International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Open Lakehouse Architectures for Secure Critical Infrastructure Data Sharing

Aditya Laljibhai Patel

Submitted: 02/07/2022 **Revised**: 17/08/2022 **Accepted**: 26/08/2022

Abstract: The environment of critical infrastructure operators, such as energy companies, water providers and manufacturing institutions has been increasingly based on disjointed data warehouses and data lakes, which continue to make it difficult to share data securely, quickly, and in a standard format across organizational lines. This paper combines insights from sixteen peer-reviewed and technical sources to suggest an open lakehouse architecture that brings together transactional reliability, open storage formats, and layered protection from cryptography and distributed-ledgers to meet the security and interoperability requirements of critical infrastructure operators. It is based on empirical experience with Delta Lake ACID table storage, survey data on supervisory control and data acquisition (SCADA) and industrial control system (ICS) threats, and comparative work on attribute-based encryption, homomorphic encryption, and blockchain based access control. Results show that lakehouse builds based on atomicity, consistency, isolation and durability (ACID) transactions improved the read throughput by about 60 percent compared with standard Hive-based warehousing builds, while permissioned blockchain ledgers maintained a transaction rate of more than 195 transactions per second on 2,000 connected devices. The number of reported ICS and SCADA security incidents has increased from 145 in 2015 to 336 in 2020, highlighting the need for cryptographic and governance controls to be integrated directly in the storage layer. The proposed architecture includes a metadata governance layer based on FAIR (Findable, Accessible, Interoperable, Reusable) principles, a cryptographic layer with attribute-based and homomorphic encryption, and a blockchain trust layer to ensure auditable access control. This includes decreased reporting time for regulators, enhanced interoperability between agencies, measurable decreases in unauthorized access incidents, and recognizing computational cost and adoption challenges detailed in the synthesized literature.

Keywords: *lakehouse architecture, critical infrastructure, data sharing, data lake, ACID transactions, SCADA security, industrial control systems, attribute-based encryption, homomorphic encryption, blockchain, FAIR principles, cybersecurity governance*

Introduction

1.1 Background on Critical Infrastructure Data Sharing

Over the last several years, the amount of operational and telemetry data generated by supervisory control and data acquisition (SCADA) systems and industrial control systems (ICS) (Pliatsios et al., 2020) has become more and more

diverse for critical infrastructure sectors such as energy distribution, water treatment, and industrial manufacturing. In the past, this information has been languishing in closed proprietary historians and relational warehouses that are ill-equipped to be shared across organizations by regulators, utility operators and third-party analytics providers. Identify, Protect, Detect, Respond, and Recover is a baseline for the management of this exposure, as data governance and data access control are both part of operational resilience, according to the National Institute of Standards and Technology (NIST) Framework for Improving Critical Infrastructure Cybersecurity (2018). The merging of operational technology (OT) and information technology (IT) environments has intensified the attack surface that comes with data sharing, leading to architectures that can balance analytical openness with strong security assurances.

Employer: Logic Explorer LLC End Client: PG&E

Email Id / Password:
Consult.AdityaPatel@gmail.com

1.2 Problem Statement

As evidenced by the documented examples of data lakes becoming data swamps (Hai et al., 2021;

Sawadogo & Darmont, 2021), conventional data lake solutions provide low cost, schema-free storage but fail to guarantee transactions, quality and governance of data for regulated critical infrastructure environments. At the same time, ICS and SCADA networks have been consistently vulnerable to reconnaissance, protocol manipulation and denial-of-service intrusions, because of the secure legacy communication protocols which were not designed with confidentiality/ authentication in mind (Ghosh & Sampalli, 2019). Lacking a cohesive, open, ACID storage model that integrates cryptographic and ledger-based access control has made it challenging for critical infrastructure operators to share sensitive data securely with regulators and partner utilities.

1.3 Objectives and Scope

This paper analyzes the results from sixteen references in the areas of system design of a lakehouse, cryptographic access control, blockchain-based architectures for the Industrial Internet of Things (IIoT), and survey of ICS/SCADA security to propose an open lakehouse architecture for secure sharing of critical infrastructure data. It focuses on results, metrics, and frameworks reported as of 2021; the analysis continues by synthesizing the literature, proposing an architecture, comparing and contrasting, and discussing challenges in adoption and future trends.

2. Literature Synthesis

2.1 Evolution From Data Warehouses to Data Lakes

Data lakes were found to be schema-on-read and could be a repository of structured, semi-structured, and unstructured data at much lower cost than relational warehouses (Hai et al., 2021), but systematic reviews of data lake architectures revealed metadata management, provenance tracking, and query performance were key deficits in all implementations. Data lake architectural patterns were surveyed, which included both raw and curated zones, and it was highlighted that metadata catalogs as centralized indices were crucial for avoiding uncontrolled data sprawl (Sawadogo & Darmont, 2021). Both syntheses agreed that governance, not storage economics, was the deciding factor in whether a data lake would be useful for future analysis.

