereignty:** Decentralized architectures empower individuals and organizations with greater control and ownership over their data, reducing reliance on centralized authorities and promoting data sovereignty.

B. Disadvantages of Decentralized Architectures:

1. **Complexity:** Decentralized architectures can be more complex to design, implement, and manage compared to centralized architectures, requiring specialized knowledge and expertise.
2. **Consensus Overhead:** Decentralized architectures incur overhead associated with achieving consensus among network nodes, which can impact performance and scalability.
3. **Security Challenges:** Decentralized architectures face security challenges such as 51% attacks, Sybil attacks, and double-spending attacks, which require robust cryptographic mechanisms and consensus algorithms to mitigate.

C. The Impact of Decentralization on Scalability, Reliability, and Security

1. Scalability:

Decentralized architectures offer inherent scalability benefits compared to centralized architectures, as they can dynamically scale to accommodate growing data volumes and user demands by adding additional nodes to the network. However, achieving scalability in decentralized architectures requires efficient consensus mechanisms, sharding-techniques, and network optimization strategies to maintain performance and throughput.

2. Reliability:

Decentralized architectures improve reliability by eliminating single-points-of-failure and distributing data storage and processing across multiple nodes. This enhances fault tolerance and availability, as the failure of a single node does not compromise the integrity or accessibility of data. However, ensuring reliability in decentralized architectures requires robust redundancy mechanisms, data replication strategies, and network monitoring to detect and mitigate failures promptly.

3. Security:

Decentralized architectures enhance security by reducing the attack surface and vulnerabilities associated with centralized points of control. By distributing data and control across a network of nodes, decentralized architectures mitigate risks such as data-breaches, un-authorized access, and censorship. However, decentralization introduces new security challenges such as consensus attacks, network partitioning, and malicious node behaviour, which require robust cryptographic protocols, consensus mechanisms, and network governance to address effectively.

Below is the steps involved in designing a mechanism or Algorithmic steps for figuring out Impact of Decentralised systems . Please refer the Fig. 4 Below

