Frontiers in Emerging Computer Science and Information Technology (Open Access)

Volume 01, Issue 02, September 2024,
Publish Date: 09-04-2024
PageNo.50-74

Quantum-Safe VPNs for Financial Institution Backbones

Ashutosh Chandra Jha
Network Security Engineer, New York, USA

Abstract
This study assesses quantum-safe virtual private networks to backbones of financial institutions that experience harvest-now-decrypt-later risk. It integrates hybrid key exchange with X25519 and Kyber, as well as KEM-authenticated TLS. A testbed is used to simulate metro, regional, and long-haul paths (10, 50, and 100 ms round-trip, 1 percent loss), measuring time to first valid packet, goodput, fragmentation, and CPU utilization, as researchers vary the MTU, NAT-T, and offload. The threat model includes passive recording, active downgrade, man-in-the-middle attack interference, and insider misuse: experimental MOS-PK 47, 48. Greater post-quantum artifacts and HelloRetry/IKE cookies challenges increase control-plane cost, driving cold-start latency to hundreds of milliseconds on very long-haul links. Throughput on the data plane remains several percent apart from classical. Jumbo frames minimise the CPU load of software gateways, and ASIC offloading can be used to prevent capacity loss. The primary failure inducers are path-MTU black-holing as enlarged handshakes cross mixed-MPLS and Internet paths. Silent drops are eliminated with MSS closer to 1360 bytes, deterministic DF policy, active PMTUD, and IKE fragmentation stabilizes rekeys. The study provides a deployment playbook that jitters and batches rekeys, pins cipher suites, and monitors golden signals, and advances a blueprint that is auditable. The ability to migrate PKI with hybrid or cross-chained hierarchies, OCSP stapling, uniform telemetry, and cohort-based promotion provides the needed crypto agility without breaking service levels. The feasibility is proven, and regular validation and reassessment drills address residual risk.

Keywords: *Quantum-safe VPNs, post-quantum cryptography (PQC), Hybrid key exchange (X25519 + Kyber), IKEv2 and TLS 1.3 / KEMTLS, Harvest-now, decrypt-later (HNDL)*

1. Introduction

Finance institution backbones are used to connect data centers (DCs), regional hubs, disaster recovery (DR) sites, branches, trading floors, and cloud edges based on carrier MPLS and Internet-based SD-WAN, as well as on private optical waves. They carry latency-sensitive protocols, such as SWIFT, FIX, market data, core banking RPC, and card authorizations, as well as bulk replication and backups. Examples of the overlay design are IPsec (IKEv2/ ESP) and TLS-based tunnels; the underlay is a BGP/IGP routing with SD-WAN controllers to manage policy and path selection. Opponents have graduated beyond criminal eavesdroppers to more involvement, including passive collection, state-level opponents that can gather at scale. Harvest now, decrypt later (HNDL) is an encryption scheme that encrypts today but can be efficiently decrypted in the (probable) future when quantum cryptanalysis can be applied. Since a lot of the financial records cannot be disclosed to anyone for many years, a legacy of risk builds up. Regulators also have a growing focus on auditability, provable key custody, and crypto agility, where schemes must rotate algorithms, keys, and certificates without impact to service. A quantum-safe VPN program thus becomes a practical necessity as opposed to a scholarly exercise.

A question is how to raise VPN control and data planes to quantum-safe levels, preferably in a hybrid way, without missing service-level objectives (SLOs) and breaking interoperability. Control-plane modifications are required to cover IKEv2 or TLS 1.3

handshakes, certification validation and access policy, key lifetime, and rekey behavior in the face of loss and jitter. The data-plane effects comprise an increase in MTU pressure due to larger key-encapsulation (KEM) and signature artifacts, packetization policy, Path MTU Discovery dynamics, and CPU or ASIC load on the edge device. Goals will be three-fold: first, generate an experimental performance and security analysis of typical hybrid key exchanges (e.g., X25519+Kyber). Second, come up with a practical MTU and fragmentation rulebook that maintains goodput and reduces black-holing on mixed MPLS/Internet paths. Third, provide an operational plan including PKI migration, monitoring, incident response, and secure roll-out patterns that fit in regulated environments.

This paper brings a comparative analysis of hybrid key exchange to IPsec/IKEv2 and TLS 1.3/KEMTLS, analyzing handshake latency, message sizes, retry and downgrade failure behavior, and data-plane goodput in the presence of realistic impairment. It also gives an MTU and fragmentation playbook that quantifies MTU and fragmentation overheads of post-quantum artifacts, suggests using DF and MSS clamping, and reports PMTU black hole defences on MPLS and Internet routes. It details PKI migration patterns, including hybrid certificates, cross-signing, OCSP stapling at scale, and telemetry to prove policy and detect downgrade attempts. It also provides operator runbooks with drills to control changes, rollback, and crypto-agility, to allow upgrades to be introduced during maintenance windows without impairing availability.

It has Layer 3 overlays utilizing IPsec and Layer 4/7 overlays using TLS 1.3 or KEMTLS as service meshes and proxy-terminated links. It presupposes the availability of edge routers and firewall-based hardware offload support of the classical primitives, with the post-quantum backing to be implemented in software or emerging accelerators. The issuance of keys is done in FIPS-grade HSMs; the process of enrollment and rotation is automated by using an enterprise PKI, where the distribution of CRLs and OCSP is provided. The fleet is a heterogeneous, multi-vendor, and multi-domain; change control is canary promotion with tested acceptance and defined rollback. Traffic profiles/classes: market data, trading, retail payments, and replication; Service-level objectives: latency budgets, packet-loss ceiling, recovery timings; compliance logs: must be tamper-evident.

This study is presented in different chapters. Chapter 2 presents an overview of appropriate standards, algorithms, and relevant measurements so far. Chapter 3 describes data sets, pre-processing, visual analytics, threat model, and measurement testbed. Chapter 4 describes a post-quantum architecture of a VPN: libra options, hybrid protocol, PKI migration, interoperability, operational or compliance considerations. Controlled experiments of the control-plane and data-plane performance, reliability, and cost or benefit are given in Chapter 5. Chapter 6 concludes the evaluation with a trade-offs synthesis, a deployment blueprint proposal, a limitation and threats to validity analysis, and finally, closing comments on standards maturation, hardware offload, and hybrid-to-PQ-only migration.

2. Literature Review

2.1 Quantum Threat Timeline & HNDL

Financial institutions use networks in which the confidentiality should live beyond the systems that uphold it. Payment histories, loan documentation, cardholder personal information, trade data, and risk model flows have secrecy periods that are decades because of retention regulations and litigation windows. The HNDL threat potential reshapes backbone security: recovering tactics is now that well-heeled malefactors stage VPN-transported traffic today and wait to decrypt it later in time, as exponentially scaled quantum cryptanalysis materializes. In such a framing, time-to-compromise no longer correlates with time-of-capture; the pertinent question is not whether ciphers currently in use can resist attack within a given year, but whether information obtained within a given year can continue to remain secret over the next 1020 years.

Institutions thus translate the time periods of the threat into a map of asset exposure. Tunnels carrying inter-datacenter replication, database backups, SWIFT messages, FIX/FAST market data, core banking RPCs, or compliance telemetry may use shared internet connections during brownout or disaster-

recovery drills. Such windows introduce more control-plane renegotiation and presentation of downgrade surfaces, with the data itself being high-value and often patterned--which makes forced harvesting of ciphertext appealing. Perfect forward secrecy (PFS) is still required but still not sufficient: even without any keys being exfiltrated, a future quantum attacker would break classical ECDH assumptions in recorded sessions today [6]. A practical HNDL calculus categorises by the confidentiality horizon and migration criticality of flows. Urgent deals with traffic S employing 10+ years of secrecy requirements and Internet access; Near-term deals with flows a on private backbone that lose secrecy during failover; Opportunistic is fluid requirements such as developer or staging traffic.