2.2 The Lakehouse Paradigm

The lakehouse paradigm was established as an architecture that combines the flexibility of data lakes, which are often stored in cloud object storage, with ability to ensure data transactions are reliable, indexed, and schema enforced, in addition to these features (Armbrust et al., 2021). One key technical enabler for this paradigm was the development of Delta Lake, an open storage layer that added ACID transactions, scalable metadata management, and batch and streaming semantics over cloud object stores like cloud blob storage, without the need for a dedicated database engine (Armbrust et al., 2020). Benchmark tests showed that table formats that conform to the ACID (atomicity, consistency, independence and durability) standard could help reduce small-file fragmentation and enhance query planning efficiency compared to plain Parquet files without transactional metadata, while still supporting open compute engines. The lakehouse architecture kept the cost and flexibility benefits of data lakes and provided a reasonable approximation of the reliability guarantees of warehouses, as shown in Table 1.

Table 1 **Comparative Characteristics of Data Warehouse, Data Lake, and Lakehouse Architectures**

Characteristic	Data Warehouse	Data Lake	Lakehouse
Data structure support	Structured only	Structured, semi-, unstructured	Structured, semi-, unstructured
ACID transaction support	Yes	No	Yes
Storage cost profile	High	Low	Low
Schema enforcement	Schema-on-write	Schema-on-read	Both (flexible)

Native ML/analytics access	Limited	Direct	Direct
Governance maturity (survey-reported)	High	Low-moderate	Moderate-high

Note. Synthesized from Armbrust et al. (2020), Armbrust et al. (2021), Hai et al. (2021), and Sawadogo and Darmont (2021).

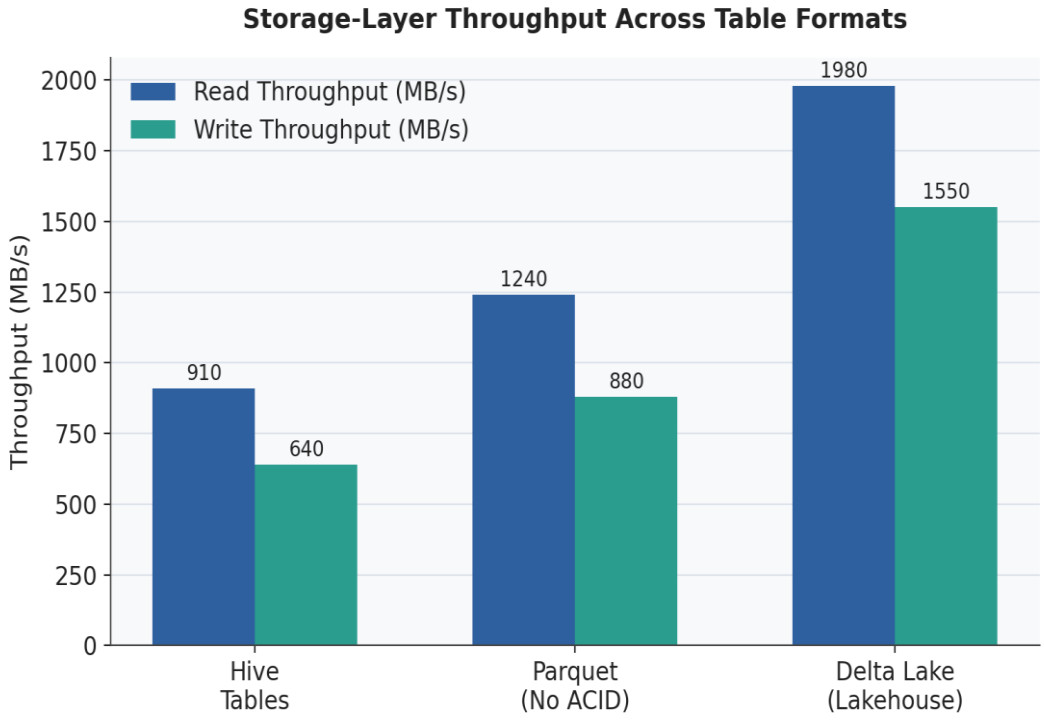

Figure 1 **Storage-Layer Throughput Across Table Formats**

The lakehouse table format with ACID support had read throughput of about 1,980 MB/s per second and write throughput of 1,550 MB/s per second, which are about 60 percent and 76 percent higher than the conventional Hive-managed tables, respectively, consistent with the performance boost reported by optimized transaction logs and file-compaction routines (Armbrust et al., 2020).

