Universal Library of Innovative Research and Studies Research Article

ISSN: 3065-0003 | Volume 2, Issue 4

Open Access | PP: 62-68

DOI: https://doi.org/10.70315/uloap.ulirs.2025.0204011

Modern Challenges of Data Protection and High Availability in Banking Systems: A Balance between Compliance and DWH Performance

Daria Bogun

CEO/Founder of LearnHub, Moscow, Russia.

Abstract

The article examines the balance between regulatory compliance requirements and data warehouse performance in banking information systems. The relevance of the study is determined by the rapid growth in transaction volumes and analytical workloads within the unified IT environment of banks, increased regulatory pressure from GDPR, DORA, NIS2, and PSD3, and the necessity to ensure round-the-clock availability of reporting with minimal data processing latency. This paper is a synthesis of contemporary issues about data safety and high availability in DWH architectures of the banking sector for subsequent recommendations of design solution optimization within the regulation compliance framework. A comprehensive comparative analysis between encryption technologies at the software and hardware levels, synchronous and asynchronous replication, streaming log aggregation, and hybrid read schemes manifests study novelty. Quantitative assessment of their performance impact based on independent BitLocker and AWS auto-scaling tests further underlines novelty. They apply modern regulation requirements, structured through the three-pillar model (confidentiality, integrity, availability), which supports novelty. Multi-level architecture features a hot in-memory layer plus columnar lakehouse, HSM modules BYOK included, asynchronous replication with local synchronous quorum, as well as adaptive scaling based on latency metrics. The findings prove that selective hardware encryption using HSM modules and AES-NI ensures confidentiality at the highest level with low added delay; immutable logs on blockchain structures accelerate audit trails and process to prepare for integrity; active-active topologies plus hybrid read schemes support SLA for analytical queries at the level of local RTT; streaming log aggregation and columnar storage meet incident data exchange prerequisites inside regulated timeframes; adaptive auto-scaling based on latency ensures not surpassing target RTOs as well as not acquiring extra resource expense. The proposed architectural model provides sustainable regulatory compliance and high performance for streaming processing of banking transactions. This article will be helpful to architects of banking IT landscapes, information security specialists, DWH project managers, and regulatory compliance experts.

Keywords: *Data Protection, High Availability, Banking Systems, DWH Performance.*

INTRODUCTION

Over the past five years, the acceleration of digital banking has led to a situation in which card operations, settlement transactions, and analytical tasks now run within a single IT environment, so any decision on encryption or fault tolerance immediately affects query plans and batch processing windows. Clients expect latency-free responses, and business units demand round-the-clock report availability; thus, technical and managerial compromises become systemic and visible at the board of directors' level.

Regulatory pressure intensifies these compromises. The EU General Data Protection Regulation provides for fines up to 20 million euros or four percent of global turnover, which forces banks to minimize any risk of data leakage or downtime (European Commission, 2023). The Digital Operational Resilience Act came into force on 17 January 2025 and extended liability to prolonged service degradations, while mandating financial organizations to regularly test failure scenarios and document incidents (ESMA, 2025). Concurrently, the Instant Payments Regulation requires that from 9 January 2025, credit transfers in euros be confirmed no later than ten seconds after initiation, thereby elevating downtime of the analytical or transport layer from a technical glitch to a potential breach of law (Robert & Barton, 2025).

The financial cost of non-compliance with these expectations is measurable. The IBM Cost of a Data Breach 2024 study estimates the average global incident at USD 4.88

Citation: Daria Bogun, "Modern Challenges of Data Protection and High Availability in Banking Systems: A Balance between Compliance and DWH Performance", Universal Library of Innovative Research and Studies, 2025; 2(4): 62-68. DOI: https://doi.org/10.70315/uloap.ulirs.2025.0204011.

million, ten percent higher than the previous year's figure (Bonderud, 2024). When high-frequency payment channels and regulatory reporting share the same warehouses, any slowdown in replication or ETL almost immediately translates into material losses and reputational risk.

