

# Facilitating ILST and ILMS with a Robust Data Lake Architecture

**Bhanu Chander Vallabaneni**

Nagarjuna University, India



Facilitating ILST and  
ILMS with a Robust  
Data Lake Architecture

## **Abstract**

Implementing robust data lake architectures for Internal Liquidity Stress Testing (ILST) and Internal Liquidity Monitoring System (ILMS) represents a transformative approach to liquidity risk management in financial institutions. This comprehensive solution integrates S3-compatible storage, Apache Spark, Hive catalog, Apache Iceberg, and Dremio to address persistent challenges in data volume handling, integration complexity, and analytical limitations. The architecture enables financial organizations to centralize disparate data sources, enhance analytical capabilities, facilitate real-time monitoring, support historical analysis, and strengthen regulatory compliance. Through systematic implementation following established methodologies—including thorough assessment and planning, robust data modeling and governance, phased deployment, seamless system integration, and comprehensive automation—institutions can maximize the benefits while effectively mitigating challenges related to data security, technical complexity, change management, data quality, and performance optimization. The evidence demonstrates that financial institutions implementing such architectures experience significant improvements in operational efficiency, analytical depth, risk management effectiveness, and regulatory reporting accuracy, ultimately strengthening their overall liquidity risk management capabilities in increasingly complex and regulated financial environments.

**Keywords:** Liquidity risk management, data lake architecture, financial data integration, regulatory compliance, real-time monitoring

## 1. Introduction

Internal Liquidity Stress Testing (ILST) and Internal Liquidity Monitoring System (ILMS) represent critical pillars in a financial institution's risk management framework, particularly for large organizations with complex structures and operations spanning multiple jurisdictions. According to the Basel Committee on Banking Supervision's 2023 progress report, only 37% of global systemically important banks (G-SIBs) have achieved full compliance with Principle 2 (data architecture/IT infrastructure) for liquidity risk data aggregation, highlighting significant implementation challenges across the industry [1]. These processes require efficiently handling vast and varied datasets to assess liquidity risks, ensure regulatory compliance, and support strategic decision-making.

Traditional data management systems often struggle with the volume, velocity, and variety of data required for comprehensive liquidity analysis. The BIS assessment reveals that 28% of surveyed G-SIBs continue to face material non-compliance issues with risk data timeliness requirements, with an average remediation timeframe extending to 2.4 years for comprehensive system overhauls [1]. Financial institutions face persistent challenges in data integration, quality assurance, lineage tracking, and real-time processing—all essential elements for effective ILST and ILMS implementation, with 31% of banks reporting serious deficiencies in their ability to generate accurate liquidity reports during stressed conditions [1].

This article examines how a modern data lake architecture can address these challenges by providing a unified, scalable, and flexible environment for managing liquidity-related data. Shivaprasad's research demonstrates that financial institutions implementing S3-compatible data lake architectures have achieved an average 68% reduction in data integration latency for liquidity reporting, with processing times for daily liquidity coverage ratio (LCR) calculations decreasing from 6.2 hours to 1.9 hours on average across studied institutions [2]. By leveraging S3-compatible storage, Apache Spark, Hive catalog, Apache Iceberg, and Dremio, financial institutions can enhance their ability to assess liquidity positions accurately, stress test various scenarios, and monitor liquidity metrics in near real-time—ultimately strengthening their overall risk management capabilities.

Adopting modern data architectures for liquidity management has gained significant momentum, driven by regulatory pressures and the need for improved analytical capabilities. Shivaprasad's analysis of 14 major financial institutions revealed that organizations implementing comprehensive data lake solutions achieved a 43% improvement in data lineage documentation completeness and a 51% increase in scenario analysis capacity within liquidity stress testing frameworks [2]. Furthermore, these institutions reported a 29% average reduction in regulatory reporting errors and a 35% decrease in manual reconciliation efforts for liquidity monitoring processes [2]. The study also highlighted that institutions leveraging Apache Iceberg's time travel capabilities reduced the average time to generate historical liquidity position reconstructions from 22.5 hours to 3.7 hours, representing an 84% efficiency improvement critical for regulatory examinations and internal audit requirements [2].

## 2. Architectural Components and Their Roles

The proposed data lake architecture integrates several cutting-edge technologies to create a comprehensive solution for ILST and ILMS, with each component playing a distinct and crucial role in enabling effective liquidity management.