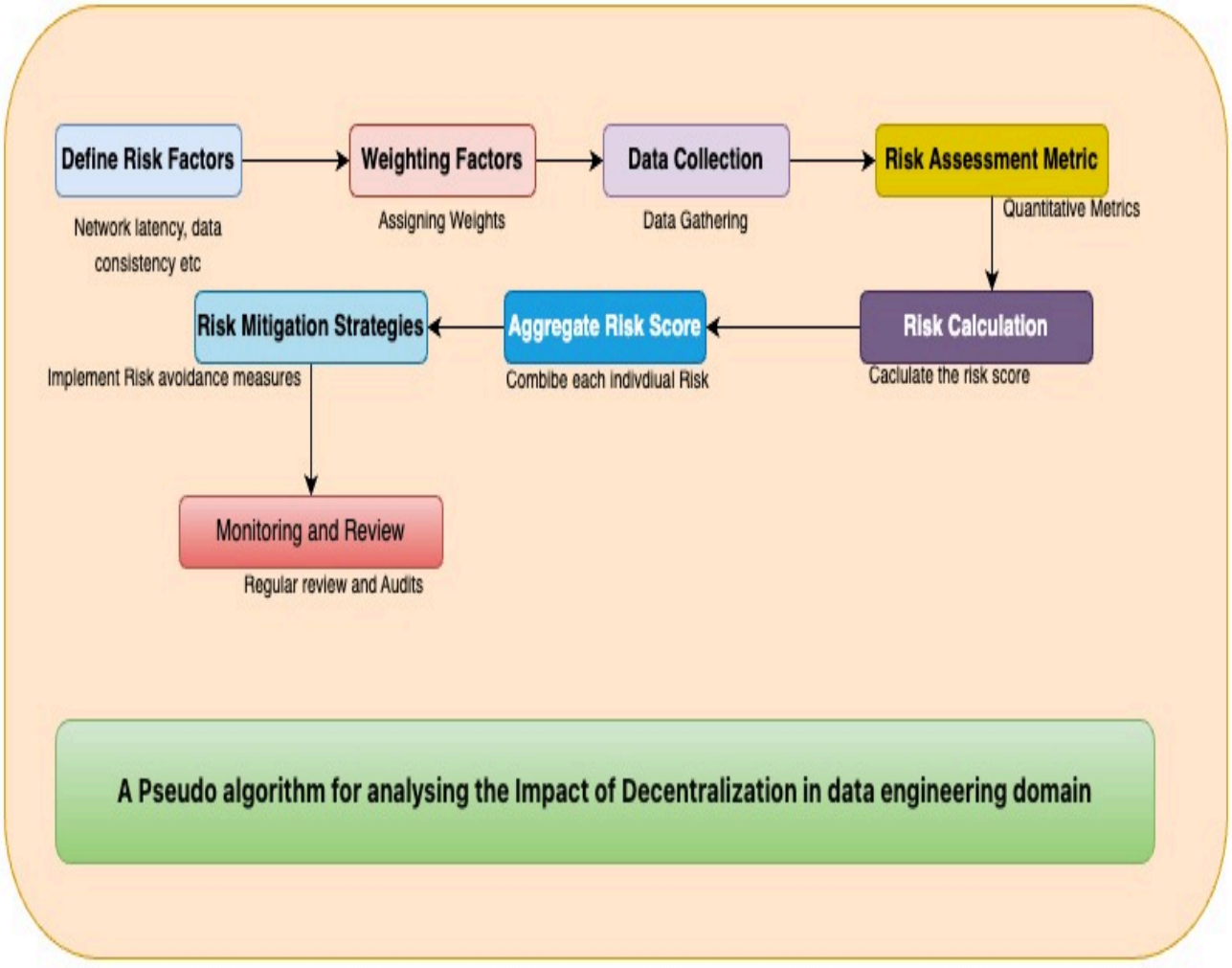

Fig. 4 Impact of Decentralization

Let’s analyse the risk factor here as shown in the below image Fig. 5

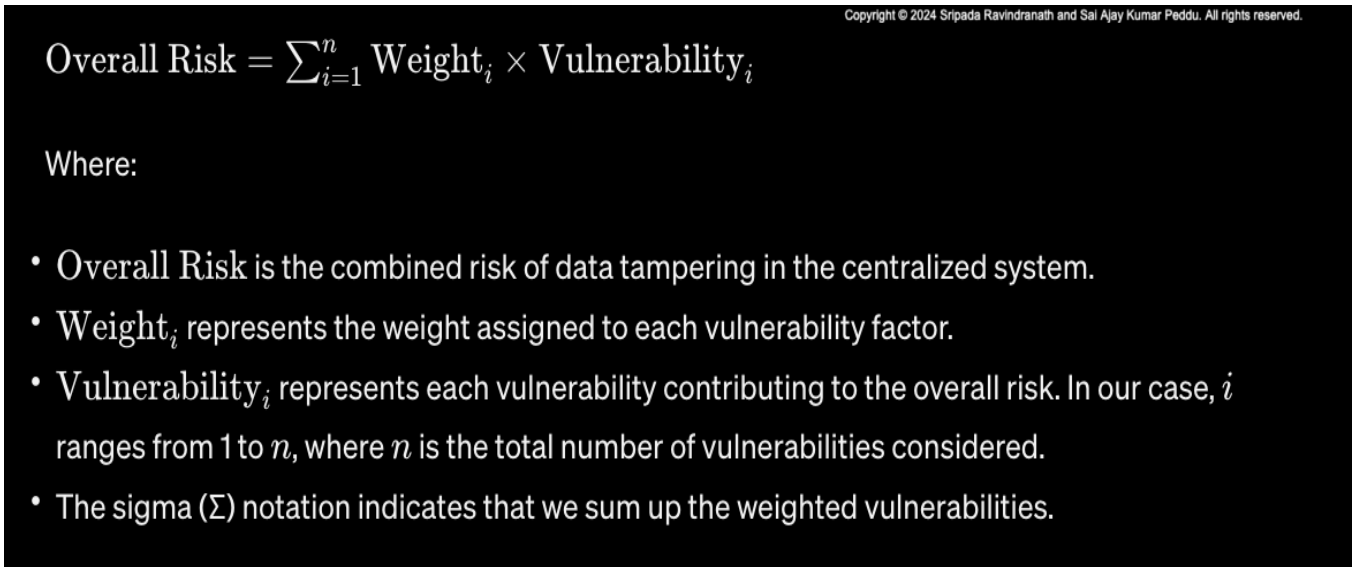

Fig. 5 Risk Analysis of Decentralized systems

In conclusion, decentralized data storage and processing offer significant advantages over centralized architectures in terms of scalability, reliability, and security. While centralized architectures provide simplicity and centralized control, decentralized architectures empower individuals and organizations with greater resilience, scalability, and data sovereignty. By understanding the comparative analysis of centralized and decentralized architectures, organizations can make informed decisions when designing and implementing data engineering solutions that prioritize provenance, integrity, and decentralization.

6. Case Studies and practical Implementation

Practical Usage: Real-world examples demonstrate blockchain's effectiveness in various sectors, such as supply-chain-management (Walmart), healthcare data management (Mayo Clinic), and financial services (JPMorgan Chase). These implementations highlight blockchain's ability to enhance transparency, streamline operations, and improve data security in data engineering applications.

Best Practices: Organizations embarking on blockchain projects should start by identifying clear and compelling use cases aligned with their business objectives. Collaboration with industry partners and stakeholders is crucial for ecosystem building and addressing interoperability challenges. Prioritizing data privacy and security through robust encryption and access control mechanisms is essential to ensure compliance and protect sensitive information.

Lessons Learned: Scalability and performance limitations of blockchain technology should be carefully evaluated, and organizations should explore scalability solutions such as sharding and off-chain scaling to accommodate large-scale data processing requirements. Continuous evaluation and improvement are essential, with organizations monitoring key performance indicators, gathering user feedback, and iterating on blockchain solutions to optimize effectiveness and usability over time.

