Resilient Architecture for Saudi Banking Treasury Systems: Integrating RPO/RTO, High Availability, Failover Clustering, and Disaster Recovery under Vision 2030.

Abstract:

Saudi banking treasury systems increasingly depend on always-on digital infrastructure as payment velocity, intraday liquidity management, market connectivity, and automated risk controls expand under Vision 2030. This review develops an integrated resilience architecture for treasury platforms by examining Recovery Point Objective (RPO), Recovery Time Objective (RTO), high availability, failover clustering, replication, backup integrity, and disaster recovery as one control system rather than isolated technical features. A structured review of literature and regulatory guidance published between 2020 and 2025 was conducted across banking resilience, cyber recovery, cloud reliability, database protection, and Saudi financial-sector modernization. The synthesis shows that resilient design requires four linked decisions: translating business impact into service-tier RPO/RTO targets; eliminating local single points of failure through redundant application, database, network, and messaging layers; separating high-availability mechanisms from geographically independent disaster-recovery capabilities; and validating recovery through rehearsed, observable, governed failover procedures. Strong infrastructure may still fail operationally when replication lag, identity dependencies, third-party services, runbook drift, data-integrity uncertainty, or untested decision rights are overlooked. The proposed framework therefore combines clustered local continuity, controlled cross-site replication, immutable recovery points, automated health detection, treasury-aware recovery sequencing, and post-incident learning. For Saudi banks, this architecture supports digital trust, risk discipline, scalable financial infrastructure, and the resilience objectives of Vision 2030.

Keywords: banking treasury systems; RPO; RTO; high availability; failover clustering; disaster recovery; operational resilience; Saudi Arabia; Vision 2030

1. Introduction

Treasury platforms sit near the operational centre of a modern bank. They support liquidity positions, cash forecasting, money-market activity, foreign-exchange dealing, securities funding, balance-sheet steering, counterparty exposure, and settlement obligations. An outage can therefore move from a technology problem to a liquidity, market, compliance, or reputational problem. As banking becomes more digital and interconnected, operational resilience increasingly depends on continuing critical services during disruption and restoring normal operations within predetermined tolerances [1–4]. Saudi Arabia intensifies this requirement because Vision 2030 and the Financial

Sector Development Program are expanding digital financial infrastructure, payment volumes, platform integration, and the competitive role of financial technology [25–27,29,30].

Traditional disaster-recovery thinking often begins with a secondary data centre and ends with a periodic failover exercise. That model is too narrow for treasury systems. A service can appear technically available while producing stale positions, delayed market data, incomplete settlement messages, inconsistent ledgers, or unusable authentication paths. Conversely, a protected database is of limited value if application clusters, network routes, security controls, market gateways, reference data, and operator procedures cannot recover in the correct sequence. The resilience problem is therefore architectural and socio-technical: systems must remain available when components fail, preserve data integrity across recovery boundaries, switch processing roles safely, and provide enough observability to distinguish transient faults from events that justify failover [2,3,19].

Banking and digital-transformation studies show that resilience outcomes depend on governance, preparedness, and recovery design rather than backup technology alone [5–8,16,20]. Financial-sector guidance emphasises impact tolerances, severe but plausible disruption, third-party dependencies, response coordination, and tested recovery [1–4,18,21,22]. Cloud and database engineering adds the technical distinction that high availability uses fault isolation and automated local recovery, whereas disaster recovery requires an independent failure domain, replicated state, recovery orchestration, and explicit RPO/RTO choices [9–11,19,24].

This review bridges these perspectives. It asks how banking resilience principles can be translated into a treasury architecture that is technically defensible, operationally testable, and aligned with Saudi priorities. The paper treats RPO and RTO as design inputs, distinguishes clustering from disaster recovery, examines replication trade-offs, and integrates data protection, security, observability, runbooks, and governance into one model.

1.1 Aim and Objectives

The aim of this review is to develop a resilient reference architecture for Saudi banking treasury systems that converts business continuity requirements into measurable technology controls. Four objectives guide the analysis. First, the study synthesises evidence from 2020–2025 on operational resilience, cyber recovery, high availability, failover, database replication, and banking continuity. Second, it clarifies how RPO and RTO should be derived from treasury business impact and then mapped to application, database, messaging, and infrastructure design. Third, it evaluates the complementary roles of high availability, failover clustering, cross-site replication, immutable backup, and disaster recovery, including the failure modes that appear when these mechanisms are treated independently. Fourth, it proposes an implementation and maturity pathway aligned with Saudi banking modernization, regulatory expectations, and Vision 2030. The intended contribution

is a review-based decision framework that allows architects, treasury technology leaders, risk teams, and continuity managers to discuss resilience using a common set of service tiers, dependencies, recovery evidence, and design trade-offs.