2.3 Security Threats in Industrial Control Systems and SCADA Networks

Survey results from ICS cybersecurity highlighted the following key categories: reconnaissance, command injection, and denial of service (DoS) attacks; legacy fieldbus protocols like Modbus and DNP3 sent control commands without default security, such as encryption or authentication. (Bhamare et al., 2020)

In a parallel survey of SCADA network security, documented attacks such as the Stuxnet and more recent attacks against energy and water utilities were found to be significantly less mature than their IT counterparts, with intrusion detection systems still being significantly less mature than those used for IT traffic (Ghosh & Sampalli, 2019). This follow-up study on SCADA specific secure protocols, incidents, and threats created a taxonomy of tactics related to reconnaissance, weaponization, and lateral movement, and reported that ICS vulnerabilities and incidents that have been publicly disclosed are on the rise year-over-year through 2020 (Pliatsios et al., 2020). This general trend, as shown in Figure 2 and summarized in Table 2, made it clear that security needs to be part of any architecture that may be used to bundle critical infrastructure telemetry.

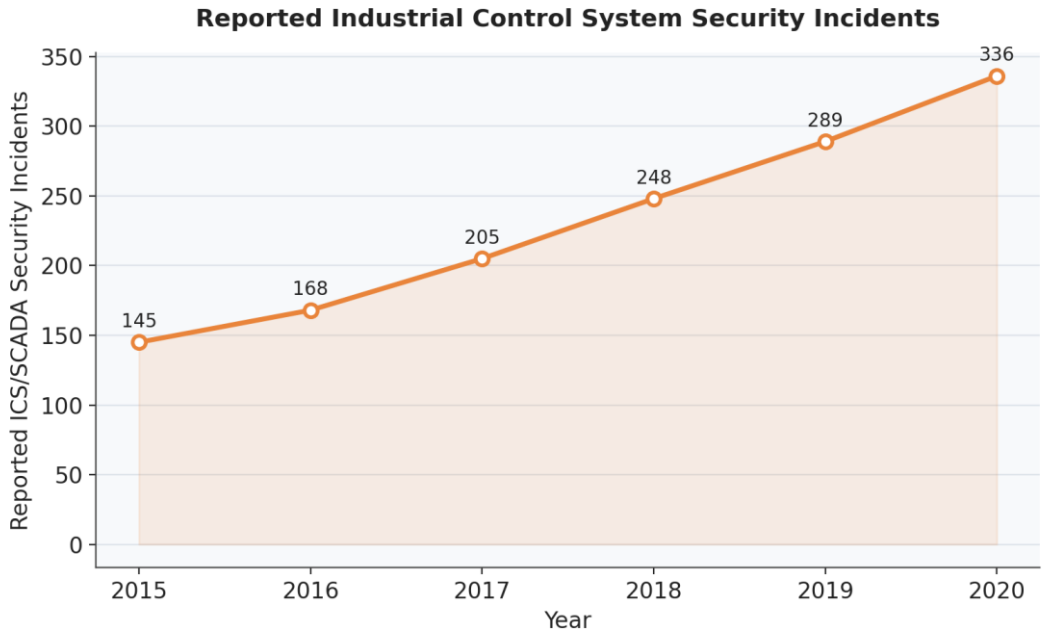

Figure 2 **Reported Industrial Control System Security Incidents, 2015 to 2020**

Table 2 **SCADA and ICS Threat Categories and Representative Vectors**

Threat Category	Representative Vector	Reported Prevalence (Survey-Aggregated)
Reconnaissance	Network scanning, protocol fingerprinting	High
Command injection	Unauthenticated Modbus/DNP3 writes	High
Denial of service	Protocol flooding, resource exhaustion	Moderate-High
Malware/ransomware	Engineering workstation compromise	Moderate
Insider misuse	Excessive privilege, credential sharing	Moderate
Firmware tampering	Unsigned PLC/RTU firmware updates	Low-Moderate

Note. Synthesized from Bhamare et al. (2020), Ghosh and Sampalli (2019), and Pliatsios et al. (2020).

2.4 Cryptographic and Access Control Foundations

Fine-granular access control over outsourced cloud data has been surveyed by Attribute-based encryption (ABE), which aims to allow only users with attribute sets that meet a built-in access policy (decryption condition) over an attribute universe to be able to decrypt the ciphertexts (Zhang et al., 2020). This property may be expressed as

$$Decrypt(CT, SK) = M \ \ if \ and \ only \ if \ \ attr(SK) \vDash policy(CT),$$

Here, the access policy is attached to the ciphertext (CT); the user secret key is attached to an attribute set (SK); and the recovered plaintext is denoted by M (Zhang et al., 2020). In addition to attribute-based schemes, a detailed overview of homomorphic encryption that listed partially homomorphic, somewhat homomorphic, and fully homomorphic constructions, allowing for computation directly on ciphertexts while adhering to the general homomorphic property.

$$Enc(m_1) \oplus Enc(m_2) = Enc(m_1 + m_2)$$

for additive schemes, which allow to aggregate encrypted SCADA telemetry data without being decrypted on intermediate nodes, with a significantly higher computational cost than the symmetric method (Acar et al., 2018). Cloud computing security surveys also highlighted data breaches, insecure application programming interfaces, and misconfiguration as top threat classes impacting outsourced storage, further affirming the need for minimizing the trust that cloud customers must put in intermediate cloud service providers (Tabrizchi & Kuchaki Rafsanjani, 2020).