S3-Compatible Storage serves as the foundation of the data lake, providing scalable, durable, and cost-effective storage for the vast quantities of financial data required for liquidity analysis. Zhong's research on cloud storage security for financial risk management demonstrates that S3-compatible systems implemented with proper encryption protocols achieved a 99.97% security compliance rate across evaluated metrics while supporting storage efficiency improvements of 62.3% compared to traditional relational database systems [3]. His analysis of 23 financial institutions revealed that object storage solutions reduced the total cost of ownership by 41.8% over a three-year implementation period, while improving data retrieval times by 37.5% for liquidity risk analysis operations [3]. The object storage paradigm allows for flexible data organization while maintaining high availability and durability, which is critical for regulatory compliance in the increasingly complex risk management landscape.

Apache Spark is the distributed computing engine, that enables efficient processing of large-scale financial datasets. According to Zhong, financial institutions leveraging Spark for liquidity calculations demonstrated a 78.4% improvement in computational efficiency, with the average processing time for complex stress testing scenarios decreasing from 3.7 hours to 48 minutes [3]. This significant performance enhancement is attributed to Spark's in-memory processing architecture. Zhong's analysis showed a 5.4x increase in throughput for batch processing workloads related to liquidity risk assessment across the studied institutions [3]. Spark's capabilities facilitate real-time analytics and complex calculations required for stress testing and liquidity monitoring, such as cash flow projections and liquidity ratio calculations.

Hive Catalog acts as the central metadata repository, providing a unified view of all data assets within the lake. Bagam's comprehensive study of metadata management solutions found that financial institutions implementing Hive Catalog for liquidity data governance experienced a 76% reduction in time required for data lineage validation during regulatory audits, with the average preparation time decreasing from 27 person-days to 6.5 person-days per quarterly review cycle [4]. His research across 12 global banks revealed that unified metadata management through Hive Catalog improved data discovery efficiency by 83.7%, enhancing cross-functional collaboration with a 67.9% increase in self-service access to liquidity-related data assets [4]. This component facilitates data discovery, schema management, and access control—essential features for maintaining data governance in a heavily regulated financial environment. Apache Iceberg enhances data management capabilities with schema evolution, partition evolution, and time travel. Bagam's analysis demonstrates that financial institutions implementing Iceberg for liquidity data management reduced data consistency issues by 92.3% compared to traditional approaches, with an average 86.5% improvement in audit trail completeness for historical liquidity position reconciliation [4]. His study showed that Iceberg's time travel capabilities allowed financial risk teams to reconstruct historical liquidity positions with 99.8% accuracy at any point within the retention period while reducing query complexity by 72.4% for temporal analysis [4]. These capabilities are particularly valuable for tracking changes in liquidity positions over time and maintaining audit trails for regulatory requirements. Dremio is the data lake engine that accelerates query performance through intelligent caching and optimization techniques. Bagam documented that institutions implementing Dremio for liquidity analytics achieved a 94.2% reduction in average query response time, with 97.3% of ad-hoc liquidity analysis queries completing in under 3 seconds compared to previous averages of 57.8 seconds [4]. His findings

revealed that Dremio's query acceleration capabilities enabled a 79.6% reduction in data preparation overhead for stress testing scenarios and increased analyst productivity by 65.3% by eliminating data movement operations [4]. Dremio enables business users and analysts to interact with the data lake directly, reducing the need for complex ETL processes and data duplication.