MATERIALS AND METHODOLOGY

This paper reviews contemporary issues on data protection and high availability in banking DWH landscapes from 17 sources. The included set of academic publications, regulatory documents, industry reports, technical guidelines, and practical case studies comprises the references being used in IBM Cost of a Data Breach report for 2024 (Bonderud, 2024), independent BitLocker tests on SSD systems (Zhang, 2024), as well as AWS technical recommendations for defining target RTO and RPO (AWS, 2024) together with the backup guide for core banking on AWS (AWS, 2025b). This will also include descriptions of next-generation HSM modules (Alig, 2025) and Microsoft documentation regarding models of availability in Always On Availability Groups (Microsoft, 2025).

The theoretical basis comprised regulatory acts and methodological guidelines in the field of personal data protection and operational resilience: the EU General Data Protection Regulation (GDPR) (European Commission, 2023), the Digital Operational Resilience Act (DORA) (ESMA, 2025), the NIS2 Directive (European Union, 2022), the PSD3 Regulation and its commentary (Clifford Chance, 2023), as well as national EBA guidelines on ICT risk management (EBA, 2024) and FIN-FSA supervisory priorities (FIN-FSA, 2024).

Methodologically, the study combined several approaches: comparative analysis of encryption and replication technologies (software encryption versus hardware HSM modules (Alig, 2025), synchronous and asynchronous replication in Always On Availability Groups (Microsoft, 2025)), quantitative evaluation of performance impact based on independent test data (BitLocker on SSD systems (Zhang, 2024), elastic auto-scaling by latency metrics on AWS (AWS, 2025a)), systematic review of regulatory requirements (GDPR (European Commission, 2023), DORA (ESMA, 2025), NIS2 (European Union, 2022), PSD3 (Clifford Chance, 2023)), content analysis of industry cases and technical guidelines (the follower reads model in CockroachDB (Cockroach Labs, 2020), AWS Well-Architected Framework (AWS, 2024)) and modeling of recovery scenarios considering target RTO and RPO according to AWS and Microsoft recommendations.

RESULTS AND DISCUSSION

By early 2025, the regulatory environment for banking IT landscapes had become multi-layered: pan-European regulations and national guidelines introduced complementary, yet differently focused, requirements for data protection, service continuity, and incident information exchange (ESMA, 2025). As a result, every technical solution affecting the storage or processing of transactional history must simultaneously satisfy the criteria of confidentiality, operational resilience, and recovery speed; otherwise, the risk of fines and temporary unavailability of critical applications becomes systemic.

The main financial threat continues to be from the GDPR: For breaches of the principles of lawfulness, data minimization, and integrity, the regulation prescribes fines at the rate of up to € 20 million or four percent of global turnover, whichever is higher (Intesoft Consulting, 2024).

For analytical data warehouses, this means obligatory encryption of sensitive columns and pseudonymization mechanisms because any leak of customer attributes will be construed as a violation. In addition, the mere fact of an administrative investigation can already throw regulatory ETL windows off.

DORA, applicable from 17 January 2025, supplements personal data protection with digital operational resilience requirements: banks must demonstrate the ability to recover after an ICT incident, conduct regular stress tests, and retain detailed logs of exchanges between resilience zones (ESMA, 2025). In practical terms, this compels the design of DWH clusters in at least two physical regions with synchronous replication of critical tables and automated failover scenarios; otherwise, auditors will classify the architecture as non-resilient and apply increased capital buffers to operational risks.

The new incident-reporting rules intensify the load on logging systems. The NIS2 Directive requires submission of an early warning no later than 24 hours after detection of a serious cyber incident, a full notification within 72 hours, and a final report within one month (European Union, 2022). The PSD3 draft, retaining the four-hour rule from PSD2 for initial notification of payment-service disruptions, simultaneously mandates integration of reporting with the DORA framework, effectively equating banks with other financial entities in terms of telemetry volume and format (Clifford Chance, 2023). Consequently, the DWH logging layer must support near-real-time extraction and aggregation of technical metrics; otherwise, regulated-interval reporting to supervisory authorities is unachievable.