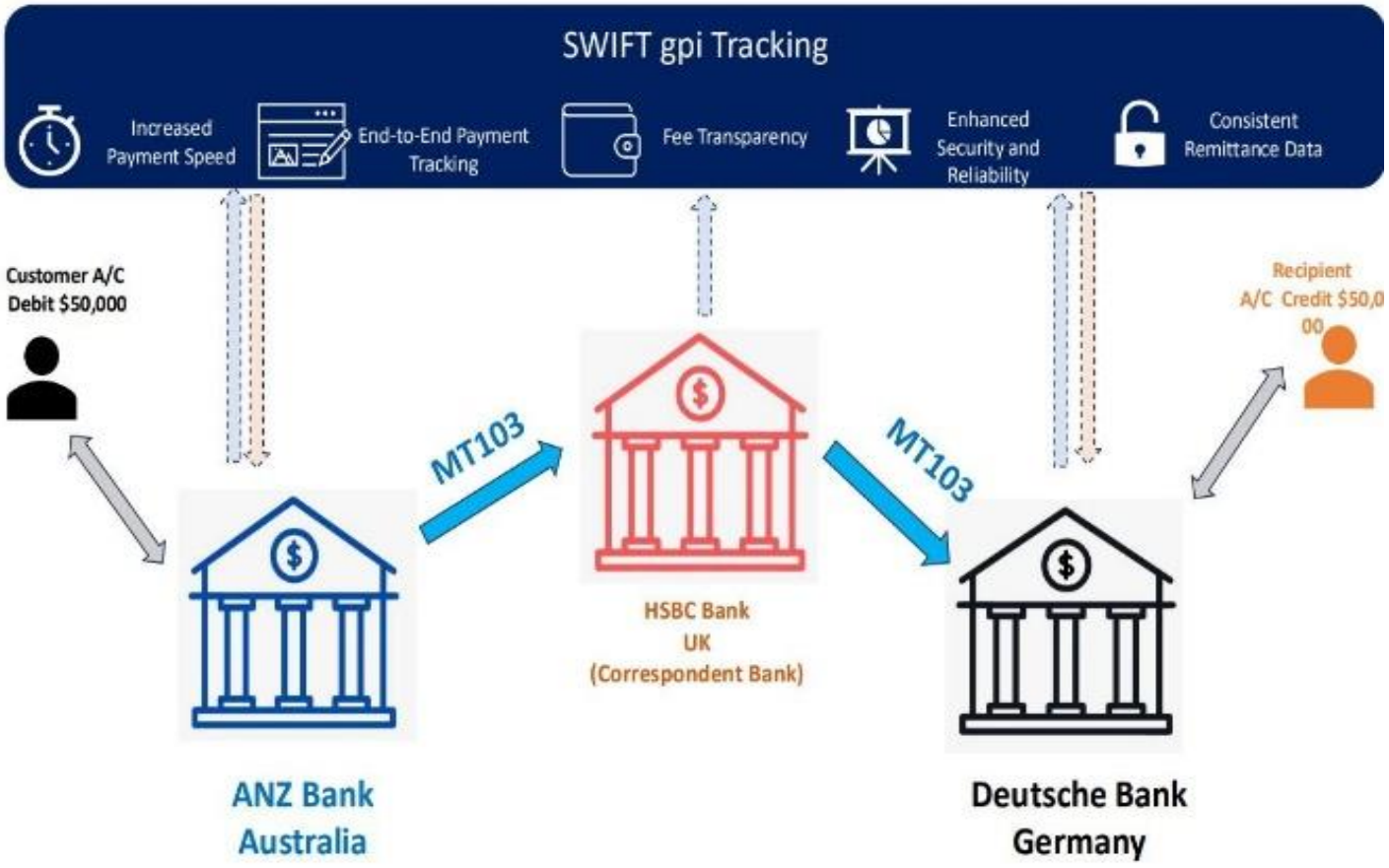

Figure 1: SWIFT MT103 flow across banks highlighting harvest-now decrypt-later risk

As shown in the figure above, a SWIFT gpi MT103 payment flows through originating, correspondent, and beneficiary banks with an end-to-end view of the flow process, showing where backbone VPN connections can renegotiate and provide downgrade surfaces when failing over via the internet. It has a high-value, repeatable SWIFT payload that is appealing to harvest-now-decrypt-later adversaries, so perfect forward secrecy is needed but not sufficient since classical ECDH over recorded material can be defeated at some time by future quantum capability. A practical HNDL calculus prioritizes such flows as urgent and near-term, with the categories of opportunistic and private-backbone traffic and staging.

Migration sequencing is based on this classification and on operational constraints: the market opening and closing times, payment cutoffs, batch windows, and regulatory blackout periods constrain change windows. Establishing pilot sites, followed by canary releases and finally full E2E rollout, can all be additionally isolated under crypto-agility controls, where the issuance of the certificates can be policy-based, the profiles versioned, and there is automated rollback and an inventory of cipher/policy compatibility to hedge against revisions to the parameters or profiles used during the transition.

2.2 Classical VPNs & Where PQC Fits

Financial Backbone overlays are typically a mix of IPsec tunnel mode (used as L3 backbone tunnels),

interspersed with TLS 1.3 (usually mTLS) service meshes, API gateways, partner links, and brokered interconnects. WireGuard is presented as a site-to-site adjunct and tunneling access because of its clean handshake and keying design [37]. The points of insertion of post-quantum cryptography (PQC) vary by protocol. In IKEv2, PQC appears in the key-exchange payloads (hybrid DH+KEM) and in authentication (PQ signatures or KEM-based authentication); rekey timers, anti-DoS cookies, and fragmentation processing are now first-order issues because messages become larger.

In TLS 1.3, PQC can be added via hybrid KeyShare values or KEM-based authentication (e.g., KEMTLS variants) in which peers demonstrate a KEM public key through encapsulation/decapsulation instead of conventional signatures. A more direct consequence of the separation of control and data planes is that larger handshakes may become more likely to IP-fragment and cause PMTU misfeatures with MPLS labels, GRE/VXLAN overlays, and cloud load balancers that break MSS/DF. Operators thus configure and tape off MTU/MSS clamping, turn on handshake-sensitive retry, and test HelloRetryRequest frequency under loss/Latency.

The parsimonious nature of WireGuard also links cryptographic decisions to minimum feature sets, making direct PQ use nearly infeasible; currently, most organizations will likely use IPsec/TLS to control the PQ-based plane and run WireGuard as an inner overlay network until PQ-enabled versions are available. Telemetry should be extended to record negotiated suites, parameter identifiers, transcript hashes, certificate sizes, handshake byte counts, and exact failure taxonomies (e.g., PMTU blackhole, cookie-challenge rate-limit, downgrade refusal). Signals to SLOs include time-to-first-valid (TTFV) handshake, rekey success rate, and path-specific MTU hit rates, which allow for a safe scale rollout.

2.3 PQC Building Blocks

The most feasible PQC building blocks to use with VPNs are lattice-based key-encapsulation mechanisms (KEMs) to establish keys and lattice- or hash-based signatures to authenticate endpoints. Most notably, Kyber families are preferentially used as KEMs since they offer balanced security/performance, whilst the standard parameter sets still have relatively desirable public keys and ciphertexts that minimize fragmentation risk in long-RTT environments with lossy channel characteristics. In signatures, both Dilithium and Falcon trade-offs: Dilithium has much larger signatures but has simple constant-time implementations; Falcon offers more minor signatures, and verification is fast, but more complex deployments are required. These primitives cause dynamic changes when incorporated into IKEv2/TLS: ClientHello and ServerHello increase, chains of certificates become longer when PQ signature is used, and the retry state becomes broader. Two integration patterns prevail in the examples. Second, hybrid key exchange is constructed by pairing classical ephemeral (e.g., X25519) with a KEM (e.g., Kyber), deriving a session secret that is only broken when both parts are broken. This mitigates short-term risk and facilitates interoperability in the case of cross-fleet transitions.

PQ authentication can also be in combination with signing on certificates/handshakes or authenticated encryption KEMTLS, where instead of signatures, authenticated encryption is used; the latter shifts signature quota on the wire to the certificate and stack change. KEM-set choice is an optimization problem: the larger KEMs can mitigate cryptanalytic risk, at the cost of longer handshake payload and CPU consumption, which limits TTFV under high RTT. The PMTU planning should include outer encapsulations (MPLS, GRE, VXLAN) and security associations to prevent fragmentation chains [7]. Since production payments and PII traffic is treated as sensitive information in production setting, institutions can leverage synthetic yet realistic documentation to assess system functionality in the lab; generative modeling in privacy-sensitive field provides valuable insights on how to produce high-fidelity synthetic data that replicates the structural features of the original, without leaking secrets, and this same concept can be applied to PQ handshake and certificate-path testing [35].

2.4 Standards Status

Standards maturity shapes what enterprises can purchase and how multi-vendor fleets can be used together. Hybrid key-exchange and PQ-auth specifications in the IPsec/IKEv2 stack identify their

payload specifications, transcript binding, and rules governing negotiation semantics and downgrade resistance; vendors implement these to provide larger trade tolerance, capability discovery, cookie-challenge behaviour, and rekey policies. TLS 1.3 hybrid KeyShare drafts and KEM-auth variants discuss capability advertisement, HelloRetryRequest triggering, transcript coverage, and certificate usage, and certificate profile work has attempted to describe mixed classical/PQ chains, OCSP stapling behavior, and CRL size limits in the context of PQ signatures. In addition to the wire image, the matter of operational standardization counts.

To permit dashboarding and audit records across a fleet, institutions require consistent telemetry: negotiated suites and parameters, certificate and chain sizes, HelloRetryRequest reasons, PMTU-related retries, and categorical error codes must all be emitted in uniform ways to allow the operation of fleet-wide dashboards and compliance records without brittle parsers. Language in procurement should specify hybrid or PQ Modes, emphasize PMTU/fragmentation stress testing, and specify interop matrices involving mixed-vendor tunnels and asymmetric capability negotiation. The program should be explicitly cost-conscious: enlarged handshakes, more frequent retransmission under loss, and certificate distribution updates impose compute and bandwidth headroom burdens. The approaches that have been popularized in microservices engineering, namely incremental rollout, automated canary analysis, and cost-constrained scaling, can serve as a master plan of balancing scalability and spend during cryptographic transitions [4].

2.5 QKD vs PQC & Evidence from Prior Studies

Quantum key distribution (QKD) is an information-theoretically secure form of key exchange over optical fibre. However, it requires either dark fibre or trusted fibre, dedicated endpoints, carefully controlled environmental conditions, and is limited in distance by the lack of trusted repeaters. Such constraints are not well-suited to heterogeneous financial backbones that traverse metro rings, long-haul carriers, cloud interconnects, and even partner networks. PQC, in contrast, aims to enhance the software stacks and commodity hardware, leveraging existing IP/optical backbones and intersecting with existing operational models - change windows, SLOs, and telemetry. In the shorter-to-medium time horizon, PQC therefore offers the physically more feasible pathway to quantum-safe backbones, with QKD to be reserved for niche. In these fiber-heavy routes, organizations can justify the additional lifecycle expenditures.

As shown in Figure 2, a software-defined networking stack over optical infrastructure, a quantum key distribution layer includes controllers that send keys to classical and data channels, via OpenFlow handshakes and gateways. The architecture demonstrates the requirement of QKD to dedicated fiber, stability, and short range without trusted repeaters, which creates operational complexity when heterogeneity is found in the financial backbones [2]. Comparatively, post-quantum cryptography retrofits existing IP and optical routes with software endpoints and commodity silicon; therefore, PQC is the reasonable near-term solution, except for niche fiber-rich routes.