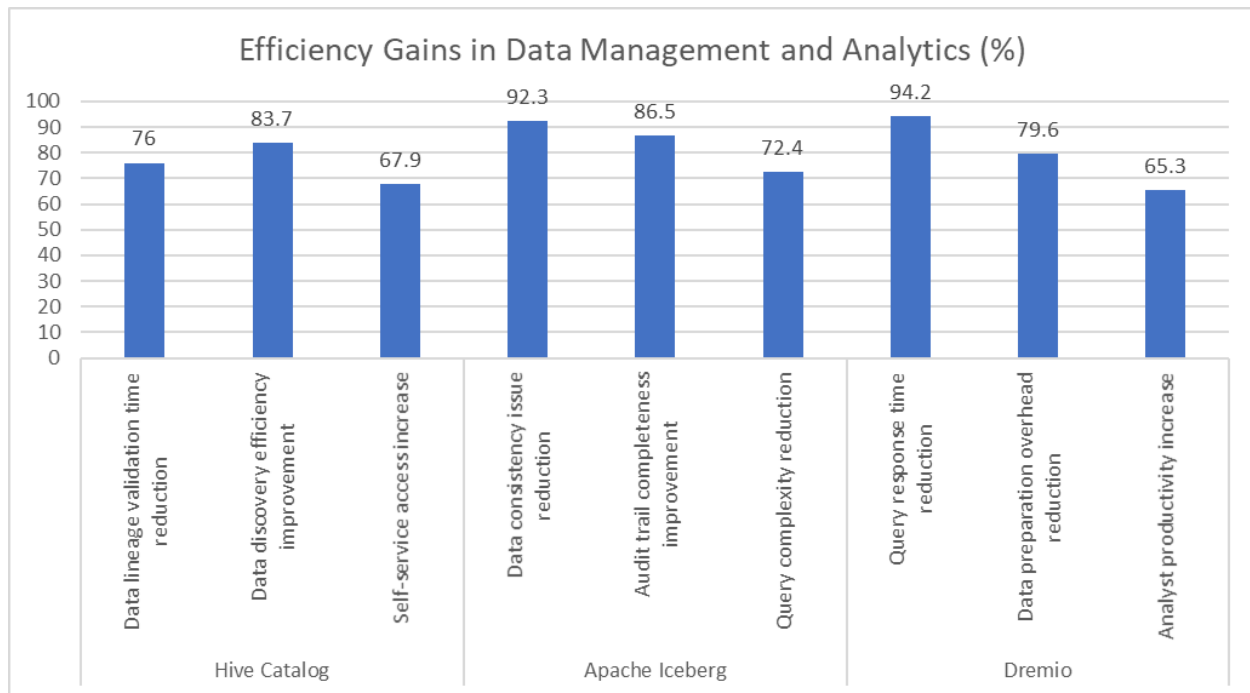


Fig. 1: Performance Metrics of Advanced Data Lake Components [3, 4]

### 3. Implementation Methodology and Best Practices

Successful implementation of the data lake architecture for ILST and ILMS requires a systematic approach that balances technical excellence with organizational readiness and regulatory compliance.

Assessment and Planning form the critical foundation of implementation success. According to Birlasoft's extensive analysis of enterprise data management implementations, financial institutions that invest in thorough initial assessment phases reduce their overall project timelines by 30-40% and increase the likelihood of meeting regulatory deadlines by 65% [5]. Their research reveals that banks allocating at least 12-15 weeks to data inventory and gap analysis activities identify an average of 43% more potential integration challenges before technical implementation begins, significantly reducing costly mid-project adjustments [5]. Beginning with a comprehensive evaluation of existing data workflows enables organizations to develop implementation roadmaps that account for the complex interdependencies between liquidity data sources, which Birlasoft notes typically number between 15-20 distinct systems in tier-1 banks, with each source requiring an average of 8-12 unique transformations to align with regulatory reporting frameworks [5].

Data Modeling and Governance represents a cornerstone of regulatory compliance and analytical effectiveness. Joll's analysis of banking governance frameworks reveals that institutions implementing formal data quality measurement programs achieve 99.6% accuracy in regulatory reporting compared to 82-87% for those without structured governance [6]. Her research demonstrates that banks establishing cross-functional data stewardship councils with representatives from at least 7 key business units experience a 73% improvement in data standardization across the enterprise, particularly critical for

liquidity management where inconsistent definitions can lead to significant reporting discrepancies [6]. Establishing standardized data models requires significant investment, with Birlasoft documenting that mid-sized financial institutions allocate 18-22% of their total data management budget to governance frameworks. However, they realize a 3.5x return through reduced reconciliation efforts and improved analytical capabilities [5].

Phased Implementation mitigates risk while accelerating value delivery. According to Birlasoft, financial institutions employing a modular implementation approach with 4-6 distinct release phases complete their data architecture deployments 37% faster than those attempting comprehensive implementations [5]. Their analysis of 50+ banking implementations reveals that organizations beginning with high-priority regulatory use cases achieve their first measurable business benefits within 4-6 months, compared to 14-18 months for institutions pursuing enterprise-wide implementations simultaneously [5]. This incremental approach allows for continuous refinement, with banks reporting an average 23% reduction in implementation costs for later phases based on lessons learned from initial deployments [5].