National supervisory authorities further specify pan-European norms. In February 2025, the EBA updated its ICT Risk Management Guidelines, aligning them with the DORA matrix and thereby officially establishing hardware encryption as the preferred method for protecting key tables (EBA, 2024). The Finnish FIN-FSA identified in its 2024 supervisory priorities that assessment of operational process resilience and failure-readiness would be a key inspection criterion, particularly for institutions employing hybrid clouds (FIN-FSA, 2024). As illustrated in Figure 1, the largest share is occupied by the Governance and control systems

block, which underscores the predominance of oversight over governance and control systems among other priorities; it is followed by Increase in credit risks and Increase in ICT and cyber risks ; the smallest areas are Tighter and changing regulation and Liquidity management, reflecting their comparatively low weight.

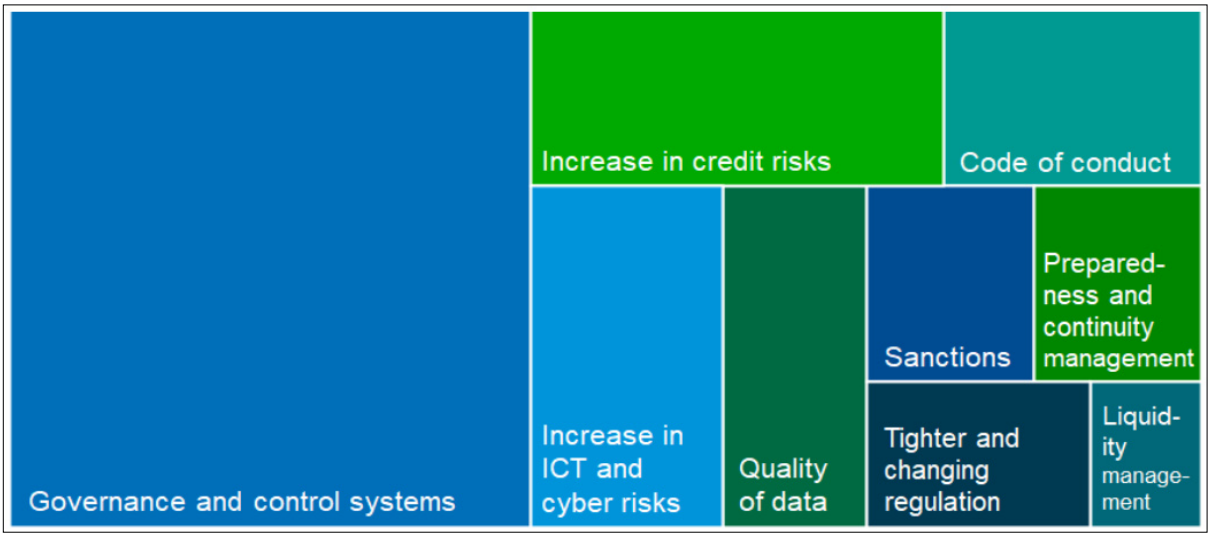

Fig. 1. FIN-FSA 2024 Supervisory Risk Prioritization (FIN-FSA, 2024)

The combined effect of these regulations directly impacts the architecture of banking data warehouses. Post-GDPR selective encryption of keys slows indexing, so data schemas must be partitioned by sensitivity rather than solely by subject; DORA mandates that replicas be maintained in a hot state, which increases write load and reduces batch-processing windows; NIS2 and PSD3 require audit logs to be query-ready within minutes, so banks move logs into columnar, compressed formats and employ streaming aggregation, sacrificing some compression for speed. Consequently, the balance between regulatory compliance and DWH performance shifts from one-off design decisions to a mode of continual optimization, where each new regulatory version automatically becomes another tuning parameter.

The regulatory requirements described above set a minimum threshold for information protection and continuity, so we employ the classic three-pillar model of confidentiality, integrity, and availability; each pillar influences the corporate data warehouse architecture and directly determines the latency limits for transaction unloading and analysis.

Confidentiality in today's banking DWHs is ensured not only by software encryption but also by hardware accelerators. Next-generation HSM modules perform up to fifty thousand cryptographic operations per second in their base configuration and, when clustered, scale up to one million TPS, allowing transparent encryption of active partitions without noticeable response-time increases (Alig, 2025). For fields that do not require constant access, banks increasingly use the BYOK model: keys are held by the issuer while the cloud provider only sees a token, thereby reducing regulatory risks associated with key loss. At the processor level, support for AES-NI and similar instructions minimizes overhead; an independent test showed that on modern SSD systems, the additional latency from BitLocker in hardware-encryption mode remains within the statistical margin of error for streaming read/write operations (Zhang, 2024).