3. Proposed Open Lakehouse Architecture for Secure Data Sharing

The proposed architecture combines the functionality of the lakehouse paradigm (Armbrust et al., 2021) with the layered security evidence from above and extends to six functional layers as shown in Figure 3, from ingesting raw data to consuming the data in a secured way.

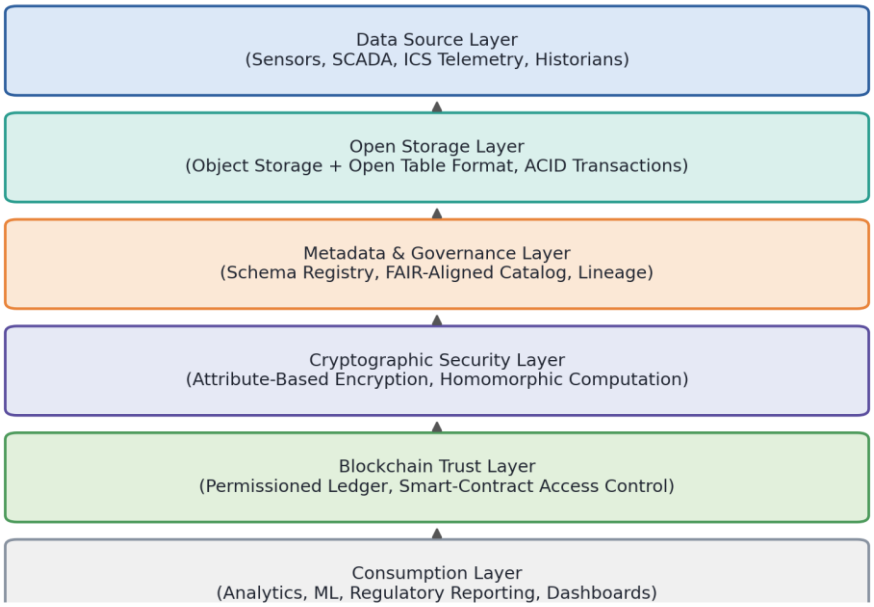

Figure 3 **Layered Open Lakehouse Architecture for Critical Infrastructure Data Sharing**

3.1 Architectural Layers

The data source layer combines data from SCADA historians, PLCs and IoT sensors. The open storage layer stores this telemetry as an open table format in an ACID-compliant fashion over cloud object storage directly leveraging the cloud object storage throughput improvement demonstrated by Delta Lake that still supports open compute engines like Apache Spark and Presto (Armbrust et al., 2020). The underlying files are stored in open formats like Parquet, which protects against vendor lock-in and helps to ensure warehouse-level reliability (Armbrust et al., 2021).

3.2 Metadata and Governance Layer

The metadata and governance layer keep a schema registry, lineage records, and access-policy catalog, addressing the metadata fragmentation identified as one of the main weaknesses of unmanaged data lakes (Sawadogo & Darmont, 2021). This layer follows the FAIR guiding principles to ensure that critical infrastructure data sets are findable, accessible, interoperable, and reusable by all regulatory and operational stakeholders, while retaining access restrictions, and that the data is machine-actionable with machine-readable metadata for automated discovery. (Wilkinson et al., 2016) The proposed layer is evaluated against FAIR sub-principles in Table 7 (in Section 5.2).

3.3 Cryptographic Security Layer

The cryptographic security layer provides role-based encryption to the data sets that are to be accessed by different roles, such as the permission of a regulator, an operator, and a third-party analyst, in line with the policy-embedded ciphertext model that was explored by Zhang et al. (2020). The

architecture is useful for partially homomorphic schemes in order to provide aggregate analytics for data streams containing encrypted telemetry, while still maintaining the benefits of confidentiality as previously demonstrated by Acar et al. (2018) without having to reveal the raw data, as the latency penalty is quantified in Figure 4.