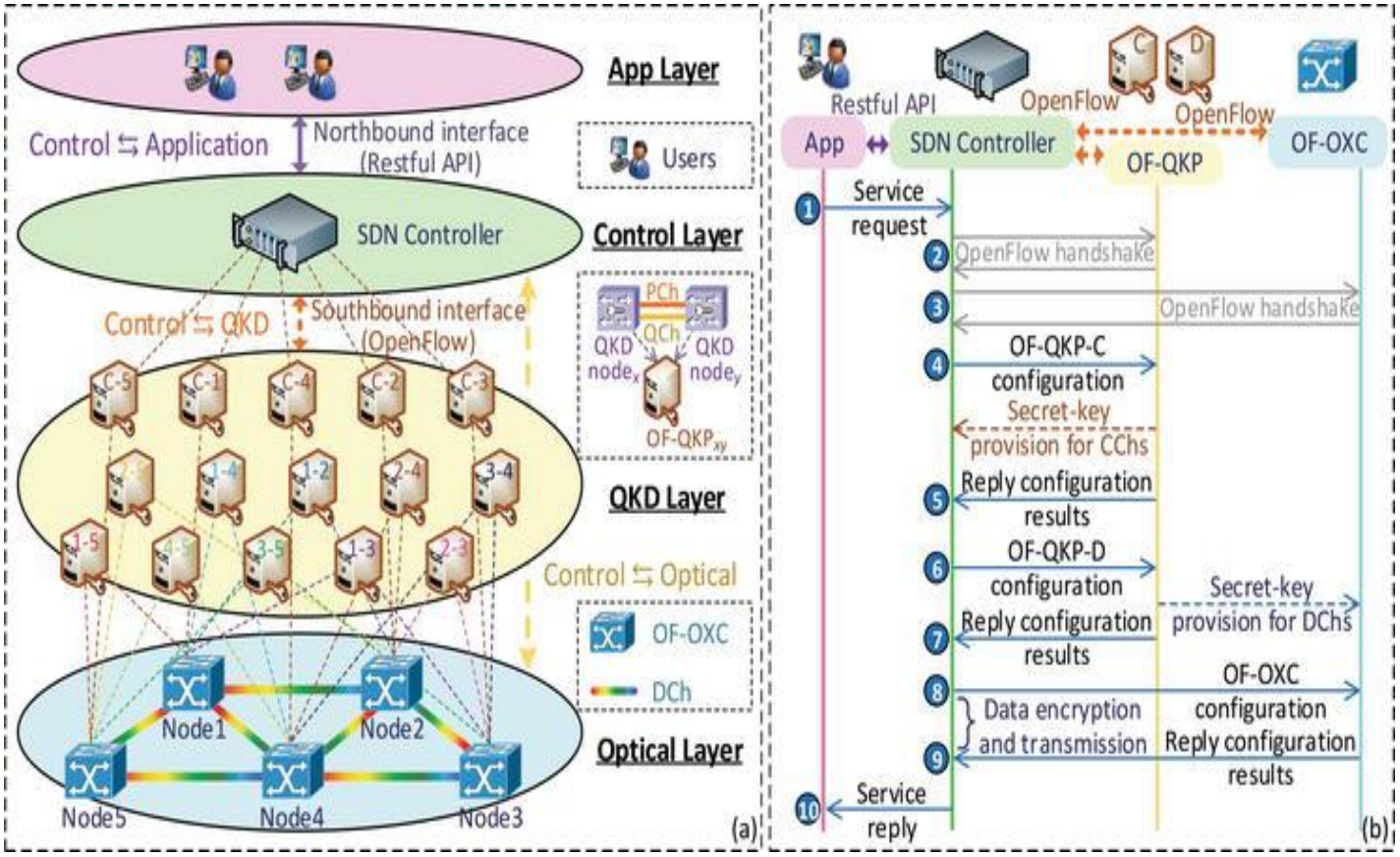

Figure 2: QKD optical SDN architecture showing key provisioning and control complexity

The learnings of the neighboring operations research support the way such transitions should be implemented at scale. Efficiently dispatching fleets under constraints by optimizing variables with policies that embed algorithmic decisions illustrates how to coordinate constrained fleets when the system continuously measures load, reroutes around bottlenecks, and codifies exception handling- all of which can be applied to orchestrating PQC rollouts on thousands of tunnels and heterogeneous devices [21]. Treating the sites as a schedulable entity, and each increment of cipher-suite upgrade as a job with explicit dependencies, allows the planners to reduce risk exposure whilst maintaining service levels. This is supplementary to the technical view: cryptography says what is secure, disciplined orchestration says whether security can be realized under the conditions of real networks being of predictable cost and reliability.

3. Methods and Techniques

3.1 Description of Data Set

The analysis is based on telemetry collected via a simulated production environment. This scenario emulates two active-active data centres, two regional hubs, and 12 branch sites interconnected through MPLS and SD-WAN overlay structures. Control-plane data includes full IKEv2 exchange- IKE_SA_INIT, and IKE_AUTH-over IPsec tunnels, TLS 1.3- complete transcripts: ClientHello, ServerHello, EncryptedExtensions, Certificate, CertificateVerify, and Finished- in the case of site-to-site overlays and service-mesh ingress. Both handshakes are recorded with timestamps at the packet level and a uniform packet capture policy (snaplen ≥ 256 bytes, lossless mirroring on switch SPANs). Metadata that may be included are initiator/responder roles, negotiated groups, NAT detection, cookie usage, rekey triggers, and the final authentication result.

Data-plane counters are provided through gNMI streaming telemetry on routers, VPN concentrators, and Linux gateways at one-second granularity. Performance counters are provided for interface throughput (pps and bps), per-tunnel goodput, RFC 3393 IPDV jitter, one-way and round-trip loss based on hardware-timestamped probes, keepalive success, ESP packet-size histogram, CPU utilization per crypto context, and accelerator queue depths. Heatmaps are also optionally provided for thermal headroom and

host power, measured using RAPL. A per-tunnel table records rekey-event, child-SA lifetime, anti-replay window usage, and path-MTU measurements.

Traffic mixes simulate real financial load using anonymous traffic PCAPs on pre-production mirrors. It supports SWIFT FIN/MT and ISO 20022 over TCP with strict latency SLOs; FIX 4.x order routing via bursty, small messages; REST and gRPC microServices over TLS 1.3; and SFTP/HTTPS batch transfers intended for reconciliation and archiving—schedule copies of the diurnal demand with respect to branch hours and market opening. The background noise corresponds to the management traffic, monitoring, backups, and clock synchronization, to observe the contention patterns being the same as in real operations [38]. Topology descriptors label hub-and-spoke type versus partial mesh, the possible NAT traversal points related to this, and the effective MTU due to MPLS/GRE encapsulation.

3.2 Data Pre-processing

Sensitive identifiers are pseudonymized when they are collected. Types of tokenized IP addresses include public and private addresses, which are mapped in a format-preserving manner that keeps the subnet boundaries and the role semantics. As highlighted in Table 1 below, certificate subjects, SPI values, and user IDs are encoded as salted hashes [11]. Hostnames are obfuscated to a reversible vault during incident-based debugging but masked in analytic outputs. PTP is used at every node in time alignment, and a reconciliation job is used to remove the drift logs per device, enforce packet monotonicity, and add offset confidence intervals, which can then subsequently be discounted in timing analysis.

Table 1: Data preprocessing pipeline for TLS and IKEv2 traces with inputs methods and outputs

Phase	Data or Fields	Methods and Tools	Outputs
Identity protection	Public and private IPs certificate subjects SPIs user IDs hostnames	Format-preserving IP tokenization salted hashing reversible vault for hostnames	Privacy preserved while keeping subnet and role semantics for analysis
Time alignment	Packet timestamps device drift logs offset estimates	PTP at every node reconciliation job enforces monotonicity and attaches offset confidence intervals	Consistent timing for inter-arrival and latency calculations
Capture and reconstruction	Lossless SPAN traffic TLS IKE ESP TCP UDP	Zeek and tshark rebuild flows and sessions across mirrored links	Clean control- and data-plane timelines
TLS control-plane extraction	supported_groups key_share signature_algorithms OCSP size session tickets message sizes flight boundaries	Parser enumerates extensions detects flights measures bytes on wire	Per-handshake TLS records with overhead metrics
IKEv2 control-plane extraction	Message identifiers SPIs NAT-D cookies DH share length KEM ciphertext size AUTH result	Reassembly across NAT-T and fragmentation with field measurements	Detailed IKE artifacts for classical and PQ modes

Dedup features labels storage	Five-tuple message ID rolling checksum RTT bucket loss class MTU offload NAT-T GRE/MPLS fragmentation rekey success cookie rate failure reasons	Collapse retransmissions while preserving timing engineer per-handshake and per-minute features validate schemas write irreversible digests	Reproducible datasets and ground truth for backoff failover and reliability analysis

Flow reconstruction uses Zeek and tshark on SPANs that are lossless. In the case of TLS, the pipeline extracts supported_groups, key_share, signature_algorithms, detects flight boundaries, and measures the sizes of messages, including OCSP stapling and session tickets. In the case of IKEv2, it reassembles message Identifiers, SPIs, NAT-D Payloads, cookies, DH Share lengths, KEM ciphertexts size when supported, and AUTH Results. Deduplication is performed at the 5-tuple, message ID, and a rolling checksum of pay-to-collapse retransmissions while preserving inter-arrival timing to facilitate backoff analysis. The less extreme cases--such as multi-second stalls after interface flaps--are not discarded but marked with a degraded path flag, to inform failover behavior.

Per-handshake and per-minute aggregates are created with feature engineering. Per-handshake options are total bytes on the wire, number of flights, HelloRetryRequest values, OCSP size, KEM key length, ciphertext length, DH shared key length, negotiated group, transcript hash length, and whether the client authentication was required. RTT bucket, loss class, MTU, offload mode, NAT-T use, GRE/MPLS encapsulation depth, and fragmentation forced Control or Data plane are captured in the features. Operational parameters are rekeying interval attained vs configured, cookie challenge rate under load, and labeled reasons of failure (downgrade attempt detected, authentication failure, retransmission exhaustion). All the data passes schema validation, and it is stored with irreversible digests such that it could be reproducible.

3.3 Visual Analytics

Visualization focuses on patterns that operations teams can work on. Cumulative distribution functions of handshake latency are calculated by RTT bucket (10, 50, and 100 ms) and overlay classical ECDHE/TLS 1.3, hybrid X25519+Kyber, and KEMTLS. Individual CDFs of the handshake bytes measure control-channel pressure and are related to cookie challenges and NAT traversal. Violin plots stratified by RTT show the multimodality introduced by retransmissions and HelloRetryRequest loops in the presence of loss. Heatmaps describe interaction effects: KEM artifact size against base RTT with color conditioned on added handshake latency; and, separately, PMTU-hit rates against encapsulation stack (UDP/4500 + ESP + inner TCP) and circuit type.