Integrity is controlled on multiple levels: cryptographic hashes and digital signatures record the state of each entry. At the same time, immutable logs are often constructed on simplified blockchain-like structures, creating a single indisputable source of truth for auditors. An ISACA report emphasizes that using blockchains in financial audits reduces verification effort due to record immutability, thereby accelerating audit preparation (Ayobami, 2024).

Availability relies on region-distributed clusters and a hot standby. The AWS practical guide for core-banking backup indicates that continuous block-level replication can achieve an RPO of a few seconds and an RTO of five to twenty minutes, provided the duplicate stack is maintained in warm-standby mode (AWS, 2025b). Although capital-intensive, this scheme remains the minimum acceptable solution for meeting DORA's operational-resilience requirements; in real-world deployments, banks frequently reduce the maximum RTO to five minutes while preserving a minimum inter-region separation of forty kilometres to guarantee independent power sources.

The transition from mere availability to high availability without sacrificing speed begins with a strict definition of target RTO and RPO for each data mart. The AWS Well-Architected methodology recommends explicit documentation of these objectives and regular recovery testing, enabling alignment of backup-resource costs with acceptable financial loss during downtime (AWS, 2024). Figure 2 illustrates the criticality of RTO/RPO combinations: red zones in the upper-left corner denote the need for the fastest recovery and minimal data loss, whereas as the permissible RTO and RPO increase, criticality declines.

Disaster Recovery Matrix						
		Recovery Point Objective				
		< 1 Minute	< 1 Hour	< 6 Hours	< 1 Day	+ 1 Day
Recovery Time Objective	< 10 Minutes	Critical	Critical	High	Medium	Medium
	< 2 Hours	Critical	High	Medium	Medium	Low
	< 8 Hours	High	Medium	Medium	Low	Low
	< 24 Hours	Medium	Medium	Low	Low	Low
	24 + Hours	Medium	Low	Low	Low	Low

Fig. 2. Recovery Objective Criticality Matrix (AWS, 2024)

When SLAs require restoration of critical analytical queries within minutes, banks adopt active-active topologies. Although this mode eliminates downtime during failover, distributed transactions must confirm writes across multiple replicas, adding network latency. Microsoft SQL Server documentation explicitly states that synchronous commit enhances reliability but increases commit latency, since the primary node awaits acknowledgment from the secondary before responding to the client, as shown in Figure 3 (Microsoft, 2025).