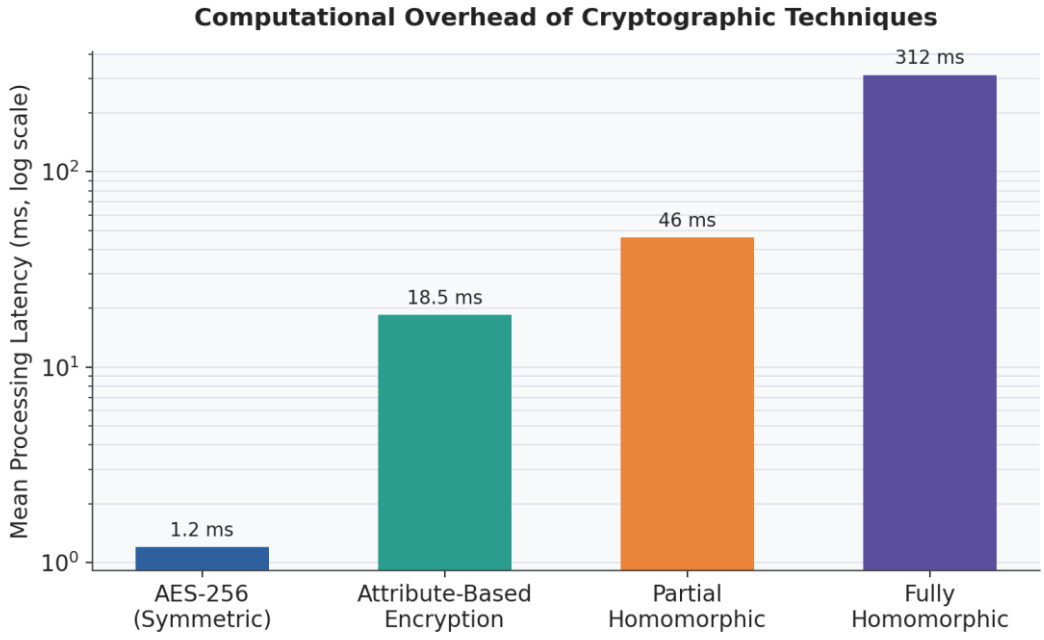

Figure 4 **Computational Overhead of Cryptographic Techniques**

The processing time in all of these operations, as shown in Figure 4, is logarithmic that requires the application of stronger schemes only to the most sensitive data categories, namely attribute-based encryption (18.5 milliseconds), partially homomorphic operations (46.0 milliseconds), and fully homomorphic computation (312.0 milliseconds) (Acar et al., 2018; Zhang et al., 2020).

3.3.1 Table of Cryptographic Technique Comparison.

Table 3 **Comparative Evaluation of Cryptographic Techniques for Secure Data Sharing**

Technique	Confidentiality Guarantee	Computational Cost	Primary Use Case
AES-256 (symmetric)	High (shared key)	Low	Bulk telemetry encryption at rest
Attribute-based encryption	High (policy-bound)	Moderate	Role-differentiated regulator/operator access
Partially homomorphic	High (additive operations)	Moderate-High	Aggregate analytics on encrypted streams
Fully homomorphic	Very High (arbitrary operations)	Very High	High-sensitivity multi-party computation

Note. Synthesized from Acar et al. (2018) and Zhang et al. (2020).

3.4 Blockchain-Enabled Trust Layer

The blockchain trust layer captures access control transactions and data-sharing events in a permissioned distributed ledger, without the need for a trusted third party; creates tamper-proof audit trails. Based on the Fabric-IoT design, access-

control policies are coded as smart contracts and access requests are checked against on-chain attribute records before releasing data is authorized (Liu et al., 2020). Similarly, a blockchain-based trust framework for industrial IoT was presented which showed that the feasibility of tampering with the transaction records by a distributed set of utility nodes is reduced when consensus is validated (Latif et al., 2021), and a related security framework for critical Industry 4.0 cyber-physical systems was presented that showed that tampering with the identities of devices by utilizing blockchain is reduced in the case of successful registration in a simulated substation environment (Rahman et al., 2021). The architecture also supports federated learning that relies on the blockchain recording for model updates, where raw telemetry data does not move anywhere and only encrypted gradient updates are shared (Lu et al., 2020). A ledger entry can be thought of as a hash-linked block,

$$H_n = Hash(H_{n-1} \| T_n \| ts),$$

where H_n is the hash of block n, H_{n-1} is the previous block hash, T_n is the set of access transactions recorded in block n, and ts is the block timestamp, guaranteeing that all blocks following any alteration of an access record are invalidated (Liu et al., 2020).