2. Review Methodology

A structured narrative review was selected because the research question spans peer-reviewed banking studies, technical architecture guidance, supervisory publications, and national financial-sector documents. A purely empirical meta-analysis would exclude much of the operational evidence needed to evaluate failover behaviour, recovery patterns, and regulatory expectations. The review period was restricted to 2020–2025 to capture post-pandemic operational-resilience practice, recent cloud and database architectures, contemporary cyber-recovery expectations, and the current phase of Saudi financial-sector digitalization.

Search terms were organised into five groups: bank operational resilience and business continuity; bank disaster recovery and treasury continuity; RPO/RTO and recovery objectives; high availability, failover clustering, replication, and multi-region architecture; and Saudi banking, financial stability, cyber resilience, and Vision 2030. Sources were retained when they informed at least one architectural decision involving impact tolerance, local availability, cross-site recovery, data integrity, cyber response, third-party resilience, or Saudi context. Duplicate and technically superficial sources were excluded.

The final evidence base combines international supervisory guidance, Saudi regulatory and policy sources, peer-reviewed studies, and selected vendor architecture documents. The mixed-source design combines regulatory expectations with implementation evidence; the 2026 study informed structure, not evidence. Evidence was coded into six themes: service criticality, local availability, cross-site recovery, data recoverability, governance and testing, and Saudi strategic alignment. These themes were then mapped into the architecture model and maturity roadmap.

3. Treasury Systems as Critical Banking Infrastructure

Treasury applications differ from many enterprise systems because their value is time-sensitive. A delayed liquidity position, foreign-exchange exposure, or settlement instruction can change financial risk. Architecture must therefore protect not only uptime but also the timeliness, completeness, and integrity of positions and transactions. Banking resilience research demonstrates that continuity capability can moderate the financial damage of disruptive events, while treasury-specific work highlights the need to preserve payment and operational functions during abnormal conditions [16,20]. Treasury should consequently be treated as an important business service with modelled dependencies.

A useful dependency map begins with business capabilities rather than servers. Treasury dealing may depend on market-data feeds, pricing engines, limit services, trade capture, confirmation, settlement messaging, cash accounts, identity services, databases, middleware, network routes, time synchronisation, and general-ledger interfaces. Intraday liquidity may additionally depend on payment-system connectivity and near-real-time account balances. A resilient design identifies which dependencies can degrade gracefully, which can be temporarily substituted, and which must recover before the service is usable.

This view aligns with operational-resilience guidance that focuses on maintaining critical operations through severe disruption [2,4,21]. It also corrects a common mistake: defining availability at the infrastructure layer while ignoring business-service availability. A database cluster can be healthy while stale market prices prevent trading, or an application server can be reachable while an identity outage blocks users. Metrics should therefore include business signals such as successful trade booking, position refresh age, payment-message completion, reconciliation status, and end-to-end user access.

Saudi banking adds strategic importance. National policy promotes advanced, digital, and stable financial infrastructure, while SAMA reporting highlights cyber risk, third-party dependence, and resilience as technology adoption deepens [12–15,27,30]. Each critical treasury capability should therefore be linked to a measurable outage tolerance, maximum acceptable data loss, tested recovery method, accountable owner, and evidence that its dependency chain can meet the same target.

Table 1. Evidence synthesis and architectural implications for resilient banking treasury systems.

Evidence theme	2020–2025 synthesis	Architecture implication	Treasury risk controlled
Operational resilience	Critical services must continue through severe but plausible disruption [1–4,21].	Define service tolerances, dependencies, decision rights and evidence.	Extended outage and governance ambiguity
Bank DR capability	Banking studies link continuity maturity with reduced disruption impact [5,16,20].	Treat DR as a business-service capability, not a backup task.	Liquidity, settlement and service interruption
RPO/RTO	Recovery objectives must reflect tolerated downtime	Tier services; align every mandatory dependency to	Unrealistic recovery commitments

Evidence theme	2020–2025 synthesis	Architecture implication	Treasury risk controlled
	and data loss [19,24].	the same target.	
High availability	Redundancy, fault isolation and automated recovery reduce local interruption [19,24].	Cluster or distribute compute, database, messaging and network layers.	Single-component and zonal failure
Cross-site DR	Independent standby environments support site or regional recovery [10,11,19].	Separate HA and DR failure domains; pre-stage capacity and configuration.	Data-centre or regional outage
Cyber recovery	Response, restoration and improvement require protected recovery and testing [1,3,22,28].	Use immutable backups, segregated credentials and clean recovery procedures.	Ransomware, destructive compromise, loss of trust
Saudi context	Digital growth increases systemic dependence on resilient financial infrastructure [12–15,17,27,29,30].	Integrate resilience metrics with bank governance and strategic modernization.	Operational and financial-stability risk

4. Translating Business Impact into RPO and RTO

RPO and RTO convert business impact into engineering constraints. RPO defines the maximum tolerable loss of committed data measured backward from a disruption, while RTO defines the maximum acceptable period before a service is restored to an agreed operating level. Contemporary platform guidance shows that different recovery mechanisms provide materially different combinations of downtime and data loss [19,24]. Treasury objectives must therefore be service-specific because market activity, settlement cut-offs, regulatory reporting, and overnight processing have different time sensitivities.