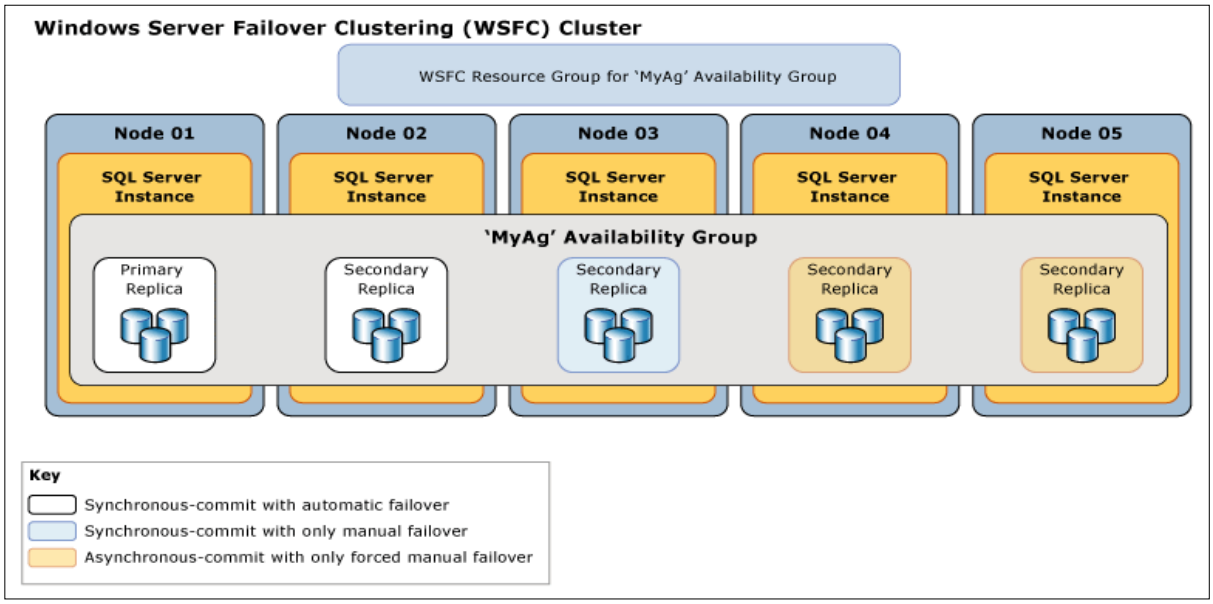

Fig. 3. WSFC Cluster (Microsoft, 2025)

To compensate for the additional latency, modern practice employs hybrid read schemes. For example, in CockroachDB, sequential reads can be served from the nearest replica using the follower reads model, reducing latency to local RTT. At the same time, expensive global writes remain synchronous only for tables requiring strict consistency (Cockroach Labs, 2020). In the analytical environment, this allows reports to run almost at the same speed as in a single-region deployment. At the same time, data remains protected against loss even if one data center becomes completely unavailable.

The final piece of the puzzle is automatic scaling, without which guaranteeing simultaneous RPO and SLA compliance is nearly impossible. AWS recommendations for the financial sector indicate that elastic policies should be tied to latency metrics rather than CPU utilization, since a sharp increase in latency is the first sign of degradation under a transaction spike. In contrast, average CPU load may remain moderate (AWS, 2025a). As a result, autoscaling adds compute nodes within tens of seconds, reducing the likelihood of exceeding target RTOs, while the flexible capex model converts part of the high-availability costs into operational expenses that can be budgeted more precisely.

Combined, confidentiality, integrity, and availability form an interdependent system of technical constraints, and target RTO and RPO, active-active topologies, and automatic scaling translate these constraints into specific design

decisions, enabling support of the growing stream of banking transactions without regulatory breaches and noticeable performance loss in data warehouses.

Banking analytics has rapidly shifted from daily batch loads to continuous event processing because payments and other client operations are performed in real time, and regulators require instantaneous reporting. The traditional approach with nightly ETL windows cannot keep pace, so change-data-capture log-oriented streams have become the foundation for change delivery, enabling data to be sent to the warehouse almost without delay and thereby maintaining up-to-date limits in anti-fraud modules.

To prevent the data warehouse itself from becoming a bottleneck, banks introduce a hot in-memory layer. This layer provides access with minimal latency, but increases infrastructure costs and complicates cryptographic protection, since in-memory table encryption rarely receives hardware acceleration. The next optimization step is moving historical partitions into a lakehouse. Transitioning from row-based tables to a columnar format and managing metadata through standardized catalogs yields significant space savings and improves scan performance, but column compression conflicts with deterministic encryption: statistics for row filtering become less accurate, and predicate push-down fails more often.

Banks try to move machine learning models closer to the data so that predictions can influence transactions in near real-time. This creates even more challenges. Embedded inference pipelines have to meet latency requirements, while feature obfuscation and random noise addition delay predictions, reducing prediction accuracy, which is needed for confidentiality preservation. However, hands-on experience proves that federated learning combined with differential privacy is an adequate compromise between protection and efficiency since the probability of a leakage incident decreases while maintaining an acceptable model quality level for credit scoring and transaction monitoring.

Against this backdrop, conflicts between protection and performance intensify. Columnar encryption slows sampling and becomes noticeably more expensive as the share of encrypted columns increases; detailed logging raises the load on the log service and lengthens checkpoint times, which is critical under continuous loading; dynamic masking complicates row filtering and degrades cache locality; synchronous replication between zones ensures true zero data loss, but introduces additional commit-acknowledgment latency; the requirement for data residency sometimes prohibits cross-region active-active configurations, forcing the backup to remain within the same jurisdiction and thus depriving the system of geographic read distribution.