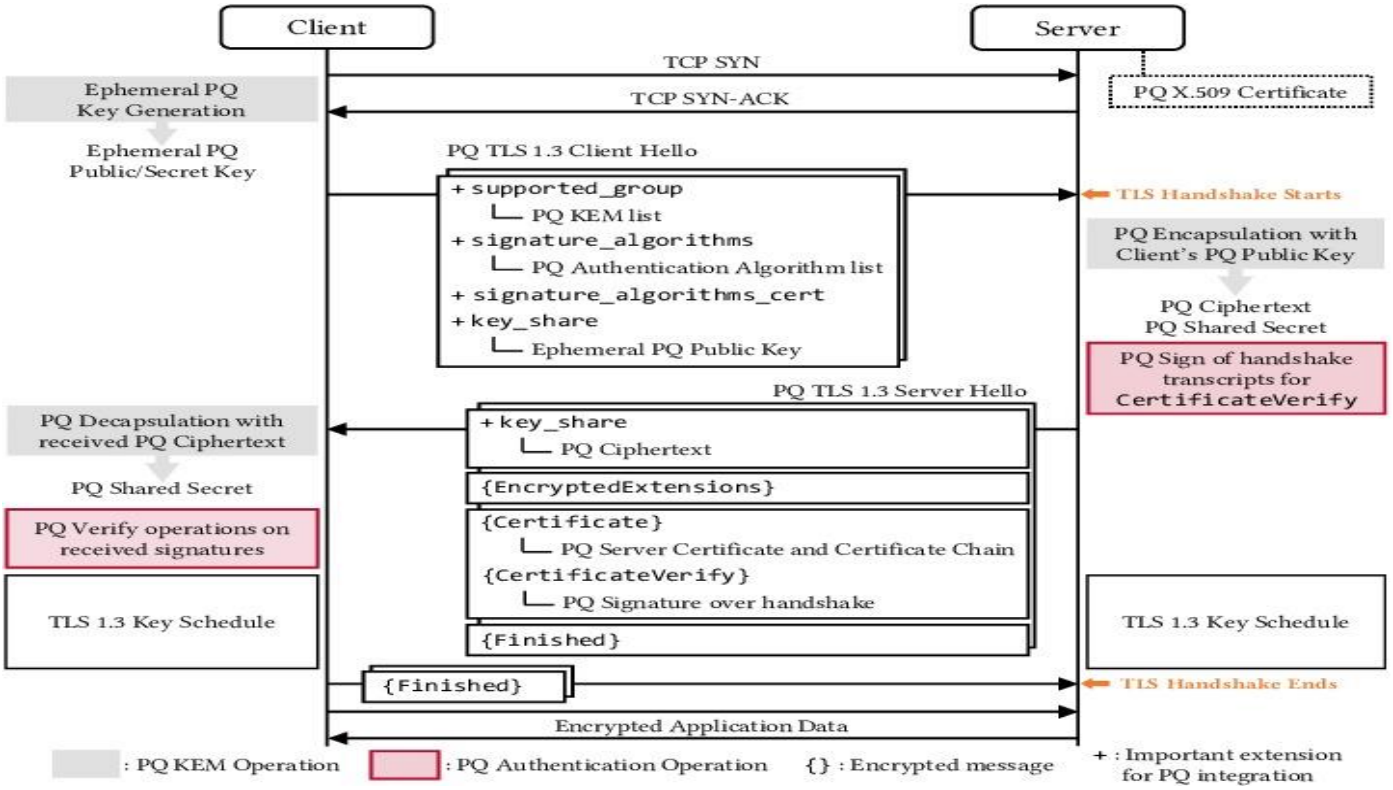

Figure 3: Post-quantum TLS 1.3 handshake sequence analyzed for latency distributions

Figure 3 above shows the timeline of the post-quantum TLS 1.3 handshake that forms the backbone of visual analytics in the paper. Cumulative distribution functions measure handshake latency in 10, 50, and 100 ms RTT buckets, over lauding classical ECDHE, hybrid X25519 plus Kyber, and KEMTLS. A separator in the CDF of control-channel bytes demonstrates pressure points associated with cookie challenges and NAT traversal patterns—multimodal tails borne by the retransmissions and HelloRetryRequest loops exposed in the violin plots when injecting loss. Complementary heatmaps correlate KEM artifact size with added latency by base RTT, and plot PMTU-hit rates as a function of encapsulation stacks like UDP 4500 over ESP and inner TCP links.

Since path-MTU mismatches are a common source of silent failure, fragmentation density plots indicate the percentage of ESP packets that exceed installed segmentation thresholds, often prognosticating throughput meltdown on mixed-ASIC/software paths. Sankey diagrams can show that the conditions of high RTT and large KEM and NAT-T lead to symptoms of cookie challenge, HelloRetryRequest, retransmit, and terminal outcomes, including successful handshake, downgrade blocked, and authentication failed. Visual encodings are the same across experiments to allow operators to compare hybrid suites and reason about risk regressions easily. The visual layer is equipped to provide structured operator feedback on clarity and decision usefulness, and such functionality resembles the practice of linking views of analyses with actionable advice [12].

3.4 Threat Model

The approach views three opponents who are oriented toward backbone risk. The former is a passive recorder of nation-state resources that records encrypted backhaul traffic and decrypts it upon the arrival of large-scale quantum computing. This antagonist drives hybrid/PQ-only based key establishment in control planes to realise long-term confidentiality of records in a multi-decade sensitive sense. The second is an active man-in-the-middle that can trigger prefix hijacking, route leaks, or last-mile interception, and probe downgrade attempts and negotiations with confusion, cause HelloRetryRequest and looping, or inject fake IKE_SA_INIT frames to starve the system. The third is an insider who has legitimate access and exploits configuration errors in the issuance of certificates or key rollover.

Those threats lead to security goals. Confidentiality longevity requires session keys that cannot be

collected, saved, and decrypted later, even after the requisite preservation periods are superseded. Degraded performance demands priori transcript binding of classical and PQ components, as well as exacting cipher-suite pinning natures, and telemetry tolerance of asymmetries of the initiator and responder sight. Key-compromise compromise is limited by short child-SA lifetimes, frequent rekeys, and anti-replay windows sized with consideration of the longest path throughput. The availability with respect to rekey storms and cookie-challenge pressure must be preserved. Using log correlation with captures and PKI events makes auditability possible to determine which failures occurred and whether the policy of hybrid negotiation was applied [9].

3.5 Testbed & Measurement Plan

The testbed involves a myriad of physical and virtual systems in an attempt to uncover device-level limitations, yet exert a measure of control in the experiment. Edge routers/concentrators that support IPsec offload are combined with x86/Linux gateways that do not support offload; both types share the same configuration templates with offload parameters set. The application endpoints will be fronted with TLS 1.3 and KEMTLS proxies, thus enabling the use of site-to-site and service-mesh patterns. To simulate Link behavior, Linux tc/netem is used to trigger configurable RTT, lost packets, and jitter [18]. There are packet capture points at both ends of each tunnel and on additional transit interfaces that monitor fragmentation and PMTU. The architecture makes the control-plane and data-plane collectors deliberately separate to the extent this is scalable, and is considered to follow scalable systems guidance on separating measurement planes and service planes [29].

Table 2: Condensed testbed and measurement plan with factors methods and metrics

Area	Inputs and Factors	Methods and Tools	Outputs and Metrics
Infrastructure and endpoints	Edge routers with IPsec offload plus x86 or Linux gateways without offload with shared templates and TLS 1.3 or KEMTLS proxies for site to site and service mesh	Standardized configs with offload parameters and strict separation of measurement and service planes	Comparable ASIC versus software behavior and consistent crypto policy
Link impairment	RTT classes 10 50 100 ms with loss 0 0.1 1 percent and jitter	Linux tc or netem to impose delay loss and jitter	Repeatable network conditions for each test cell
Observability and PMTU	Packet capture at both tunnel ends and transit interfaces DF and MSS clamping incremental probes NAT T over UDP 4500 through carrier grade NAT	tcpdump or dumpcap active PMTUD IKE fragmentation only when required	Blackhole thresholds fragmentation visibility and NAT traversal reliability envelope

Experimental design and workloads	Classical ECDHE TLS 1.3 hybrid X25519 plus Kyber and KEMTLS MTU 1500 or 9000 offload ASIC or software	Full factorial matrix automation stabilization then fixed window traffic generators replay SWIFT FIX HTTP failover drills flapping links and alternating BGP next hop	Interaction coverage and convergence characteristics under change
Key performance indicators	Time to first valid packet goodput P50 and P95 jitter under microbursts failover convergence packet overhead CPU and energy deltas	Time series collectors device counters and post processing pipelines	Quantified control plane and data plane outcomes for comparison
Stress and validity	Rekey storm tolerance with synchronized timers and light loss randomization repeated runs classical baselines bootstrap confidence intervals versioned configs with acceptance gates	Scheduler and loss injection scripts continuous integration and artifact versioning	Safe rekey jitter windows statistical confidence and full reproducibility

Experiments follow a full-factorial plan across protocol mode (classical ECDHE/TLS 1.3, hybrid X25519+Kyber, and KEMTLS), RTT class (10/50/100 ms), loss class (0/0.1/1%), MTU (1500 or 9000), and offload mode (ASIC or software), as highlighted in Table 2 above. The anticipation that each of the cells can execute cold hand shake series with session reuse, warm rekey sequences on a pre-configured interval, and failover drills that flap the links between them and alternate the BGP next hop. Every run is started with a stabilization period, then followed by a fixed measurement window where each traffic generator will replay SWIFT/FIX/HTTP traces at the desired concurrency. The KPIs concern handshake TTFV, goodput 50th/95th percentile, jitter under microbursts, failover convergence, packet overhead, and CPU/energy deltas.