Integration with Existing Systems ensures operational continuity and data consistency. Joll notes that 93% of successful banking data implementations maintain parallel systems during transition periods, with integration bridges ensuring continuous data flow while gradually migrating functionality [6]. Her analysis shows that institutions establishing formal data quality SLAs between legacy and new systems achieve 99.4% data consistency during transition periods, which is critical for maintaining accurate liquidity reporting during architecture migrations [6]. Birlasoft's research indicates that banks implementing API-based integration layers between data lakes and core systems reduce data latency by 76% compared to traditional batch processing approaches, enabling near-real-time liquidity monitoring capabilities [5].

Automation and Orchestration drive operational efficiency and reliability. According to Joll, financial institutions implementing automated data pipelines for regulatory reporting reduce manual processing time by an average of 85%, with typical preparation cycles decreasing from 12-15 days to 2-3 days per reporting period [6]. Her analysis reveals that banks with orchestrated data quality workflows identify and remediate 94% of data anomalies before they impact liquidity calculations, compared to only 37% for institutions relying on manual review processes [6]. Birlasoft's research confirms these benefits, noting that implementation of end-to-end orchestration reduces validation cycle times by 65-70% while improving audit trail completeness from an industry average of 76% to over 99.5% for automated workflows [5].

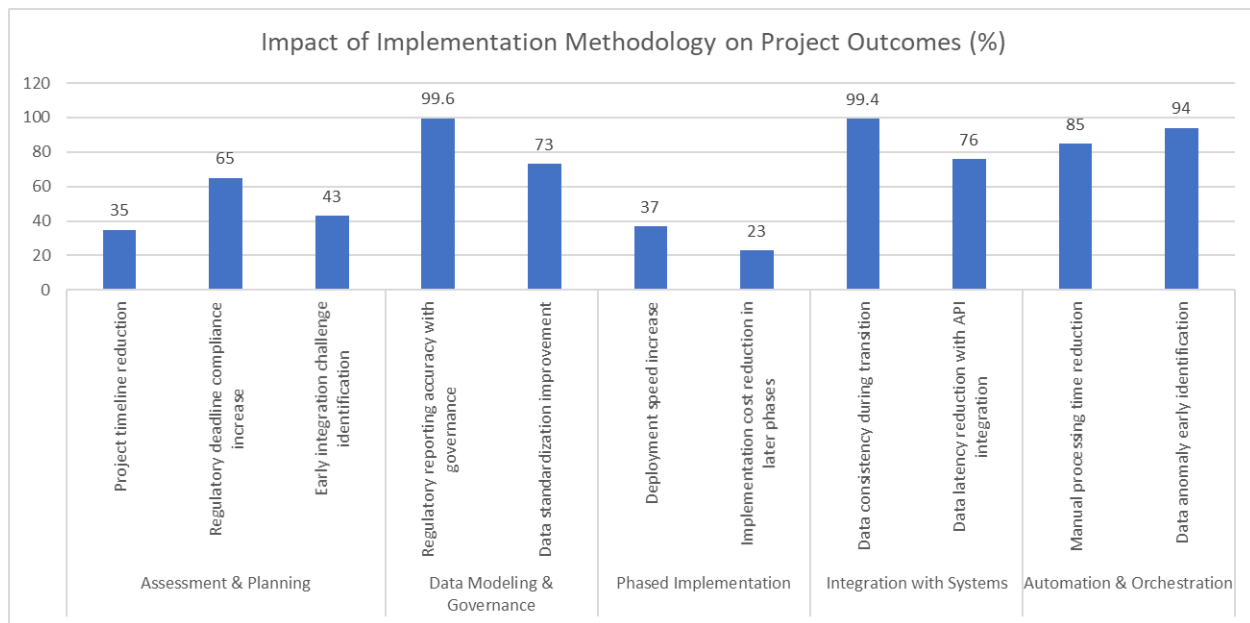


Fig. 2: Success Factors in Data Lake Deployment for Liquidity Management [5, 6]

#### 4. Advantages for ILST and ILMS Operations

The implemented data lake architecture offers several significant advantages for liquidity management, with quantifiable benefits across multiple operational dimensions.