The first design step is a treasury business-impact analysis that identifies the consequence of lost transactions and lost time. A front-office trade-capture service during active markets may require near-zero data loss and recovery measured in minutes. A reporting mart may tolerate a longer RPO if it can be rebuilt from authoritative sources. Static reference data may have low write frequency but high recovery priority because multiple components depend on it. Objectives should consequently be assigned to services and data domains rather than inherited from one enterprise tier.

RPO determines replication architecture. Synchronous replication can approach zero committed-data loss because a write is acknowledged only after secondary protection, but distance and network latency can affect transaction performance. Asynchronous replication reduces latency coupling but creates a window in which acknowledged transactions may not yet exist at the recovery site. The correct choice depends on business value, acceptable loss, and latency budget. Across long distances, active-active processing can improve availability but complicates consistency, ordering, and split-brain control [19]. Near-zero objectives therefore shift cost into coordination, network design, testing, and governance.

RTO determines automation. Recovery measured in hours may permit manual restoration, whereas recovery measured in minutes requires pre-provisioned capacity, automated detection, scripted failover, configuration parity, and clear decision rights. Failover time must include database promotion, application reconnection, cache warm-up, message replay, security validation, route changes, integrity checks, and business acceptance.

The core principle is dependency consistency. A five-minute application RTO is meaningless if identity, market data, or network failover takes thirty minutes. Every mandatory dependency must meet or beat the same service tolerance, or the published recovery objective is not credible.

5. High Availability and Failover Clustering

High availability addresses frequent, relatively local failures without invoking full disaster recovery. Its purpose is to keep a service running when a server, process, disk path, network interface, availability zone, or software instance fails. Reliability guidance recommends removing single points of failure, automating recovery, testing fault behaviour, and maintaining enough spare capacity for failover [19,24]. Treasury HA should span application, database, messaging, integration, and network layers because service resilience is limited by the weakest mandatory dependency.

Failover clustering is one HA mechanism. Multiple nodes participate in a cluster and controlled ownership determines which node serves the workload. Health monitoring detects failure and moves the service or resource group to another node. Clustering is useful for stateful workloads that cannot be treated as independent stateless instances. However, a cluster in one data centre can still share power, storage, network, directory, or control-plane dependencies. It protects against component failure but not necessarily site loss, cyber compromise, corrupted shared storage, or administrative error propagated across nodes.

Active-active patterns distribute work across multiple healthy instances and can reduce recovery time because capacity is already serving traffic. Active-passive designs keep standby capacity ready until failure. Active-active can improve continuity and utilisation, yet treasury systems require

deterministic state management. If two instances can update the same trade, limit, or settlement state, concurrency and ordering controls must be explicit. Where consistency is difficult across distance, local active-active combined with remote active-passive DR may be safer than global active-active processing [19].

Health detection is equally important. False-positive failover can create a larger incident than the original fault. Monitoring should combine process health, service response, transaction success, replication state, dependency status, and synthetic business tests. Quorum and fencing must prevent split-brain. Planned switchover should be normal operational practice because controlled role reversal validates cluster behaviour and maintenance procedures [11].

The distinction is therefore practical: HA reduces interruption inside an operational site or region; DR restores service after failure exceeds that boundary. Treasury designs need both.

6. Replication, Data Integrity and Disaster Recovery

Disaster recovery begins where local high availability ends. The recovery environment must survive the failure mode that disables the primary, which normally requires geographic, administrative, and infrastructure separation. Database technologies using managed standby patterns demonstrate the core model: maintain a transactionally consistent secondary, monitor replication, and promote it during an unplanned event or reverse roles during a controlled switchover [10,11]. Cloud guidance likewise distinguishes local availability from regional recovery and recommends defined objectives, configuration parity, automation, and regular testing [19,24].

Replication is not backup. Logical corruption, malicious deletion, erroneous batch processing, or compromised credentials can copy bad data to the standby. Treasury therefore needs two recovery dimensions: continuity copies for rapid failover and protected recovery points for historical restoration. Immutable or isolated backups reduce the risk that ransomware or privileged compromise destroys both production and recovery data. Retention should reflect business cycles such as end-of-day positions, month-end reporting, settlement windows, and regulatory evidence.

Data integrity matters because the fastest recovery is not always the safest. After an incident, operators must know whether the standby contains complete and trustworthy transactions. Reconciliation should compare trade counts, sequence numbers, ledger totals, cash positions, message queues, and replication checkpoints before business release. With asynchronous replication, the runbook should identify the last safely applied transaction and determine whether missing events can be reconstructed from journals, source systems, counterparties, or payment records.