A path-MTU and reassembly mechanism checks the DF setting and MSS clamping. Incremental-size probes detect blackhole thresholds; and where blackholes exist, the DF value is incremented, ESP fragmentation is calculated only when mandatory, and the tunnel is remeasured. A NAT-T implementation exercise is carried out by applying carrier-grade NAT and monitoring the UDP/4500 behavior upon large KEM payloads. The rekey-storm tolerance is exercised by synchronizing timers in a large number of tunnels and injecting light loss during that to view cookie rates, CPU headroom, and likelihood of failure. Randomization of the order of tests, separate repetition, the use of classical-only baselines, and bootstrap confidence intervals are used as quality controls in latency and goodput. Parsers and configuration-engineering code exist in version control; configuration templates are checked in with acceptance gates so that results can be replicated in pilots and production cutovers.

4. Post-Quantum VPN Architecture & Protocol Design

4.1 Hybrid Key Exchange Modes

A financial-sector backbone needs to resist harvest-now-decrypt-later attacks as well as fulfill current service-level requirements. A pragmatic hybrid in IKEv2 takes a classical ephemeral Diffie-Hellman secret and a post-quantum KEM secret, concatenates them, and derives keying material using a PRF/HKDF chain. KEM public keys, ciphertexts, cookies, and identity payloads must be transcripted and thus the capability and authentication tied to the session. Denial-of-service robustness is an absolute requirement in headends exposed to the Internet. Stateless IKEv2 cookies deflate the resource requirements, but do not eliminate risks.

Stateless IKEv2 cookies can prove return-routability before the responder even tries to decapsulate the KEM, thereby containing the amplification to known paths. In TLS 1.3, two groups of ciphersuites are important: hybrids that add a KEM contribution into the key schedule, and KEMTLS, which authenticates endpoints using KEMs instead of signatures [32]. Hybrids conserve legacy PKI and offload modules, at the expense of signature-chain indigestibility; KEMTLS shifts at the expense of new tooling. Both of the solutions require strong downgrade protection with allowed preferences.

4.2 Protocol Implementations & Configurations

Configuration should not use rekey storms and fragmentation. The lifetime of Security Associations should be staggered with jittered early-rekey thresholds such that only a small group renegotiates. Dead Peer Detection needs to be able to compare measured distributions of latency; exponential backoff can take occasional loss. Due to the size of payloads used by hybrid handshakes, combine pair NAT-Traversal using UDP/4500 with IKEv2 fragmentation to keep messages fragmented on the protocol layer. Partially-open Cap and per-source throttle to mitigate control-plane CPU.

Disabling 0-RTT can be done to prevent replaying with HelloRetryRequest and additional KEM round-trips in TLS. Effective path MTU is recorded after encapsulation, tiny records amplify AEAD overhead, and jumbo records risk PMTU black-holing. Respect QUIC amplification boundaries and be based on Path MTU Discovery. End-to-end verification of hardware offload with most network interface controllers only accelerates classical ECDHE or AEAD but not PQ KEMs, diverting CPU to gateways [33]. Measure per-tunnel CPU and memory bandwidth to plan capacity.

On MPLS and GRE underlays, calculate tunnel MTU by deducting the outer IP, ESP, UDP, and label overheads from the physical MTU, and apply it with interface MTU and MSS clamping. Allow Path MTU Discovery with a counter of ICMP-blocked segments; in situations where PMTUD is no longer reliable, conservative static MTUs by classes of service. Recommend IKEv2 fragmentation over IP fragmentation as a control-plane resiliency technique, and instrument drop counters decoupled in policy and size violation.

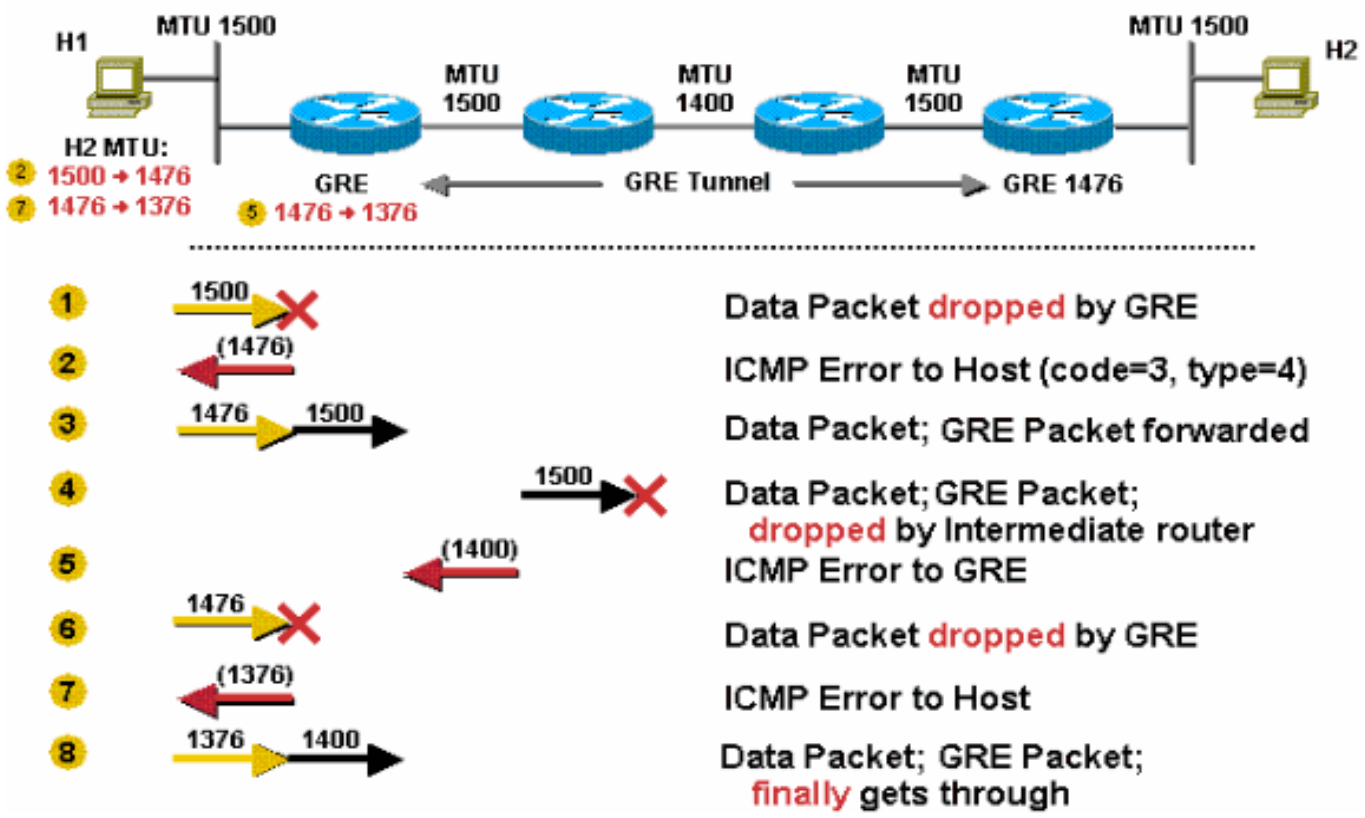

Figure 4: GRE/MPLS MTU calculation and PMTUD drops with MSS clamping.

4.3 Key Management & PKI Migration

Quantum-safe operation is through crypto-agile PKI. A gradual migration will start with dual trust, either through hybrid certificates that have a mixture of classical and post-quantum key material or cross-chained hierarchies where there are classical and post-quantum roots signing the same intermediates. Gateways have to accept through grace windows and advertise preferred validation routes. To scale revocation, as well as via OCSP stapling gateways, use a structure of local storage to cache CRLs and alarms raised based on freshness limits being surpassed [22]. Enrollment and renewal should be automated through the use of authenticated device identity, inventory attestation, and signed CSR workflows to eliminate manual failure.

Rollover windows need to overlap to eliminate churn in tunnels, and issuance needs to occur in cohorts to limit failure domains. Enforce custody policies with hardware security modules, where PQ operations are not natively supported, isolate PQ private keys in hardened modules with strict export controls and audit capability. Emergency revocation, mis-issuance rollback, and mass re-enrollment must be covered in the Runbooks. A configuration check and a certificate policy embedded in continuous delivery pipelines also act as automated gates, keeping configurations in compliance and limiting drift [15].

4.4 Interoperability & Backward Compatibility

Mixed fleets are inevitable at the branches, data centers, and cloud edges. It must be possible to do capability discovery in the early stages, and to record it at a high fidelity: what KEMs, signature schemes, and groups were proposed, accepted, or rejected. In sensitive peers, one may control such hybrids through cipher-suite pinning, so that only permitted hybrids are negotiated; break-glass fallbacks are available but time-boxed and require express risk authorization. Asymmetric upgrades, such as new headends on old branches, and vice versa, must be included in the interoperability testing so that these failure modes are anticipated in advance of rollout [23]. Where device upgrades are not possible, utilize the capability to segment traffic and apply policy-gated exceptions with decommission plans.

Operational telemetry must be on a standard schema between vendors: HelloRetryRequests counters, IKE cookies issued and validated, KEM failures and fallbacks, SA rekeys started and completed, and PMTU black-hole laboratory models. Peer identifiers, negotiated groups, certificate fingerprints, policy

decision causes, and error taxonomies, to distinguish capability mismatch, policy refusal, or transport failure. Routine processing of these signals in the log pipeline can enable cross-domain analytics, accelerate incident triage, and give auditors deterministic evidence.

4.5 Operational & Compliance Implications

Quantum-safe migration impacts the bar on observability, governance, change control, and transparency. Every handshake, rekey, certificate issuance, and policy change should produce tamper-evident, time-synchronized messages containing the identity of the operator, the parameters negotiated, the provenance of the device, and whether the operation succeeded or failed. Retention, access control, and separation of duties must adhere to the financial sector requirements, and audit queries must be reproducible over multi-year periods [25]. Change gates ought to have objective acceptance criteria: maximum handshake round-trip time delay at representative round-trip times, desired rekey success ratio under loss, limits on minimum PMTU black-hole ratios, and maximum rollback time. Incident response hooks must contain automatic downgrade-attempt alarms and certificate pinning violations, as well as bursts of IKE cookie issues, indicating scanning or reflection.