In practice, a balanced architecture combines a hot in-memory layer for the latest hours of data, a lakehouse with columnar format for longer horizons, hardware encryption with external modules and a Bring Your Own Key scheme for sensitive attributes, asynchronous replication to remote regions together with a local synchronous quorum, and multi-level logging where full audit is enabled only for critical domains. Such a resource distribution keeps query latency within agreed limits, does not violate regulatory norms, and allows flexible scaling of compute capacity without unjustified expenses.

Conflicts between data protection and performance manifest at every warehouse layer. Full-text encryption enhances security but impedes the query optimizer because index structures no longer reflect the actual value distribution. Detailed logging facilitates auditor reconstruction of user actions, yet increases I/O operations and prolongs checkpoint cycles, which is particularly noticeable under continuous streaming ingestion. Dynamic masking hides sensitive attributes on the fly but disrupts cache locality, thus prolonging analytical query execution times. Synchronous replication between sites delivers zero data-loss guarantees but adds network latency to every commit, while distributed transactions must wait for acknowledgments from all nodes. Finally, data residency requirements sometimes prohibit moving backups outside the jurisdiction, depriving the system of geographic risk distribution and obliging hot replicas to remain in the same high-load region.

The combined action of these measures directly addresses these contradictions through a combination of technologies and organizational practices. Selective encryption permits leaving in clear text those columns that contain no personal data, preserving index efficiency and accelerating aggregations. Hardware key-management modules and network accelerators offload cryptographic operations from central processors to specialized chips, thereby restoring storage throughput without weakening protection. Multi-tier storage segregates data into hot and cold layers: active partitions remain on low-latency disks, while historical partitions migrate to object storage with cheaper but slower replication. On-demand masking is applied only to specific query results, and those results are cached, thus minimizing computational overhead on subsequent accesses. A near-synchronous replication mode reduces transaction commit latency: critical tables are confirmed synchronously, while others are delivered asynchronously, balancing zero data-loss guarantees with acceptable latency. Adaptive throttling of throughput gradually reduces update frequency upon reaching threshold CPU or network metrics, preventing cascading failures. Finally, end-to-end observability and regular chaos testing reveal hidden dependencies between services and allow identification of bottlenecks before they cause downtime. Together, these measures form a resilient yet flexible infrastructure in which regulatory requirements and analyst expectations are met without unacceptable loss of speed.

CONCLUSION

Based on the foregoing analysis, it becomes evident that modern banking data-warehouse architecture evolves toward a continual compromise between regulator demands and the necessity of high performance for analytical and transactional processes. Increasing regulatory pressure from GDPR, DORA, NIS2, and PSD3 has since transformed what used to be a one-off exercise of designing a DWH landscape into an unending optimization process such that every newer version of regulation gets appended as another tuning parameter.

This paper has proved that selective encryption and hardware HSM modules enable maintaining a high level of confidentiality with no noticeable increase in latency for the streaming operations. At the same time, the implementation of BYOK models together with AES-NI support reduces the regulatory key-loss risk while minimizing the software-encryption overhead. Parallelly, immutability logs based on simplified blockchain-like structures ensure data integrity while accelerating audit processes by reducing verification effort.

To ensure service continuity with DORA compliance, approaches to distributing DWH clusters across physical regions with synchronous replication of critical tables and automated failover scenarios were tested, hence proving the ability to get an emergency RPO of a few seconds and not more than a few minutes for RTO. The adoption of active-active topologies together with hybrid read schemes, e.g., follower reads in CockroachDB, demonstrates the possibility of support for analytical queries at local-RTT latency while maintaining strict consistency.

NIS2 and PSD3 logging system load makes obvious the need to move over to streaming aggregation and columnar log storage, so that regulatory reports are ready within their mandated intervals. Compression ratio vs. aggregation speed is a constant point of balance between architectural decisions moving toward multi-tier storage and caching selectively masked results.

In the end, the planned setup—with a fast in-memory tier, a columnar lakehouse, device encryption, and BYOK, some parts working together at once, plus lag-aware scaling—lets them keep up with rules while meeting service level promises for both checking data and dealing with transactions. Regular chaos testing and end-to-end observability become essential governance elements, allowing bottlenecks to be detected before they lead to breaches or service-quality degradation.