7. Conclusions:

The Overall processing capacity increases with increase in the best practices mentioned in this paper. Tighter the module processing better will be the throughput and latency of the downstream layer. Few of the key observation made in this analysis is provided below in the Table 1.

Criteria	Traditional Data Engineering Techniques	Blockchain-Based Solutions
Provenance	Provenance tracking relies on centralized databases and audit trails.	Provenance is inherently built into the immutable ledger of blockchain, providing transparent and traceable data lineage.
Integrity	Data integrity is ensured through cryptographic hashing and access controls in centralized databases.	Blockchain offers tamper-proof data storage and cryptographic verification mechanisms, enhancing data integrity.
Decentralization	Traditional techniques may rely on centralized servers and databases, leading to single points of failure and vulnerability to attacks.	Blockchain decentralizes data storage and processing, reducing reliance on intermediaries and enhancing resilience and security.

Table. 1 Comparison between Traditional technique vs Blockchain Based solutions.

In conclusion, the overall choice between traditional data engineering techniques and blockchain-based solutions depends on the specific requirements and objectives of the organization. Traditional techniques mainly excel in these scenario's where centralized control, efficient performance, and established infrastructure are prioritized.

On the other hand, blockchain-based solutions offer significant advantages in environments where transparency, immutability, and decentralized governance are paramount. Industries seeking to adopt blockchain for data engineering should carefully assess their needs, considering factors such as data sensitivity, scalability requirements, regulatory compliance, and network participation.

Collaborative efforts between data engineers, blockchain developers, and domain experts are crucial for designing tailored solutions that leverage the strengths of both methodologies effectively. By embracing a strategic-approach and evaluating trade-offs, industries can harnessing transformational potential of blockchain technology to enhance provenance, integrity, and decentralization in their data engineering practices.

REFERENCES

1. X. Wang, J. Jia, Y. Cao, J. Du, A. Hu, Y. Liu, and Z. Wang, "Application of data storage management system in blockchain-based technology," in 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), 2023, pp. 1-6.
2. L. Yang, S. Fang, L. Sheng, L. Dandan, S. Hangxuan, W. Yingying, N. Liu, and X. Chen, "Research and Application of Archive Data Management System Based on Blockchain," in 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), 2023, pp. 1-6.
3. X. Gao, "Research On Privacy Protection Scheme For Educational Data Based On Blockchain," in 2023 3rd International Conference on Computer Science and Blockchain (CCSB), 2023, pp. 1-5.
4. Q. Guo, S. Chen, J. Wang, and X. Pan, "Research and Design of Electric Power Engineering Project Management System Bsed on Blockchain Technology," in 2022 International Conference on Blockchain Technology and Information Security (ICBCTIS), 2022, pp. 1-6.
5. Z. Wu, H. Pan, X. Si, H. Qian, and F. Cai, "A Fair and Reliable Data Trading Scheme based on Blockchain," in 2022 2nd International Conference on Computer Science and Blockchain (CCSB), 2022, pp. 1-5.
6. X. Mogeng, X. Tong, Y. Fu, and Y. Wang, "Research and Application of Key Technologies on Scientific Marine Data Integration, Submission and Credible Sharing," in 2023 IEEE 14th International Conference on Software Engineering and Service Science (ICSESS), 2023, pp. 1-5.
7. P. Fei, W. Zeng, and P. Zhang, "Application of blockchain technology in power grid engineering safety and big data storage management," in 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), 2021, pp. 1-6.

8. M. Xie, Z. Liao, and L. Huang, "Data Security Based on Blockchain Digital Currency," in 2020 3rd International Conference on Smart BlockChain (SmartBlock), 2020, pp. 1-5.
9. Z. Han, L. Li, and Y. Ding, "Research on the Application of Blockchain Technology in Defense Engineering," in 2022 2nd International Conference on Computer Science and Blockchain (CCSB), 2022, pp. 1-5.
10. X. Yang, Y. Li, L. Chen, W. Feng, and Z. Yan, "TDL-Chain: An Intelligent Data Transmission Control System in Tactical Data Link Based on Blockchain," in 2020 IEEE International Conference on Blockchain (Blockchain), 2020, pp. 1-6.
11. S. Peng, D. Sun, L. Zhu, H. Zhou, X. Zhang, and C. Cui, "Enhancing Cross-Border Data Sharing in Blockchain Networks: A Compliance-Centric Approach Ensuring Anonymity and Traceability," in 2023 3rd International Conference on Computer Science and Blockchain (CCSB), pp. 1-6.
12. M. Singh, "Using Blockchain Technology to Secure Autonomous Vehicles," in 2020 IEEE International Conference on Smart BlockChain (SmartBlock), 2020, pp. 1-6.
13. Al-Azzoni, S. Iqbal, and N. Petrović, "Data Analytics on Blockchains," in 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2023, pp. 1-6.
14. Wang, J. Si, X. Zang, S. Ding, C. Liu, J. Pan, and J. Shen, "A Blockchain Based Data Auction Mechanism," in 2023 International Conference on Blockchain Technology and Information Security (ICBCTIS), 2023, pp. 1-7.