The DR site must also avoid hidden dependence on the primary. Typical weaknesses include authentication hosted only in the primary site, DNS or certificate services without independent

recovery, monitoring that disappears with production, backup catalogues in the same security domain, and third-party connections not pre-authorised from DR addresses. Financial-sector guidance stresses ecosystem dependencies because a bank cannot recover if a necessary provider cannot support the required time [2,3,18,21].

Recovery should therefore be sequenced in waves. Foundational identity, network, time, security, and storage recover first; authoritative databases, messaging, and integration follow; core treasury applications and market connectivity come next; analytics and lower-priority reporting recover afterward. Each wave needs technical entry criteria, business validation checks, and rollback rules. This converts DR from infrastructure restoration into controlled recovery of a treasury business service.

7. Proposed Resilient Treasury Architecture

The proposed architecture combines five control planes: transaction services, local continuity, data protection, disaster recovery, and governance. At the transaction layer, treasury applications are decomposed into services with explicit dependencies and service-tier targets. Critical components run across redundant compute nodes or zones behind load balancing, while stateful services use clustering or platform-native HA. Messaging is redundant and preserves ordered delivery or replay where required. Network paths, security appliances, and market gateways are dual-homed so application HA is not undermined by a single connectivity failure.

The data layer separates operational continuity from historical recoverability. Primary databases use synchronous or low-lag replication within the local HA boundary when latency permits. A geographically separate standby receives replication according to the service RPO. Logs, journals, and transaction checkpoints support replay and reconciliation. Independent immutable backups provide recovery from corruption or destructive cyber events. The architecture therefore recognises three objectives: continue despite a component fault, fail over after a site-level event, and restore to a clean historical point when live copies are untrustworthy.

The DR layer is warm or hot for the most critical services. Application binaries, infrastructure definitions, certificates, routing rules, market connections, and configuration are pre-staged and checked for drift. Failover automation coordinates database role transition, service startup, route changes, health validation, and alerting, while final activation can remain under controlled human authority where an incorrect switch carries material financial risk.

Observability spans all layers. Technical telemetry measures node health, latency, replication lag, queue depth, backup completion, and capacity. Business telemetry measures trade-booking success, stale-position age, unmatched confirmations, payment-message progress, and reconciliation differences. A resilience dashboard therefore shows whether the service is meeting declared

recovery objectives and whether dependencies are capable of failover. It distinguishes “system available” from “business service usable”.

The governance plane defines ownership, incident severity, failover authority, communications, evidence retention, test frequency, and post-incident actions. International guidance places governance, planning, restoration, communication, and improvement in the same recovery lifecycle [1,3]. For treasury, that lifecycle necessarily coordinates technology, operations, risk, cybersecurity, counterparties, and management.

Figure 1 visualises this closed-loop architecture, while Figure 2 presents the maturity progression from component redundancy to evidence-based business-service resilience.