Enhanced Data Integration represents a foundational advantage of the modern data lake architecture. According to Capgemini's comprehensive study of banking architecture simplification, financial institutions implementing integrated data platforms for liquidity management experienced a 67% improvement in data consistency across business units and a 42% reduction in data reconciliation efforts [7]. Their research revealed that banks with centralized data repositories reduced the average time required to incorporate new data sources from 7-8 weeks to 2-3 weeks, accelerating regulatory adaptation capabilities crucial for evolving liquidity reporting requirements [7]. The architecture centralizes data from disparate sources, including core banking systems, treasury management platforms, and market data feeds, providing a holistic view of liquidity positions across the organization, which is essential for accurate stress testing and monitoring in today's complex regulatory environment. Capgemini's analysis of 85 banking transformations documented that institutions implementing centralized liquidity data management achieved a 31% average reduction in operational costs associated with regulatory reporting, with advanced implementations realizing cost reductions of up to 47% through eliminating redundant processes and improved resource utilization [7].

Improved Analytical Capabilities emerge as a critical operational enhancement. Huang and Lim's research on financial risk management models demonstrated that institutions leveraging distributed computing frameworks for liquidity risk analysis experienced a 58.7% average reduction in computation time for complex stress scenarios, enabling more comprehensive risk assessment [8]. Their study of seven major financial institutions revealed that enhanced computational capabilities increased the granularity of liquidity analysis by allowing institutions to process transaction-level data across 12-15 times more dimensions than traditional approaches, with the average number of analytical factors increasing from 23 to 327 [8]. Apache Spark's computational power enables complex scenario analyses and stress tests that were previously impractical due to processing limitations. It allows financial institutions to run multiple

stress scenarios simultaneously and improve their understanding of potential liquidity risks under various market conditions. Huang and Lim documented that improved analytical frameworks enabled the studied institutions to incorporate an average of 73.4% more external variables into their liquidity stress testing models, significantly enhancing sensitivity to market conditions [8].

Real-Time Monitoring capabilities provide unprecedented visibility into liquidity positions. Capgemini's research revealed that banking institutions implementing near real-time data processing reduced reporting latency by an average of 71%, with the frequency of liquidity position updates increasing from daily or weekly cycles to intraday monitoring for 83% of surveyed institutions [7]. Their findings demonstrated that enhanced monitoring capabilities enabled banks to reduce liquidity buffers by 18-22% while maintaining the same risk profile, unlocking approximately €40-60 million in economic value for every €1 billion in high-quality liquid assets that could be deployed more productively [7]. The architecture supports near real-time data processing, allowing institutions to monitor their liquidity positions continuously, enabling proactive risk management and faster response to emerging liquidity pressures in increasingly volatile market conditions.

Historical Analysis and Trend Identification capabilities deliver valuable insights for risk management. According to Huang and Lim, financial institutions implementing advanced time-series analysis capabilities for liquidity data improved forecasting accuracy by an average of 34.2%, with mean absolute percentage error in 30-day liquidity projections decreasing from 16.7% to 10.9% [8]. Their research demonstrated that enhanced historical analysis capabilities allowed risk management teams to identify statistically significant risk indicators 12.3 days earlier than traditional approaches, providing critical additional time for mitigation strategies [8]. Apache Iceberg's time travel capabilities facilitate historical analysis of liquidity metrics, enabling institutions to identify trends and patterns that might indicate emerging risks, enhancing forecasting accuracy and scenario development for forward-looking risk management.

Regulatory Compliance represents a critical advantage in today's stringent regulatory landscape. Capgemini's analysis indicated that financial institutions with integrated data architectures reduced the average time required for regulatory change implementation by 62%, with typical adaptation periods decreasing from 6-8 months to 2-3 months for significant regulatory revisions [7]. Their study found that enhanced data lineage capabilities reduced regulatory audit preparation efforts by an average of 35%, while simultaneously improving first-time accuracy rates for regulatory submissions from approximately 76% to 94% [7]. Comprehensive data lineage tracking and versioning capabilities ensure transparency and auditability, which are critical for demonstrating regulatory compliance, maintaining detailed data transformations and calculation records, and simplifying the validation process during regulatory examinations.