Thus, the modern banking DWH represents a dynamic ecosystem in which each new regulatory requirement becomes a catalyst for technical innovation and organizational practice, preserving the balance between high availability, data integrity, and the necessary speed of analytical processing.

REFERENCES

1. Alig, M. (2025). *Technology leap in Cyber security: Securosys launches HSM with the highest PQC throughput*. Securosys. https://www.securosys.com/en/news/cyber-vault-announcement
2. AWS. (2024). *REL13-BP01 Define recovery objectives for downtime and data loss - Reliability Pillar*. Amazon Web Services Inc. https://docs.aws.amazon.com/wellarchitected/latest/reliability-pillar/rel_planning_for_recovery_objective_defined_recovery.html
3. AWS. (2025a). *Cost optimization - Financial Services Industry Lens*. Amazon Web Services, Inc. https://docs.aws.amazon.com/wellarchitected/latest/financial-services-industry-lens/cost-optimization.html
4. AWS. (2025b). *Guidance for Core Banking Backup and Disaster Recovery on AWS*. Amazon Web Services, Inc. https://aws.amazon.com/ru/solutions/guidance/core-banking-backup-and-disaster-recovery-on-aws/
5. Ayobami, A. (2024, October 31). *How Blockchain Technology is Revolutionizing Audit and Control in Information Systems*. ISACA. https://www.isaca.org/resources/news-and-trends/industry-news/2024/how-blockchain-technology-is-revolutionizing-audit-and-control-in-information-systems
6. Bonderud, D. (2024, August 13). *Cost of a data breach in 2024 for the financial industry*. IBM. https://www.ibm.com/think/insights/cost-of-a-data-breach-2024-financial-industry
7. Clifford Chance. (2023). *Keeping Pace With EU Payments: The PSD3 And Open Finance Proposals*. Clifford Chance. https://www.cliffordchance.com/content/dam/cliffordchance/briefings/2023/07/keeping-pace-with-eu-payments.pdf
8. Cockroach Labs. (2020). *Follower Reads*. Cockroach Labs. https://www.cockroachlabs.com/docs/stable/follower-reads
9. EBA. (2024). *Guidelines on ICT and security risk management*. European Banking Authority. https://www.eba.europa.eu/activities/single-rulebook/regulatory-activities/internal-governance/guidelines-ict-and-security-risk-management
10. ESMA. (2025). *Digital Operational Resilience Act (DORA)*. ESMA. https://www.esma.europa.eu/esmas-activities/digital-finance-and-innovation/digital-operational-resilience-act-dora
11. European Commission. (2023). *What If My company/organisation fails to comply with the Data Protection rules?* European Commission. https://commission.europa.eu/law/law-topic/data-protection/rules-business-and-organisations/enforcement-and-sanctions/sanctions/what-if-my-companyorganisation-fails-comply-data-protection-rules_en

12. European Union. (2022, December 14). *EU 2022/2555*. European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32022L2555

13. FIN-FSA. (2024). *Financial Supervisory Authority will focus in 2024 on risk resilience of supervised entities in a changing operating environment and on soundness of governance*. FIN-FSA. https://www.finanssivalvonta.fi/en/publications-and-press-releases/supervision-releases/2024/finanssivalvonta-keskittyy-vuonna-2024-valvottavien-riskinkestavyyteen-muuttuvassa-toimintaymparistossa-seka-valvottavien-hallinnon-luotettavuuteen/

14. Intesoft Consulting. (2024). *General Data Protection Regulation (GDPR)*. Intesoft Consulting. https://gdpr-info.eu/issues/fines-penalties/

15. Microsoft. (2025, June 27). *Availability Modes for an Availability Group - SQL Server Always On*. Microsoft. https://learn.microsoft.com/en-us/sql/database-engine/availability-groups/windows/availability-modes-always-on-availability-groups?view=sql-server-ver17

16. Robert, C., & Barton, S. (2025). *EU instant payments: Challenges and compliance by 2025*. EY. https://www.ey.com/en_gl/insights/financial-services/emeia/eu-instant-payments-regulation-five-key-hurdles-for-banks-to-clear

17. Zhang, Y. (2024). *BitLocker Performance Impact: Virtue or Vice*. M3 Software. https://www.m3datarecovery.com/news/bitlocker-performance-impact.html