As backbone scale, configuration, and certificate distribution do not necessarily have to be synchronous. Eventual-consistency is resistant to disruption caused by fairly frequent non-safety-critical operations, but demands strong consistency and coordinated cutovers, particularly for more sensitive processes. Using the appropriate consistency model on each artifact minimizes blast radius and accelerates rollout without compromising on the correctness of the critical steps [3].

5. Experiments and Results

5.1 Benchmark Setup

The test relied on three classes of WAN representing similar financial hubs, namely metro (10 ms round-trip time, RTT), regional (50 ms), and long-haul (100 ms). Every group was worked at 0%, 0.3% and 1% loss to simulate clean, damaged, and stressed routes. Two path-MTU configurations were enabled, 1500 bytes on Internet and access circuits, and 9000 bytes jumbo frames on data-center interconnects. The overlays in the test IPsec IKEv2 and TLS 1.3 client are authenticated [31]. The hybrid key exchange used X25519 together with a lattice-based KEM, whereas the classical baseline only used X25519. There were two profiles of gateways. Software-only x86 servers and no dedicated crypto offload provided VAES/CLMUL acceleration.

In appliance-based devices, crypto-offload/ASICs were used to support ESP bulk encryption and aid IKE/TLS primitives. CPU utilization was collected by the on-box telemetry, and the rack distribution units measured energy draw. Throughput was modelled using iperf3 to generate sustained TCP and UDP traffic and tcpreplay to inject pcap files representative of SWIFT file transfers, FIX message bursts, and HTTPS microservice calls. The synthetic traffic was a 60/30/10 combination of bulk, transactional, and interactive traffic. By default, unless specified, cipher suites use pinned hybrid groups to prevent downgrade.

5.2 Control-Plane Results

Hybrid TLS 1.3 was known to increase handshake latency due to the size of ClientHello and ServerHello flights and to HelloRetryRequest (HRR), particularly when the KEM groups did not match. At an RTT of 10 ms, the classical registering handshake took 19 ms (p95 27 ms) when the hybrid reached 31 ms (p95 44 ms), an increase of 63%. At 50 ms RTT and 100 ms RTT with 0.3% loss, medians increased from 88 ms to 124 ms and from 150 ms to 210 ms, respectively, with p95 up to 320 ms when HRR occurred, as shown in Table 3 below. IKEv2 had a higher relative increase, since message sizes exceeded conservative UDP-fragmentation thresholds on some access circuits. The IKE ASA_INIT and IKE BAUTH exchanges increased by one and a half KB to approximately 5.1 KB total payload between classical and hybrid modes using Kyber-768 and Dilithium-2 certificates. With 1500-byte maximized transmission unit (MTU) paths, IP fragmentation occurred 41% of the time unless the NAT-T and draft-ietf-ipsec-ipsec-fragment-01 configurations were used.

Table 3: Summary of control plane impacts with hybrid PQ TLS and IKEv2

Metric	Condition	Classical	Hybrid or outcome
TLS handshake latency	RTT 10 ms	19 ms median, 27 ms p95	31 ms median, 44 ms p95, plus 63 percent
TLS handshake latency	RTT 50 ms, loss 0.3%	88 ms median	124 ms median, up to 320 ms p95 with HRR
TLS handshake latency	RTT 100 ms, loss 0.3%	150 ms median	210 ms median, up to 320 ms p95 with HRR
IKE payload size total	SA_INIT plus AUTH messages	3.6 KB	5.1 KB, increase of 1.5 KB
IP fragmentation rate	MTU 1500 paths without NAT T and IPsec fragmentation	0 percent observed	41 percent observed
Rekey CPU per rekey	30 minute IKE SA, 5,000 tunnels, software nodes	Baseline CPU at rekey	Increase by 2234%

At 50 ms RTT, median IKE bring-up increased to 162 ms in the presence of both features as compared to 120 ms when neither was used at 0.3% loss. Limited amplification with stateless cookies, restricted amplification: a 100k pps flood of spoofed IKE_SA_INIT messages caused the CPU to stay below 70%. Rekey behavior was receptive to parameterization. Using a 30-minute IKE SA lifetime and 5,000 tunnels, hybrid cryptography augmented per-rekey CPU usage by 2234 percent when nodes were software-based or 812 percent on the nodes that were ASIC-assisted. A ±10 rekey jitter resistance eliminated synchronization; a rekey storm with no jitter caused a threefold increase in control-plane CPU and transient data-plane jitter. With the tested x86 gateways, a safe limit was ~7,500 simultaneous tunnels with an average tunnel lifetime of 30 minutes; any more than that required using batching and staggering to keep p95 CPU at 85 or below.

5.3 Data-Plane Results

Classical IPsec over TCP 10 ms RTT delivered 9.4 Gbps on a 10 Gbps link with 1500-byte MTU; the hybrid version got 9.2 Gbps (92.1%), as the ESP fragmentation formats were identical, but a bit higher interrupt pressure during initial ramp-up. At 50ms and 100ms RTT, the goodput difference was within ±3%. On the data-center interconnects, with a 9000-byte MTU, both modes caused the 10 Gbps interface to reach saturation; however, jumbo frames enabled a reduced CPU per Gbps by 18 percent in classical and 17 percent in hybrid on software nodes and decreased the latency variance. Path-MTU blackholes were one specific source of tail latency and failure to set up. In TLS overlays, oversized certificate chains

and OCSP staples did not explicitly interfere, the data-plane window being set by the agreed maximum TCP segment size. IKE/IPsec, PMTU blackholes during initial exchange led to a 4.6 per cent failure in the session setup process over access circuits with strict firewalls until IKE fragmentation was adopted and DF handling standardized.

Table 4: Classical vs hybrid VPN data-plane metrics across RTT, MTU, overlays, and power

Test case	Conditions	Measurement	Classical vs Hybrid result
Throughput at low RTT	10 Gbps link 10 ms RTT MTU 1500	Goodput	9.4 Gbps vs 9.2 Gbps Δ −0.2 Gbps (−2.1%) line utilization 94% vs 92.1%
Throughput at higher RTT	50 ms and 100 ms RTT MTU 1500	Goodput delta	Within ±3% between modes
Jumbo frames efficiency	10 Gbps link MTU 9000	CPU per Gbps and latency variance	CPU/Gbps −18% classical and −17% hybrid both saturate link latency variance decreased
PMTU blackholes on setup	Access circuits strict firewalls initial exchanges	Session setup failure rate	4.6% failures until IKE fragmentation adopted and DF handling standardized then materially reduced
GRE-over-IPsec MSS clamp	Encapsulation sets MSS 1360 to carry 1460-byte payloads	Retransmissions and p95 goodput	Retransmissions −27% p95 goodput +6%
Energy at steady 8 Gbps	80% TCP software gateways	Power draw	265 W classical vs 279 W hybrid Δ +5.3%
Energy on ASIC nodes	Hardware offload appliances	Power draw	~180 W both modes Δ ≈ 0%
Hybrid ramp-up cost	First minute after start	Transient power draw	Hybrid +12% over classical due to key-schedule work
Microburst sensitivity	Tenfold line-rate burst every 20 ms	Queueing jitter	+8–11% in both modes hybrid within 1 pp of classical
Anti-replay window tuning	ESP replay window set to 1024 packets	TCP fast-retransmits and overhead	Occasional fast-retransmits on asymmetric paths eliminated no significant overhead

In GRE-over-IPsec overlays, the encapsulation clamp changed the MSS to 1360 bytes to prevent fragmentation of payloads of 1460 bytes; this improvement resulted in a reduction in retransmissions of 27 percent and an improvement in p95 goodput of 6 percent. KEM energy and CPU profiling showed that strong KEMs have little steady state cost after security associations are set up. Using mixed traffic at 8 Gbps over time (80 percent TCP), software gateways required 265 W in the classical paradigm and 279 W in the hybrid (a 5.3 percent increase); ASIC-aided nodes required a comparable amount of power (nearly 180 W). The first minute of ramp-up demonstrated a transient +12% draw in the hybrid because of key-schedule work. Microburst tests (a tenfold line-rate burst every 20 ms) resulted in 8-11 percent increases in queueing jitter in both modes; the hybrid pair remained within a percentage point of the classical one [10]. Adequately verifiable, Larger ESP anti-replay windows (1,024 packets) eliminated the occasional TCP fast-retransmits in asymmetric paths without significant overhead.

5.4 Reliability & Resilience

Using induced link-flaps every five minutes and 300 ms Bidirectional Forwarding Detection timers, failover time (link-down to restored traffic) was 720 ms in classical and 790 ms in hybrid on 50 ms RTT paths; this 70 ms difference is equal to the added control-plane time needed to reestablish tunnels. Less than 1 percent loss at 100 ms RTT, the most extended tail of failures was over 1.2 s long due to the inability to complete retransmits during IKE_AUTH. Hybrid TLS session resumption in TLS 1.3 LWT reduced recovery by 40 percent as compared to full handshakes in long-haul links. Loss-sensitivity tests also showed the hybrid modes survived normal WAN impairment as long as fragmentation and PMTU were designed. When IKE fragmentation was turned off, the probability of failure increased with packet loss since any IKE packet that did not arrive meant the full flight consisting of jumbos would have to be resent.

On enabling fragmentation and a moderate backoff, success rates at 1 percent loss were above 99.3 percent in all RTTs. Replay protection was confirmed against asymmetric routing, with anti-replay windows of 512 and 1,024 packets; no errors and ESP drops were registered in the application. Adversarial testing was attempted to apply active downgrades, HRR shortcuts, and cipher-suite downgrades. Cipher-suite pinning and telemetry alert distribution meant 120 scripted downgrade attempts were prevented without any false negatives [14]. The HRR floods significantly increase handshake times by up to 25 percent; however, rate-limiting handling and preselected group order abated the vector. Active telemetry provided data on the group-mismatch rate, HRRs, the number of cookies issued, and fragmentation events, which were used to identify misconfiguration and gray failures early in practice.