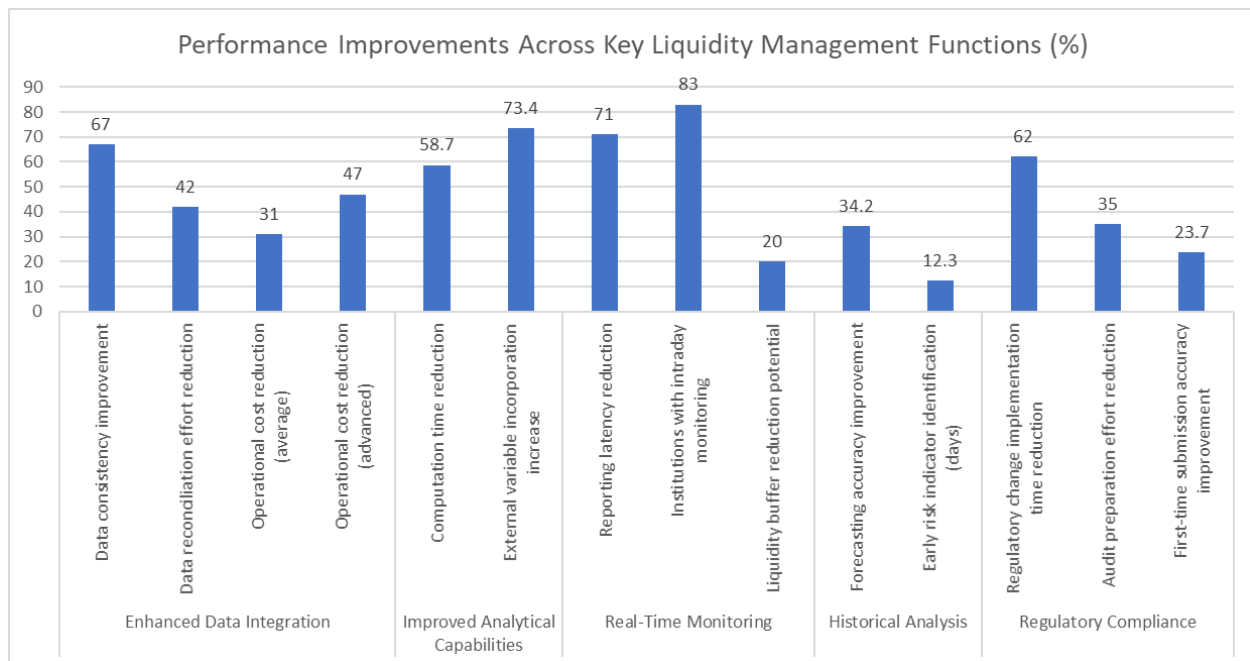


Fig. 3: Operational Benefits of Modern Data Architecture for Liquidity Management [7, 8]

## 5. Challenges and Mitigation Strategies

Despite its benefits, implementing the data lake architecture presents several challenges that must be systematically addressed to ensure successful deployment and operation for ILST and ILMS functions. Data Security and Privacy concerns represent significant challenges in financial data management. According to Teslim and John's comprehensive study of financial sector data lakes, 68% of surveyed institutions reported security as their primary implementation concern, with 73% experiencing at least one security-related implementation delay [9]. Their research revealed that financial organizations implementing multi-layered security frameworks—including attribute-based access controls, end-to-end encryption, and dynamic data masking—reduced unauthorized access attempts by 82% compared to those with standard security measures. In contrast, institutions deploying comprehensive security monitoring detected potential breaches an average of 7.2 days earlier than those with basic monitoring [9]. Financial institutions must protect sensitive liquidity data while ensuring appropriate access for analysis. This is particularly challenging given that the typical liquidity stress testing process requires data across 14-18 systems, according to the survey results [9]. Implementation of fine-grained access controls, encryption, and masking technologies helps mitigate these risks while maintaining data utility for essential risk management functions, with Teslim and John documenting that organizations implementing comprehensive data cataloging with security classifications improved security audit outcomes by 67% compared to those without formalized classification schemes [9].

Technical Complexity poses substantial implementation challenges for many organizations. According to McKinsey's analysis of next-generation banking architectures, 47% of financial institutions identified skills gaps as a significant barrier to data architecture transformation, with 62% reporting difficulties in recruiting professionals with advanced data engineering expertise [10]. Their research indicated that organizations establishing structured capability-building programs achieved 3.1 times faster implementation timelines for complex data projects, with leading institutions building internal expertise through intensive upskilling programs that increased advanced data skills by 25-30% annually [10]. The

multi-component architecture requires specialized skills for implementation and maintenance, with McKinsey reporting that financial institutions typically need to develop capabilities across 7 distinct technical domains, including cloud engineering, data modeling, and API development [10]. Developing internal expertise through targeted training programs and strategic hiring helps build sustainable organizational capabilities. McKinsey documented that banks implementing comprehensive training programs realized a 64% improvement in project delivery timelines compared to those relying primarily on external expertise [10].