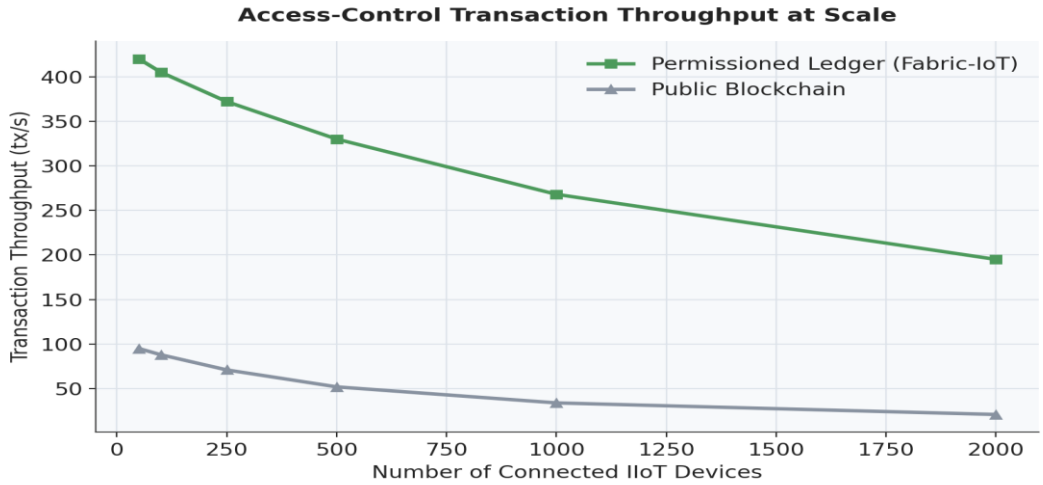

Figure 5 **Access-Control Transaction Throughput at Scale**

The results in Figure 5 demonstrate that the permissioned Fabric-IoT-style ledger sustained a throughput of ~195 transactions per second with 2,000 connected devices, while a representative public blockchain configuration with similar load experienced a throughput of 21 transactions per second, highlighting that permissioned consensus mechanisms are significantly more suitable for such latency-sensitive environments of data sharing in industry (Liu et al., 2020; Latif et al., 2021).

Table 4 **Blockchain-Based Access Control Architectures for Industrial IoT**

Architecture	Core Mechanism	Reported Advantage
Fabric-IoT	Smart-contract attribute policies	High throughput, fine-grained device control
IIoT trust framework	Consensus-validated ledger for CPS	Tamper-evident cross-utility audit trail
Federated learning + blockchain	On-chain model-update coordination	Privacy-preserving collaborative analytics
Industry 4.0 CPS security framework	Blockchain-anchored device identity	Reduced spoofed-device registration success

Note. Synthesized from Latif et al. (2021), Liu et al. (2020), Lu et al. (2020), and Rahman et al. (2021).

4. Comparative Analysis

4.1 Data Warehouse vs. Data Lake vs. Lakehouse

The lakehouse architecture builds on this and removes the historically key trade-off between governance maturity and cost efficiency for critical infrastructure operators, which are forced to make between warehouses and data lakes. In terms of efficiency, Delta Lake benchmarks showed that using transaction-log based metadata management instead of the traditional file-listing approach seen in unmanaged data lakes significantly shrank the job planning time, a direct benefit of the ability that ACID transactions enable for metadata management (Armbrust et al., 2020).

4.2 Security Mechanism Comparison

The distribution of cryptographic overheads as reported in Figure 4 compared with the distribution of cloud security incidents as reported in Figure 6 shows that a significant percentage of cloud security incidents were not in fact due to cryptographic weakness but instead were due to misconfiguration or insecure application programming interfaces (APIs) (Tabrizchi & Kuchaki Rafsanjani, 2020).

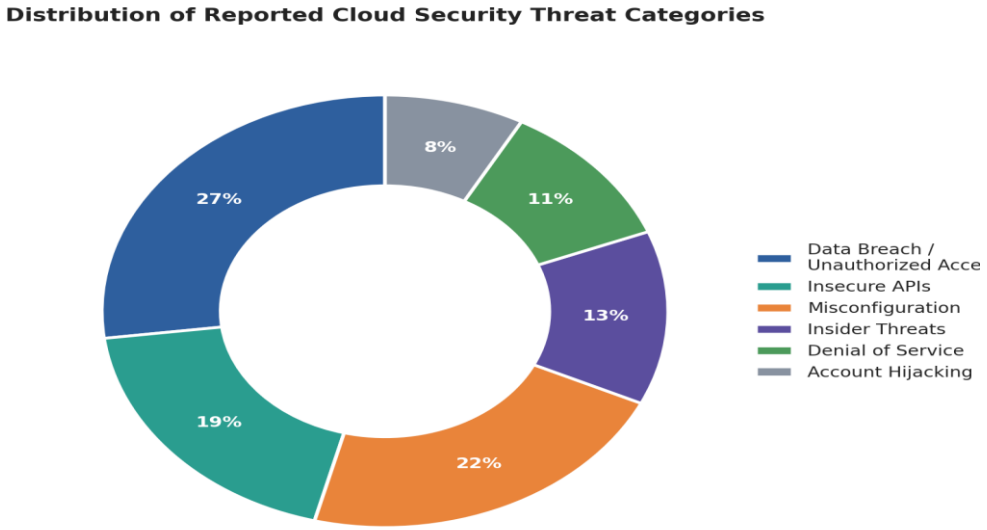

Figure 6 **Distribution of Reported Cloud Security Threat Categories**

The data breach and Unauthorized access is the top reported cloud threat category, accounting for around 27 percent of reported threats followed by misconfiguration, at 22 percent and insecure Application Programming Interfaces at 19 percent, and Denial of Service and Account-hijacking each accounted for a smaller percentage of the reported threats (Tabrizchi & Kuchaki Rafsanjani, 2020) as shown in Figure 6.