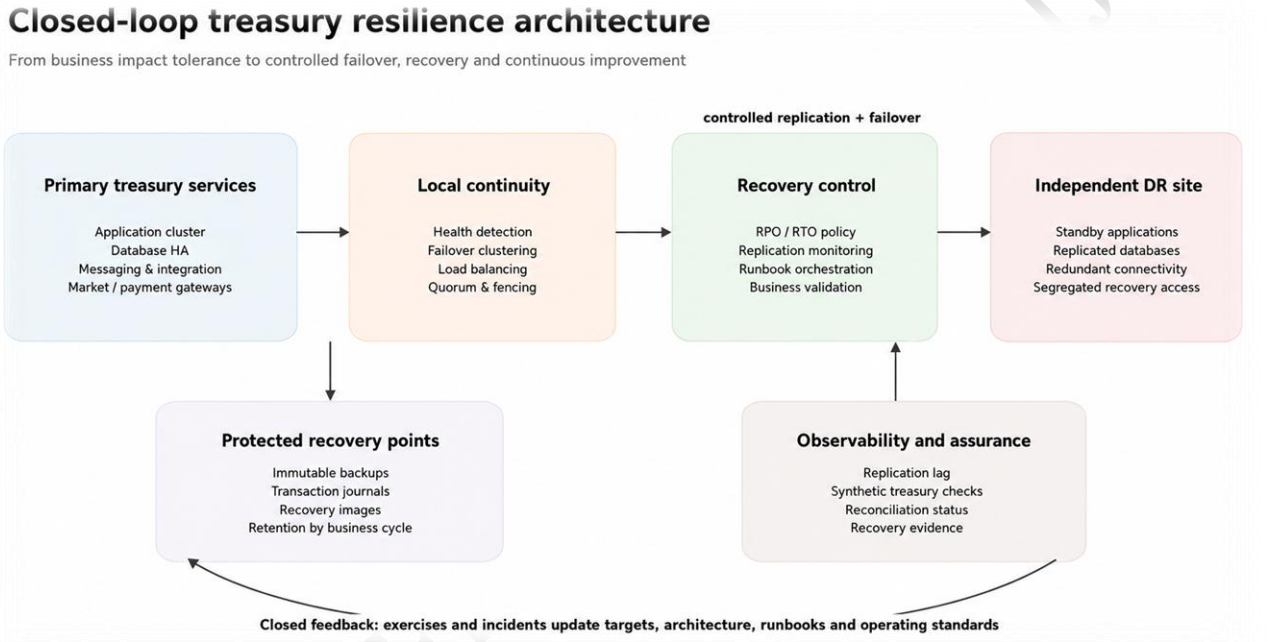

Figure 1. Closed-loop treasury resilience architecture integrating local HA, governed failover, protected recovery points, and an independent DR site.

7.1 Resilience Service Tiers and Operational Trade-offs

A practical architecture becomes easier to govern when treasury services are grouped into resilience tiers. Tier 0 should be reserved for capabilities whose interruption can create immediate liquidity, settlement, or market risk, such as payment release, intraday cash positions, critical dealing interfaces, or authoritative transaction stores. These services justify hot standby capacity, tightly controlled replication lag, automated health detection, and pre-tested failover paths. Tier 1 covers important services that can tolerate a short interruption but must return within the same operating window, including selected risk calculations, confirmation processing, collateral functions, and core treasury reporting. Tier 2 includes analytical or historical services that can be reconstructed from authoritative data and therefore tolerate longer recovery.

Tiering prevents an expensive mistake: assigning the same near-zero target to every treasury component. Resilience cost rises rapidly as RPO and RTO approach zero because the bank must duplicate capacity, maintain continuous synchronization, preserve network headroom, automate recovery, and test more frequently. A differentiated model directs the strongest controls to services where interruption produces the greatest business harm while allowing economical recovery patterns for lower-criticality functions.

The same tiering should govern change management. Tier 0 changes require evidence that rollback, replication, cluster health, and DR compatibility remain intact after deployment. Configuration drift between primary and recovery environments should be treated as a resilience defect, not as routine technical debt. Capacity reviews should also include failover conditions: a standby that can run normal average load may still fail during a disruption if transaction volume spikes or other workloads are moved onto the same environment.

Finally, treasury acceptance criteria should be defined before an incident. A recovered service should not be declared healthy only because infrastructure checks are green. Business acceptance should confirm current positions, complete trade and payment sequences, valid market data, reconciled balances, correct user entitlements, and reachable external interfaces. These criteria create a measurable handover from technology recovery to treasury operations and reduce pressure to resume processing before data integrity is established. They also provide audit evidence that recovery objectives are linked to operational outcomes, helping management compare resilience investment with treasury risk reduction.

Table 2. Illustrative treasury resilience tiers linking business criticality with RPO/RTO and recovery patterns.

Illustrative tier	Example treasury capabilities	Illustrative RPO	Illustrative RTO	Preferred pattern	Acceptance evidence
Tier 0 – critical	Payment release, authoritative trade store, intraday liquidity, core dealing	Near zero to minutes	Minutes	Local HA + hot DR + tightly monitored replication	Current positions, transaction continuity, settlement readiness
Tier 1 – important	Confirmations, collateral, key risk calculations, core treasury reporting	Minutes	< 1 hour	HA + warm/hot DR + automated recovery	Reconciled records, valid interfaces, user access
Tier 2 – recoverable	Analytics marts, historical reporting,	Hours	Several hours	Warm DR or rapid rebuild from	Successful restore/rebuild and

Illustrative tier	Example treasury capabilities	Illustrative RPO	Illustrative RTO	Preferred pattern	Acceptance evidence
	non-time-critical workflows			authoritative data	data completeness
Tier 3 – deferred	Archives, non-operational research extracts, ancillary tools	Daily or longer	Next business day	Backup/restore with documented priorities	Integrity check and controlled business release

8. Governance, Testing and Cyber Recovery

A resilient architecture is credible only when recovery can be demonstrated. Supervisory practice increasingly favours scenario-based testing because documentation cannot reveal timing problems, human ambiguity, or dependency failures [2,4,21,22]. The 2024 European banking cyber exercise focused on how banks respond to and recover from a severe but plausible attack rather than merely on preventive controls [22]. Treasury failover exercises should likewise test continuity under imperfect conditions, not only ideal infrastructure outages.

Testing should operate at four levels. Component tests verify node, process, network, and storage failover. Service tests verify a complete treasury application with database, identity, messaging, and integration dependencies. Site tests activate the recovery environment and validate published RPO/RTO. Cyber-recovery tests assume normal trust relationships may be compromised and require restoration from protected data, credential reset, and forensic validation. Threat-led testing and broader cyber-resilience practice reinforce the value of realistic exercises that expose weaknesses in detection, response, and recovery [3,23,28].