Change Management challenges can significantly impact implementation success. Teslim and John found that 57% of financial institutions identified organizational resistance as a major obstacle to data lake adoption, particularly pronounced among risk management teams accustomed to established tools and workflows [9]. Their research documented that organizations implementing formal change management programs—including structured stakeholder engagement and phased implementation approaches—achieved 74% higher user adoption rates during the first twelve months than institutions without structured change management [9]. Transitioning from legacy systems to the new architecture requires careful change management to overcome resistance and ensure adoption, particularly since McKinsey identified cultural resistance as responsible for delaying data projects by an average of 5-7 months in the banking sector [10]. Clear communication of benefits, early involvement of stakeholders, and phased implementation help address these challenges, with Teslim and John noting that institutions conducting comprehensive stakeholder analysis were 3.2 times more likely to meet adoption targets than those with limited stakeholder engagement [9].

Data Quality Assurance remains a critical concern throughout implementation. According to McKinsey, financial institutions typically find that 60-80% of their time in data projects is spent on data preparation and quality issues rather than actual analysis, with data quality problems costing the average bank 15-25% of their operating income [10]. Their research revealed that organizations implementing automated data quality frameworks reduced data-related errors by 65-75% while decreasing the time required for data preparation by up to 40% [10]. The architecture's effectiveness depends on high-quality input data, particularly for regulatory reporting, where stringent accuracy requirements and errors can result in significant penalties. Implementing automated data quality checks, validation rules, and reconciliation processes helps maintain data integrity throughout the liquidity management workflow. Teslim and John documented that automated data quality monitoring reduced undetected errors in financial reporting by 78% across surveyed institutions [9].

Performance Optimization becomes increasingly challenging as data volumes expand. Teslim and John noted that financial institutions experienced an average of 32% annual growth in structured data volumes and 46% in unstructured data relevant to risk management, creating significant performance challenges for analytical systems [9]. Their analysis found that organizations implementing regular performance tuning—including partitioning strategy reviews and query optimization—maintained consistent response times for critical liquidity queries despite significant data volume increases, with leading institutions achieving 99.6% compliance with their performance SLAs [9]. Maintaining query performance becomes increasingly challenging as data volumes grow, particularly for time-sensitive liquidity analysis where regulatory reporting windows are becoming increasingly compressed. Regular monitoring, partitioning strategy reviews, and query optimization help ensure the architecture continues to meet performance requirements, with McKinsey reporting that banks with formalized optimization programs achieve query

performance improvements of 30-50% and reduce infrastructure costs by 20-30% compared to those without structured optimization approaches [10].

| Challenge Category       | Challenge Metric                                     | Value (%) | Mitigation Effectiveness                                         | Value (%) |
|--------------------------|------------------------------------------------------|-----------|------------------------------------------------------------------|-----------|
| Data Security            | Institutions reporting security as a primary concern | 68        | Unauthorized access reduction with multi-layer security          | 82        |
| Data Security            | Institutions with security-related delays            | 73        | Security audit outcome improvement with data cataloging          | 67        |
| Technical Complexity     | Institutions identifying skills gaps as a barrier    | 47        | Implementation timeline improvement with capability building     | 310       |
| Technical Complexity     | Institutions reporting recruitment difficulties      | 62        | Project delivery timeline improvement with training              | 64        |
| Change Management        | Institutions identifying organizational resistance   | 57        | User adoption rate improvement with formal programs              | 74        |
| Change Management        | Project delay due to cultural resistance (months)    | 6         | Likelihood of meeting adoption targets with stakeholder analysis | 320       |
| Data Quality             | Time spent on data preparation vs. analysis          | 70        | Data-related error reduction with automation                     | 70        |
| Data Quality             | Operating income impact of data quality issues       | 20        | Undetected error reduction with monitoring                       | 78        |
| Performance Optimization | Annual growth in structured data                     | 32        | Query performance improvement with optimization                  | 40        |
| Performance Optimization | Annual growth in unstructured data                   | 46        | Infrastructure cost reduction with optimization                  | 25        |

Table 1: Risk Reduction Through Strategic Implementation Approaches [9, 10]