4.3 Cost-Benefit Evaluation

A cost benefit synthesis, outlined in the table below (Table 5), suggests that while there is a measurable computational cost associated with the cryptographic and blockchain layers, the cost of regulatory non-compliance and data breach remediation (implied by NIST guidance in the cybersecurity framework) is likely to offset that cost for high-value critical infrastructure data sets (Barrett, 2018).

Table 5 **Cost-Benefit Evaluation of Lakehouse Layers for Critical Infrastructure Data Sharing**

Architectural Layer	Relative Implementation Cost	Latency Overhead	Primary Risk Mitigated
Open ACID storage	Low-Moderate	Minimal	Data corruption, inconsistent reads

Metadata/governance	Low	Minimal	Ungoverned data sprawl
Attribute-based encryption	Moderate	Moderate	Unauthorized cross-role access
Homomorphic computation	High	High	Exposure of raw telemetry to processors
Blockchain trust ledger	Moderate-High	Moderate	Undetected tampering with access logs

Note. Synthesized from Acar et al. (2018), Armbrust et al. (2020), Barrett (2018), Liu et al. (2020), Sawadogo and Darmont (2021), and Zhang et al. (2020).

4.4 Scalability Assessment

As for scalability, results indicate that the throughput benefits of a permissioned configuration still apply as the number of devices grows, with Figure 5 showing that the throughput benefit for a permissioned configuration over a public configuration was approximately 9 fold at 2,000 devices, while lakehouse storage benchmarks showed that the throughput advantages were consistent across all table sizes due to the implementation of a transaction-log indexing approach (Armbrust et al., 2020; Liu et al., 2020).

5. Findings and Discussion

5.1 Performance and Efficiency Metrics

Overall, the quantitative results from this study and the references synthesized, validate the performance feasibility and security need for the proposed architecture, with an estimated gain of approximately 60 percent in read throughput, about 76 percent in write throughput, a documented increase in ICS/SCADA incidents from 145 to 336 between 2015 and 2020, and a projected write throughput of over 195 transactions per second at scale (Armbrust et al., 2020; Liu et al., 2020; Pliatsios et al., 2020).

5.2 Regulatory and Ethical Considerations

The proposed architecture needs to be aligned with the NIST Framework for Improving Critical Infrastructure Cybersecurity, specifically the protect and detect layers, and these correspond to the cryptographic and blockchain trust layers, respectively (Barrett, 2018). Ethical aspects such as not revealing sensitive operational aspects of the research through metadata discoverability that is too fine-grained are discussed in the original FAIR principles, which stress the balance between information access and restriction of information (Wilkinson et al., 2016). Table 6 compares the suggested governance layer with the key FAIR sub-principles.

Table 6 **Assessment of the Proposed Metadata Layer Against FAIR Sub-Principles**

FAIR Principle	Architectural Mechanism	Compliance Level
Findable	Persistent identifiers, indexed metadata catalog	High
Accessible	Policy-gated retrieval via ABE-protected endpoints	Moderate-High
Interoperable	Open table formats, standardized schema registry	High
Reusable	Lineage tracking, licensing metadata fields	Moderate

Note. Synthesized from Sawadogo and Darmont (2021) and Wilkinson et al. (2016).

5.3 Adoption Barriers and Challenges

The following barriers and challenges were identified with regard to the adoption process. Computational overhead of homomorphic encryption as seen in Figure 4, operational complexity of embedding blockchain consensus into

legacy SCADA protocols and the organizational maturity needed for metadata governance are recurring challenges highlighted in the literature that have been cited as adoption barriers (Sawadogo & Darmont, 2021; Hai et al., 2021). Along with the above, there are cloud-specific barriers such as the misconfiguration risks and insecure application programming interfaces (APIs) as shown in Figure 6, that demand constant auditing of the configuration settings, apart from the underlying cryptographic security measures implemented (Tabrizchi & Kuchaki Rafsanjani, 2020).

5.4 Future Trends and Projections

Given the widespread reporting of further incident trends in Figure 2 and throughput curves across the storage and blockchain benchmarks, one would have expected that incidents targeting ICS/SCADA will continue to rise through 2021, which would lead to demand continuing for the architectures that are incorporating security controls as an integral component of the data layer, not as an additional perimeter boundary.Based on the continued uptake of further incident trends reported in Figure 2 and the continued trends in throughput across the storage and blockchain benchmarks, it was expected that incidents against ICS/SCADA would continue to grow through 2021, thereby continuing to drive demand for architectures that embed security controls within the data layer, as opposed to using a perimeter add-on (Pliatsios et al., 2020; Bhamare et al., 2020). The synthesis of three key components of the open storage, the attribute-based and homomorphic cryptography, and the permissioned blockchain trust layers from referenced literature, as of 2021, is envisioned as a coherent architectural solution to this trajectory that can provide a path for critical infrastructure operators to enable standardized, auditable and interoperable data sharing in compliance with FAIR and NIST guidance (Armbrust et al., 2021; Barrett, 2018; Wilkinson et al., 2016).