Runbooks must be executable and version-controlled. Each step should identify the owner, prerequisite, action, expected result, decision threshold, escalation path, and evidence captured. Automated scripts should be stored in a protected repository accessible during primary-site failure. Manual workarounds should exist for critical treasury processes if restoration exceeds tolerance, such as controlled position reconstruction, restricted transaction modes, or alternative settlement communication.

Cyber recovery requires extra safeguards because the incident may compromise the identities and administration tools normally used for DR. Recovery accounts should use segregated credentials, strong authentication, limited standing privilege, and independent logging. Clean recovery procedures may be required before systems reconnect to production networks. Backups must be tested for restorability, not merely completion.

Post-incident learning closes the loop. Root-cause analysis should distinguish the initiating fault from conditions that amplified impact, such as poor capacity headroom, hidden dependencies,

replication delay, brittle automation, or unclear authority. Corrective actions then update design standards, monitoring, test scenarios, and recovery assumptions. This improvement cycle prevents treasury DR from becoming a static compliance artefact [1–4].

9. Saudi Vision 2030 Alignment and Implementation Roadmap

Saudi financial-sector transformation creates a business case for resilient treasury architecture. The Financial Sector Development Program seeks a diversified, efficient, digital, and stable sector supported by advanced infrastructure and stronger institutional capability [27,29]. As payment activity, digital banking, fintech participation, and data-driven services expand, tolerance for technology disruption falls because more economic activity depends on continuously available financial rails. SAMA reporting has also emphasised systemic cyber risk and growing dependence on technology and third parties [12,30].

Implementation should begin with service classification rather than a technology purchase. Phase 1 establishes a trusted baseline: inventory treasury services, map dependencies, assign owners, define RPO/RTO, identify single points of failure, and verify backup integrity. A common service catalogue and dependency model remove inconsistent recovery assumptions.

Phase 2 builds local continuity. Critical applications are clustered or distributed across failure domains, database HA is configured, messaging and networks are redundant, and synthetic business monitoring is introduced. Planned switchovers validate architecture during maintenance. The objective is to absorb ordinary component failures without declaring a disaster.

Phase 3 integrates cross-site recovery. A geographically independent DR environment is provisioned, replication is aligned with service-tier RPO, configuration drift is monitored, and failover automation is developed. Treasury validation becomes part of the runbook, including reconciliation of trades, positions, queues, and settlement status. Results are reported against measured recovery time rather than simple pass/fail completion.

Phase 4 develops adaptive resilience. Banks run severe but plausible scenarios involving cyber compromise, third-party outage, simultaneous faults, or loss of administrative trust. Recovery metrics feed executive dashboards. Critical suppliers participate in tests, and contractual service levels are compared with treasury tolerances.

This progression supports Vision 2030 by protecting trust in digital banking, improving scalability as transaction volumes grow, and strengthening risk discipline through measurable engineering and governance outcomes. The result is not merely a faster DR site but a banking capability that allows digital growth without equivalent growth in operational fragility.

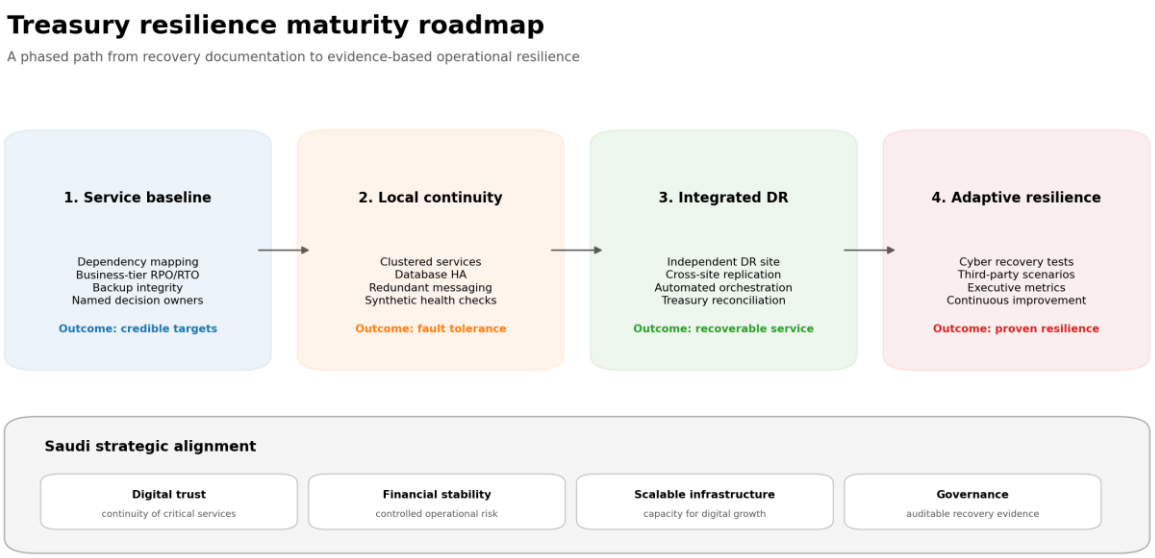

Figure 2. Treasury resilience maturity roadmap from service classification to adaptive, evidence-based recovery.