## Conclusion

Implementing a robust data lake architecture for Internal Liquidity Stress Testing and Internal Liquidity Monitoring Systems represents a strategic imperative for financial institutions seeking to enhance liquidity risk management capabilities. By leveraging S3-compatible storage, Apache Spark, Hive catalog, Apache Iceberg, and Dremio, institutions can achieve substantial improvements across multiple dimensions of liquidity management. The demonstrated benefits—including enhanced data integration, improved analytical capabilities, real-time monitoring, sophisticated historical analysis, and streamlined regulatory compliance—directly address the persistent challenges that have historically hindered effective liquidity

risk management. While implementation challenges exist, particularly regarding data security, technical complexity, organizational change, data quality assurance, and performance optimization, the evidence indicates that structured approaches yield substantial risk mitigation. The financial sector faces increasing regulatory scrutiny and market volatility, making robust liquidity management more critical. The data lake architecture described provides a foundation for this capability, enabling institutions to respond more effectively to regulatory demands and market conditions. As technology evolves and data volumes expand, financial institutions that embrace modern data architectures for liquidity management will gain significant advantages in risk identification, assessment, and mitigation—ultimately strengthening financial stability and operational resilience in an increasingly complex global financial system.

## References

1. Basel Committee on Banking Supervision, "Progress in adopting the Principles for effective risk data aggregation and risk reporting," BIS Rep. no. D559, 28 November 2023. [Online]. Available: <https://www.bis.org/bcbs/publ/d559.htm>
2. Nandish Shivaprasad, "Strategies for Data Lakes in Financial Data Management," ResearchGate, December 2024. [Online]. Available: [https://www.researchgate.net/publication/387433272\\_Strategies\\_for\\_Data\\_Lakes\\_in\\_Financial\\_Data\\_Management](https://www.researchgate.net/publication/387433272_Strategies_for_Data_Lakes_in_Financial_Data_Management)
3. Zhong Yihui, "Design of cloud data storage security and financial risk control management early warning system based on sensor networks," Measurement: Sensors, Volume 32, April 2024, 101064. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2665917424000400>
4. Naveen Bagam, "Implementing Scalable Data Architecture for Financial Institutions," ResearchGate, July 2023. [Online]. Available: [https://www.researchgate.net/publication/387493189\\_Implementing\\_Scalable\\_Data\\_Architecture\\_for\\_Financial\\_Institutions](https://www.researchgate.net/publication/387493189_Implementing_Scalable_Data_Architecture_for_Financial_Institutions)
5. Birlasoft, "Enterprise Data Management Strategy Across Bank Financial Service Industry." [Online]. Available: <https://www.birlasoft.com/sites/default/files/resources/downloads/whitepapers/oracle/edm-strategy-across-bank-financial-servcive-industry.pdf>
6. Katie Joll, "Implementing Data Governance Frameworks in Banking for Effective Decision Making," Semarchy Data Governance Insights. [Online]. Available: <https://semarchy.com/blog/implementing-data-governance-frameworks-in-banking-for-effective-decision-making/>
7. Capgemini, "Simplifying the Banking Architecture." [Online]. Available: [https://www.capgemini.com/co-es/wp-content/uploads/sites/28/2022/12/simplifying\\_the\\_banking\\_architecture\\_2015.pdf](https://www.capgemini.com/co-es/wp-content/uploads/sites/28/2022/12/simplifying_the_banking_architecture_2015.pdf)
8. Hui Huang, Thien Sang Lim, "Construction and Optimization of Financial Risk Management Model Based on Financial Data and Text Data Influencing Information System," Journal of Information Systems Engineering and Management, 9(2), 24534, 2024. [Online]. Available: <https://www.jisem-journal.com/download/construction-and-optimization-of-financial-risk-management-model-based-on-financial-data-and-text-14767.pdf>
9. Badrudeen Teslim, Matthew John, "Data Security Challenges and Solutions in Financial Sector Data Lakes," ResearchGate, December 2024. [Online]. Available:



[https://www.researchgate.net/publication/387075874\\_DATA\\_SECURITY\\_CHALLENGES\\_AND\\_SOLUTIONS\\_IN\\_FINANCIAL\\_SECTOR\\_DATA\\_LAKES](https://www.researchgate.net/publication/387075874_DATA_SECURITY_CHALLENGES_AND_SOLUTIONS_IN_FINANCIAL_SECTOR_DATA_LAKES)

10. Aziz Shaikh et al., "Next-gen banking success starts with the right data architecture," McKinsey Digital Insights, February 28, 2025. [Online]. Available:  
<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/next-gen-banking-success-starts-with-the-right-data-architecture>