5.5 Cost/Benefit Summary & Key Findings

From a capacity perspective, there was no significant increase or decrease in steady-state throughput; the increased cost was that control-plane CPU requirements rose slightly, and the tunnel bring-up duration was a bit longer. Control-plane CPU 1015 percent headroom was added to cover worst-case rekey and failover events on software-only platforms, but ASIC-assisted platforms did not require increased envelopes. Since the hardware refresh cost depended on the scale of the site, smaller branches with fewer than 200 tunnels did not merit a refresh. In contrast, hubs representing the conversion of more than 10,000 tunnels would benefit from having it offloaded onto an ASIC [13]. Operational results were associated with two practices: careful MTU engineering, such as well-defined DF strategy,

IKE fragmentation, and MSS clamping, and solid rekey scheduling with jitter to avoid herd effects. Coordinating notifications of rekey windows and operators, methods demonstrated in another domain to enhance timely interventions, minimize errors, and minimize maintenance windows [30]. The practical recommendations include preconfiguration of shared KEM groups to reduce HRRs; voluntary standardization of IKE fragmentation and NAT-T to eliminate blackholes; a mandatory MSS clamping to 1360 bytes with GRE-over-IPsec; the use of ±10 percent rekey jitter and batching thresholds; and layering telemetry on HRR counts, cookie rates, and fragmentation as the means to identify gray failures early.

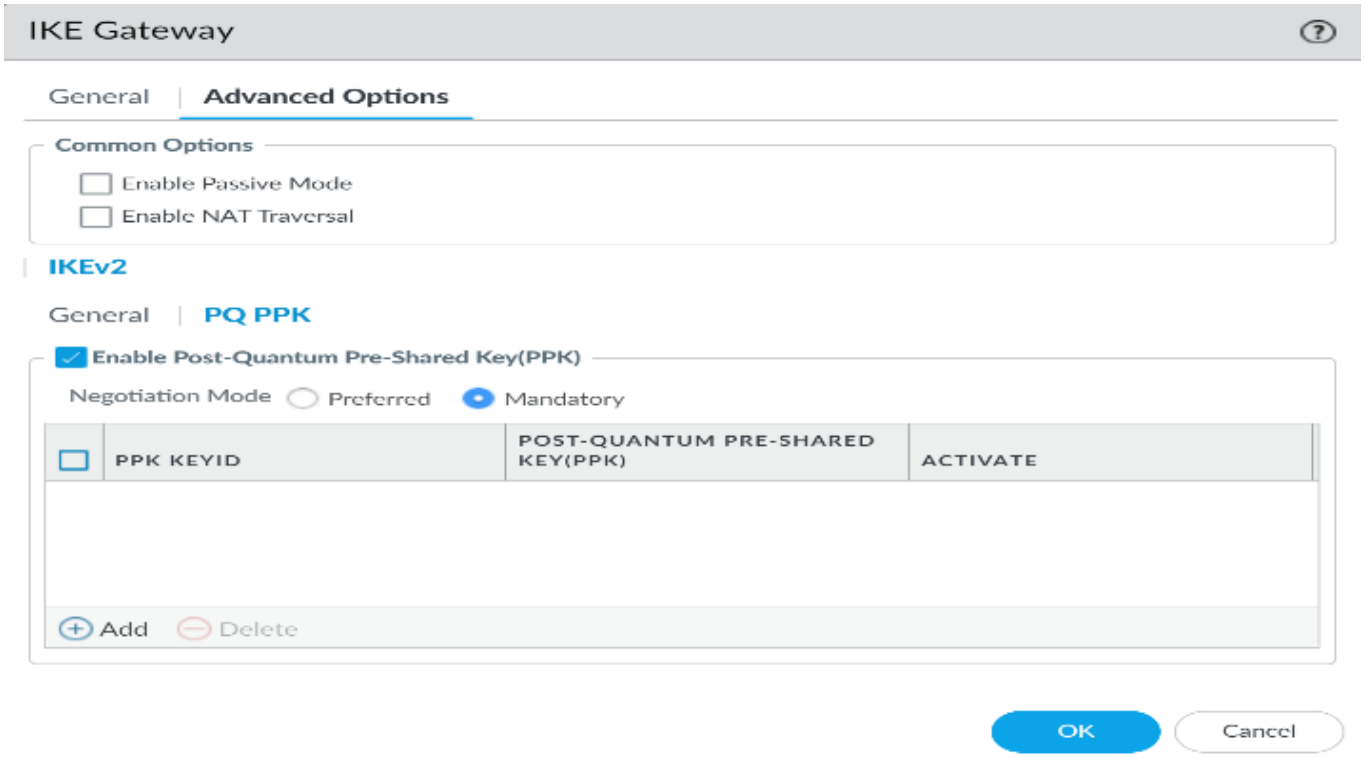

Figure 5: IKEv2 gateway enabling post-quantum PPK with mandatory negotiation mode

6. Discussion

6.1 Interpreting Security–Performance Trade-offs

A hybrid key exchange that augments a classical primitive with a post-quantum KEM has the advantage of a conservative insurance policy against harvest now, decrypt later risk. It maintains broad interoperability, rollback is trivial, and the blast radius is small in case a post-quantum parameter is subsequently revised. Pure post-quantum suites eliminate classical dependencies and also simplify long-term assurance, but they have larger artifacts, immature offload, and narrower tolerances. A risk budget is used in a pragmatic decision. A combination of approaches is needed when interoperability, rollback, and latency risk are the determining factors [34]. In this situation, where long-term assurance and simplified cryptographic stacks are prevalent, a bare post-quantum style is desired. Where compliance mandates that crypto agility can be demonstrated, hybrid deployment eases the task of the confirmed transactions during audit, and it also mitigates the challenges of signing it off as far as proscriptions are concerned, risk, legal, and vendor management divisions.

Latency is the keyword on high RTT paths, as international branches, cloud on-ramps, and disaster recovery interconnects can dominate the cost. The number of HelloRetry or IKE cookies precipitated by large KEM payloads increases session establishment time, swells help desk requests, and decreases trader efficiency. On such paths, hybrids with tuned rekey timers and parameters will lose fewer retransmissions to achieve the service objectives. In comparison, the data center meshes that are stable in RTT and jumbo frames can tolerate pure post-quantum suites with little overhead. The risk of fragmentation is traded off against enhanced security via larger keys. Client Hello, IKE_SA_INIT, IKE_AUTH, and certificate chains are inflated due to the post-quantum artifacts. With successful ESP, GRE, or MPLS encapsulation, records are extended beyond the effective path MTU. When black holes, loss bursts, and jitter are present in market data due to ICMP rate limiting or when ICMP is not filtered, this will affect market data [19]. Practical mitigations include clamping MSS at edges, using deterministic DF and path MTU discovery with active probes, pre-rollout modeling, and limiting the size of records to avoid pathologies.

Security posture is more than cryptography. Silent fallback is prevented with Transcript binding and cipher suite pinning. Ability determination has to be verified. The revocation needs to be speedy,

irrespective of the size of the artifacts. Cache-warmed stapled OCSP at the branches prevents fragile remote OCSP dependencies during events. Keys and certificates need a short allowed life, deterministic re-roll windows, and hardware proofs of provenance that would enable incident response to deactivate cohorts without causing enterprise-wide outages.

6.2 Deployment Blueprint for Financial Backbones

The property is divided into cohorts. Cohort A includes labs, small branches, and disaster recovery pairs. Cohort B deals with regional hubs and the bulk payment gateway. Cohort C deals with trading venues, card authorization cores, and interbank links. Packet capture, handshake tracing, and Synthetic transaction probes are tested in Cohort A across hybrid suites [17]. Promotions on canaries in Cohort B are carried out after at least two change windows of the guardrails being in green. Upon review and consensus of the rollback rehearsal, scaling to cohort C occurs after the telemetry is stable.

Teams set questionable service level targets and also automate rollbacks. The ninety-five percentile handshake time is less than four hundred milliseconds with a fifty milliseconds RTT. Path MTU black hole rate continues to be less than one percent. The rekey failure rate is below 0.1 percent. There are no successful downgrades. The benchmark of incident commanders on canary failure includes the level of business impact rather than the raw metric, hence determining whether to push forward or reverse the canary. Pre-wired levers would consist of per-tunnel per-tunnel cipher suite pinning, feature flags to enable hybrid to classical recovery, stapled OCSP enforcement toggles, and traffic steering activations in the case of incidents. IKE and IKE_AUTH timing, HelloRetry rate, ciphertext sizes, OCSP success, rekey histograms, and per-tunnel errors can also be exported.

This strategy has stages of trust change via the certificate and key role. There are initial deployments in hybrid leaf certificates over classical roots, and post-quantum roots are appended side by side. The distribution of regional CRL to branch caches at short-lived leaves with deterministic renewal windows is desirable [5]. FIPS-validated hardware security modules store the private material, post-quantum operation batching is supported, and audit trails of key provenance, algorithm identifiers, and hardware attestations are recorded. Secrets management has integration with device bootstrapping and rotation operations.