10. Discussion

The central design challenge is not choosing between high availability and disaster recovery but assigning each mechanism to the failure class it can control. HA and clustering handle local component failure; cross-site replication addresses site loss; backups address historical restoration and corruption; cyber-recovery procedures address loss of trust; and governance determines when and how controls are activated. Weakness appears when one mechanism is expected to compensate for another. A two-node cluster cannot protect against a shared storage fault, and a replicated standby cannot protect against corruption copied immediately.

RPO/RTO provide integration logic because they force architecture to be evaluated against business impact. Yet aggressive targets can create new risk. Near-zero RPO may require synchronous coordination that increases latency or couples failure domains. Near-zero RTO may require complex automation that is difficult to test and govern. The optimal design is therefore not the smallest possible number but the smallest defensible number whose cost, complexity, and operational evidence are proportionate to service criticality.

A second insight concerns recovery evidence. Organisations may describe sophisticated DR architecture without proving that every dependency can recover within the same tolerance. Resilience should therefore be measured through end-to-end exercises and business telemetry. The relevant question is not whether a standby database opened, but whether treasury users can authenticate, positions are current, market data is valid, payment instructions can progress, and reconciliations show no unexplained loss. This business-service perspective aligns engineering with regulatory emphasis on critical operations.

The Saudi context favours architecture that is both resilient and governable. Financial-sector modernization is increasing interconnectedness, and SAMA has highlighted technology, cyber, and operational risks as material components of financial stability [12,17,30]. Saudi banks therefore benefit from designs that make recovery state visible to management, integrate third-party readiness, and preserve auditable evidence.

Finally, the review supports a layered rather than vendor-specific framework. The same principles can be implemented with different database, operating-system, virtualization, or cloud technologies. What matters is failure-domain independence, controlled state transition, recoverable data, consistent security, observable service health, and tested decision rights.

11. Limitations and Future Research

This review has three limitations. First, the evidence base combines scholarly research, regulatory publications, and technical architecture guidance; these sources differ in method and abstraction. That diversity is necessary for architecture synthesis but limits direct comparison of quantitative outcomes. Second, banks rarely disclose detailed treasury recovery topologies, failover timings, or control weaknesses. The framework therefore generalises from banking resilience, payment infrastructure, database continuity, and treasury continuity evidence. Third, RPO/RTO values depend on each bank's product mix, market participation, data architecture, and risk appetite, so the service tiers in Table 2 are design examples rather than universal limits.

Future research should test the framework through Saudi bank case studies and measure actual recovery performance across treasury service classes. Comparative work could examine on-premises, hybrid, and cloud-native patterns; quantify the trade-off between synchronous replication latency and data-loss reduction; and evaluate automated failover under compound faults. Another priority is cyber recovery after destructive attacks, particularly methods for proving data integrity before market re-entry. Research should also examine third-party concentration across market data, identity, cloud, and payment gateways. Finally, resilience metrics could be linked to liquidity cost, failed-settlement exposure, operational loss, and counterparty impact, creating a stronger economic basis for resilience investment.

12. Conclusion

Saudi banking treasury systems require resilience architectures that combine continuity, recovery, data integrity, security, and governance. RPO and RTO should be derived from treasury business impact and enforced consistently across every mandatory dependency. High availability and failover clustering reduce local disruption, but they do not replace geographically independent disaster recovery. Cross-site replication enables rapid restoration, while immutable backups and clean recovery procedures protect against corruption and destructive cyber events. Observability,

reconciliation, decision rights, and scenario testing convert these mechanisms into dependable business outcomes.

The proposed framework links the controls in a closed loop. Local HA absorbs routine failures; an independent DR environment protects against site-level disruption; protected recovery points preserve historical integrity; automated and governed orchestration manages role transition; and business telemetry confirms that treasury is usable, not merely online. A phased roadmap moves banks from inventory and service classification through local continuity, integrated DR, and adaptive scenario-based resilience. Under Vision 2030, this approach supports a more digital and scalable financial sector without accepting equivalent growth in operational fragility. Resilience becomes an architectural property of treasury services and an auditable management capability rather than an emergency procedure invoked only after normal operations fail.