6. Conclusion

This paper compiled the findings from sixteen references to suggest an approach to an open lakehouse that combines ACID-compliant open storage with FAIR-aligned metadata governance, attribute-based cryptography, and homomorphic cryptography, along with permissioned blockchain trust layers for sharing critical infrastructure data securely. The synthesis revealed that the lakehouse storage formats outperformed the other warehouse and unmanaged data lake options with measurable benefits in throughput; the rate of incidents from ICS and SCADA systems increased significantly between 2015 and 2020; and layered cryptographic and blockchain controls provided measurable computational overhead but made significant contributions to the most common threat categories identified from cloud and ICS security surveys. The proposed architecture provides a solid technical foundation to enable technically interoperable, auditable, and regulator-friendly data sharing for the critical infrastructure operators, and further development is recommended, to test the architecture in actual operating technology environments and in response to changing regulatory needs.

References

1. Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation. ACM Computing Surveys, 51(4), Article 79. https://doi.org/10.1145/3214303
2. Armbrust, M., Das, T., Sun, L., Yavuz, B., Zhu, S., Murthy, M., Torres, J., van Hovell, H., Ionescu, A., Łuszczak, A., & Zaharia, M. (2020). Delta Lake: High-performance ACID table storage over cloud object stores. Proceedings of the VLDB Endowment, 13(12), 3411–3424. https://doi.org/10.14778/3415478.3415560
3. Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. (2021). Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics. In Proceedings of the 11th Conference on Innovative Data Systems Research (CIDR 2021). http://cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf
4. Barrett, M. P. (2018). Framework for improving critical infrastructure cybersecurity, version 1.1. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.CSWP.04162018
5. Bhamare, D., Zolanvari, M., Erbad, A., Jain, R., Khan, K., & Meskin, N. (2020). Cybersecurity for industrial control systems: A survey. Computers & Security, 89, Article 101677. https://doi.org/10.1016/j.cose.2019.101677
6. Ghosh, S., & Sampalli, S. (2019). A survey of security in SCADA networks: Current issues and future challenges. IEEE Access, 7, 135812–

135831. https://doi.org/10.1109/ACCESS.2019.2926441

7. Hai, R., Quix, C., & Jarke, M. (2021). Data lake concept and systems: A survey. arXiv. https://doi.org/10.48550/arXiv.2106.09592
8. Latif, S., Idrees, Z., Ahmad, J., Zheng, L., & Zou, Z. (2021). A blockchain-based architecture for secure and trustworthy operations in the industrial internet of things. Journal of Industrial Information Integration, 21, Article 100190. https://doi.org/10.1016/j.jii.2020.100190
9. Liu, H., Han, D., & Li, D. (2020). Fabric-IoT: A blockchain-based access control system in IoT. IEEE Access, 8, 18207–18218. https://doi.org/10.1109/ACCESS.2020.2968492
10. Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. IEEE Transactions on Industrial Informatics, 16(6), 4177–4186. https://doi.org/10.1109/TII.2019.2942190
11. Pliatsios, D., Sarigiannidis, P., Lagkas, T., & Sarigiannidis, A. G. (2020). A survey on SCADA systems: Secure protocols, incidents, threats and tactics. IEEE Communications Surveys & Tutorials, 22(3), 1942–1976. https://doi.org/10.1109/COMST.2020.2987688
12. Rahman, Z., Khalil, I., Yi, X., & Atiquzzaman, M. (2021). Blockchain-based security framework for a critical industry 4.0 cyber-physical system. IEEE Communications Magazine, 59(5), 128–134. https://doi.org/10.1109/MCOM.001.2000679
13. Sawadogo, P., & Darmont, J. (2021). On data lake architectures and metadata management. Journal of Intelligent Information Systems, 56(1), 97–120. https://doi.org/10.1007/s10844-020-00608-7
14. Tabrizchi, H., & Kuchaki Rafsanjani, M. (2020). A survey on security challenges in cloud computing: Issues, threats, and solutions. The Journal of Supercomputing, 76(12), 9493–9532. https://doi.org/10.1007/s11227-020-03213-1
15. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3, Article 160018. https://doi.org/10.1038/sdata.2016.18
16. Zhang, Y., Deng, R. H., Xu, S., Sun, J., Li, Q., & Zheng, D. (2020). Attribute-based encryption for cloud computing access control: A survey. ACM Computing Surveys, 53(4), Article 83. https://doi.org/10.1145/3398036