References

1. Financial Stability Board. Effective Practices for Cyber Incident Response and Recovery: Final Report. Basel, 2020.
2. Basel Committee on Banking Supervision. Principles for Operational Resilience. Bank for International Settlements, Basel, 2021.
3. Financial Stability Institute. Cyber Resilience Practices: Executive Summary. Bank for International Settlements, Basel, 2021.
4. Prudential Regulation Authority. Operational Resilience: Statement of Policy SoP1/21. Bank of England, London, 2021.
5. Tariku, N.; Lessa, L. Towards a Conceptual Framework for Information Technology Disaster Recovery Plan for Banks: Based on Cases from Ethiopia. International Journal of Computer Applications 2021, 174(10), 35–42. https://doi.org/10.5120/ijca2021920976
6. Chen, H.; Tse, D.; Si, P.; Gao, G.; Yin, C. Strengthen the Security Management of Customer Information in the Virtual Banks of Hong Kong through Business Continuity Management to Maintain Its Business Sustainability. Sustainability 2021, 13, 10918. https://doi.org/10.3390/su131910918
7. Verhoef, P.C.; Broekhuizen, T.; Bart, Y.; Bhattacharya, A.; Dong, J.Q.; Fabian, N.; Haenlein, M. Digital Transformation: A Multidisciplinary Reflection and Research Agenda. Journal of Business Research 2021, 122, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022
8. Thakor, A.V. Fintech and Banking: What Do We Know? Journal of Financial Intermediation 2020, 41, 100833. https://doi.org/10.1016/j.jfi.2019.100833
9. Amazon Web Services. Disaster Recovery Compliance in the Cloud, Part 1: Common Misconceptions. AWS Security Blog, 2021.

10. Oracle. Deploy a Hybrid Disaster Recovery Topology for an On-Premises Database Using Oracle Data Guard. Oracle Cloud Infrastructure Solution, 2021.
11. Amazon Web Services. Amazon RDS for Oracle Supports Managed Oracle Data Guard Switchover and Automated Backups for Replicas. 2022.
12. Saudi Central Bank. Financial Stability Report 2022. Riyadh, Saudi Arabia, 2022.
13. Saudi Central Bank. Cybersecurity Requirements for Operational Resiliency and Testing of Banking Systems and Applications. SAMA Rulebook, 2022.
14. Saudi Central Bank. Cyber Threat Intelligence Principles for the Financial Sector. Riyadh, Saudi Arabia, 2022.
15. Financial Stability Institute. Principles for Operational Resilience: Executive Summary. Bank for International Settlements, Basel, 2022.
16. Bastan, M.; Tavakkoli-Moghaddam, R.; Bozorgi-Amiri, A. Resilient Banking: Model-Based Assessment of Business Continuity Policies on Commercial Banks. Kybernetes 2023, 53(12), 5325–5357. https://doi.org/10.1108/K-07-2022-0981
17. Saudi Central Bank. Operational Risk Requirements for Banks. SAMA Rulebook, 2023.
18. Financial Stability Board. Recommendations to Achieve Greater Convergence in Cyber Incident Reporting: Final Report. Basel, 2023.
19. Istrati, E.; Iu, J.; Makota, T. Architecting Critical Payment Systems for Multi-Region Resiliency. AWS for Industries, 2024.
20. Nugraha, T.A.; Djakman, C.D. Implementation of Task Continuity Management in Banking Service Activities and Treasury Operations for Smooth Payment System. JIFA 2024, 6(2), 128–144. https://doi.org/10.22515/jifa.v6i2.7728
21. Adeney, R.; Hitchins, A.; Lane, C.; Mehta, H.; Quashie, A. Operational Resilience in a Macroprudential Framework. Financial Stability Paper No. 50, Bank of England, 2024.
22. European Central Bank. ECB Concludes Cyber Resilience Stress Test. Banking Supervision Press Release, 26 July 2024.
23. European Central Bank. Revised Eurosystem Cyber Resilience Strategy for Financial Market Infrastructures and Payment Entities. Frankfurt, 2024.
24. Microsoft. High Availability and Disaster Recovery Checklist for Azure SQL Database. Microsoft Learn, 2025.
25. Citterio, A.; King, T.; Locatelli, R. Is Digital Transformation Profitable for Banks? Evidence from Europe. Finance Research Letters 2024, 70, 106269.
26. Bueno, L.A.; Sigahi, T.F.A.C.; Rampasso, I.S.; Leal Filho, W.; Anholon, R. Impacts of Digitization on Operational Efficiency in the Banking Sector: Thematic Analysis and Research

Agenda Proposal. International Journal of Information Management Data Insights 2024, 4, 100230.

27. Financial Sector Development Program. Annual Report 2024. Kingdom of Saudi Arabia, 2025.
28. European Central Bank. TIBER-EU Framework Updated to Align with Digital Operational Resilience Requirements. Frankfurt, 2025.
29. Saudi Vision 2030. FinTech Strategy. Government of Saudi Arabia, Riyadh, 2025.
30. Saudi Central Bank. Financial Stability Report 2025. Riyadh, Saudi Arabia, 2025.