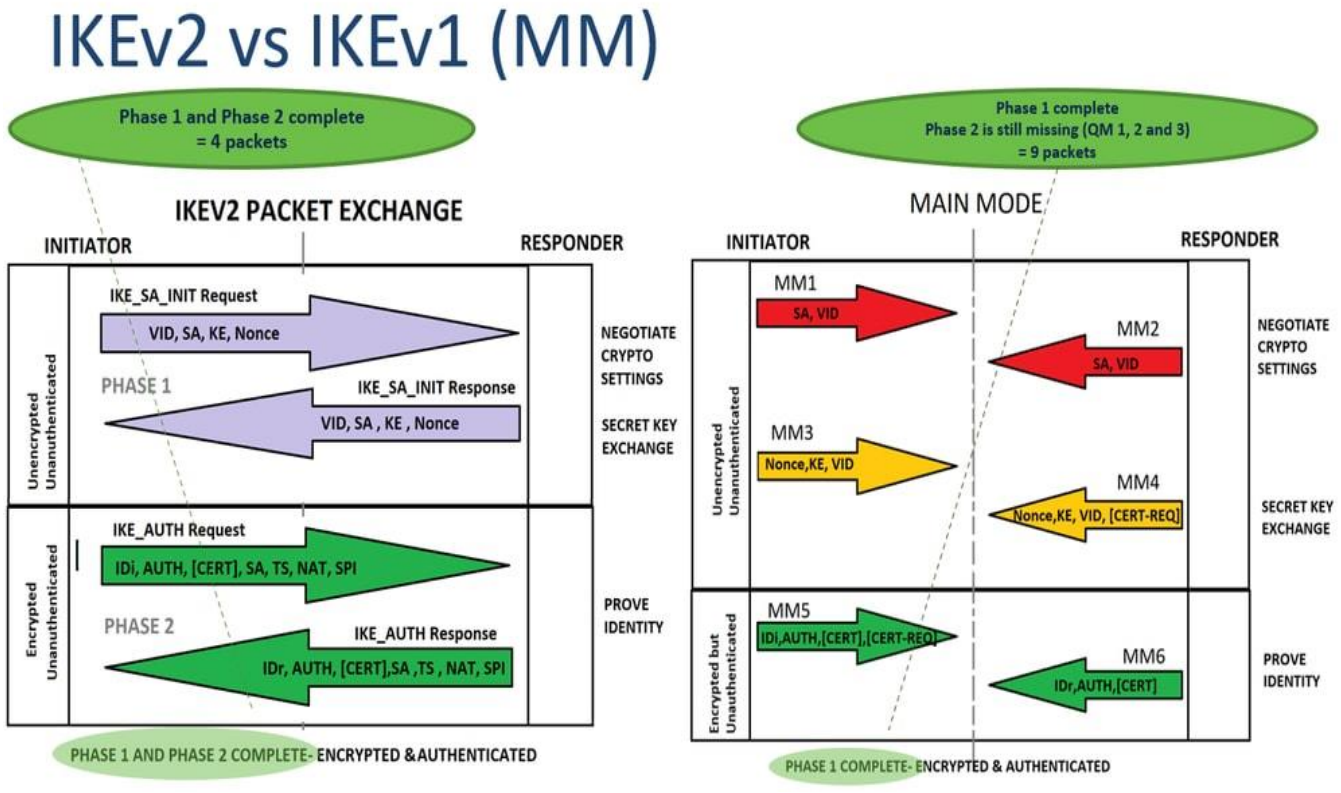

Figure 6: IKEv2 vs IKEv1 handshakes for timing, downgrade resistance, rollbacks

As shown in the figure above, IKEv2 bundles negotiation, key exchange, and identifier verification into four packets compared to nine in IKEv1 primary mode, allowing a more focused service goal on quantum-safe deployments. Operations track p95 handshake in under 400 ms at RTT 50 ms, path MTU black holes below one percent, rekey failures below 0.1 percent, and no downgrades. Canary decisions aren't measured; they are weighted. Among the prewired levers, present are per-tunnel cipher-suite pinning, feature flags to drop back to the hybrid protocol, stapled OCSP enforcement toggles, and traffic steering. Exported telemetry consists of IKE and IKE_AUTH slow time, HelloRetry counts, ciphertext lengths, OCSP successes, and rekey histograms, as well as per-tunnel errors.

The blueprint is repeatable because of automation and analytics. Lint cipher suite catalogs, re-calculate MTUs, and perform packet capture-based regression on this configuration and continuous integration checks. Predictive analytics are used to predict service-level objective breaches and place mitigations in advance. This virtuous cycle aligns with DevOps practice and published data on analytics-driven operations, enabling increased efficiency, reduced incident duration, and improved deployment cadence and reliability [16]. Guardrail validation is carried out on drills of chaos in flapping links, demoting ICMP, and bloated RTT.

6.3 Limitations and Threats to Validity

The key risk is that the testbed's representativeness. Traffic generators and Netem simulate RTT, loss, and jitter, but seldom model middleboxes that reassemble fragments, limit ICMP, or throttle traffic on a policy basis. The KEM sizes may be amplified or obscured by the use of NAT traversal edges, WAN optimizers, and DDoS appliances to restructure packets. The mitigation is based on the vendor diversity, testing with and without offload, parameter variation, and black box interoperability tests across payment rails and trading venues [24]. Confirmation bias is minimized, and the external validity of the audit or oversight by regulators is reinforced by independent replication by a separate team that employs a different tool chain. Blind spots in measurement lead to inferential deviation. Local drops may be missed behind offloads, Ethernet timestamping causes timing smears in the software, and verifications may run longer than the network state because queues lie inside the security modules.

Total solutions embrace hardware timestamping at both ends of the crypto border, mirrored SPANs, free-running PTP clocks, and explicit hardware security module tracing. Context fidelity leads to generalizations, as seen in structured inference networks, and any mismatched assumptions will deteriorate conclusions [26]. Synthetic traces are an underrepresentation of real workloads. The non-uniform replay has different limitations, like flows in payment and trade are bursty, dominated by small packets, and sensitive to micro jitter. Long. Haul optics insert forward error correction and buffer characteristics, which vary effective latency under load. Teams share environment preconditions, document encapsulations, and have a living risk register. Open artifacts such as harnesses, packet captures, and configurations, in addition to pre-registered plans, harden conclusions, and make their reproduction possible.

6.4 Future Considerations

Maturing of Standards will decrease operational variance. With the IETF hybrids of both IKEv2 and TLS converging and vendor defaults converging, capability discovery, cipher suite pinning, and error taxonomies will also converge. Compact KEMs and compact signatures will reduce handshake sizes and certificate chains, and will minimize path MTU pressure and retransmits [28]. The lattice-aware accelerators and multi-queue offload allow rekey cadence to be increased without sacrificing throughput, and this can be used to enable aggressive forward secrecy of sensitive corridors. The roadmap of hardware is essential. SmartNICs and appliances will open post-quantum instructions and hardened microcode and relocate bottlenecks to the wire speed engines. Operators will require per-suite performance profiles, firmware attestation of cryptographic units, and telemetry capability, and will stagger refresh, allowing cohorts to absorb capabilities as they become mature.

Post-quantum signatures will be integrated into the authentication chains of payments early. The size of issuer, acquirer, and gateway certificates, and

verification time, affect terminal latency and batch windows. Pilots measure the effects of client authentication on reconciliation service level agreements and determine contingency paths when stapling fails. Regulators are developing guidance on crypto agility controls, downgrade reporting, and audit-ready evidence in public key infrastructure and control planes. The continuous crypto agility seals the gap. Path rotation drills, automated certificate path linting, and rekey event and path MTU fault injections keep teams in match-fit condition. Leading organizations do not view quantum-safe VPNs as a one-off solution but as a long-term capacity in terms of engineering, operations, and governance, with re-examination of algorithms, parameters, and vendor posture every year.

7. Practical Guidelines Checklist

7.1 Readiness Checklist

Such a production backbone entails a strict inventory of hardware models, firmware baselines, driver versions, and a clear definition of the post-quantum capabilities. Teams need to have a capability matrix indicating which firewall, router, and VPN concentrator support which KEM and signature suite, the maximum tunnel count it supports, the presence of crypto offload, and known defects [27]. Hardware security modules are to be tested in terms of key size, interfaces supported, fault-tolerant capability, and operator backup and escrow processes. Performance metrics must be verified on both the control plane and data plane using packet capture, handshake logs, crypto engine counters, and path metrics.

Time discipline is required since handshake analytics and incident timelines rely on synchronized clocks, and each site must operate redundant time sources, authentication, and robust monitoring. The experience of significant telematics initiatives reveals that the increased visibility of assets and proper timing enhances reliability and decision making, which can be directly applied to backbone observability and control quality of the finance networks [20]. The risk registers must be kept up to date with component-level threats, vendor end-of-life dates, and change freeze dates. The completion of the checklist is capped by an earthy laboratory scenario that reflects the production connections and has labelled owners of documentation, rollback files, and access rights.

7.2 Recommended Configurations

Pecers must pin ciphers to remove downgrade paths, and IKEv2 should use hybrid proposals that combine a classical Diffie-Hellman group and one of the selected KEMs. IKE lifetimes must restrict exposure without churn, and most sites perform well at two to eight hours with rekey start delays. The rekey window needs to support parallel child association establishment and lubricated teardown. The MTU planning is required as bigger handshakes and certificates increase the potential fragmentation. Tunnels must impose a virtual MTU that is approximately 1,400 bytes and MSS clamping that is approximately 1,360 bytes for IPv4 and a bit less for IPv6.

The appliances must permit a PMTU discovery and black hole detection, and log a failure due to the do-not-fragment flag. KEM-based key exchanges should be favored in use by TLS overlays and early data disabled. Record sizes of TLS will need to be tailored to the path, and in several backbones, caps of about four kilobytes can be helpful [1]. CSP stapling needs to be turned on, and CRL caches need to be checked. Backoff policies, as well as retry policies, should prevent synchronized storms following certificate or responder outages.

7.3 Rollout Playbook

The program commences with a lab pilot in which a branch to a data center and a hub-to-hub pattern is replicated with realistic delay and loss. Preanticipated flight validation should cover handshake time at 50 milliseconds round-trip, steady-state goodput at line rate with mixed flows, and anticipated failover due to link flaps and divergence of the routes. Acceptance objectives must limit the time to complete the handshake to a few hundred milliseconds at the fifty-millisecond mark. They should ensure that path MTU-related drops are below one in a hundred packets. Canary enabled should be deployed on low-risk sites and work their way through desirable traffic patterns to full deployment.

Each wave must have soak periods such that exit criteria are in place to verify the error budgets, rekey success rates, and verify device headroom. The

change tickets should include backout steps, a tested configuration, and a communications plan to everyone and the on-call teams [36]. The automated tests of configurations must be similar to the automation that checks the running and validates: rollback initiation is triggered when the alarms are violated. Post-change reviews are expected to compare tunnel health and application SLOs with baselines, indicating what was found and used to refine subsequent waves.

7.4 Monitoring and IR Runbooks

Golden signals are expected to monitor time to first valid packet on new tunnels, handshake failure percentage, rekey storming percentage, path MTU hit percentage, packet loss above guidelines, and CPU / accelerator saturation. The dashboard should display per-site distributions and short and long windows to provide a view of the tail risks [8]. Limit and rate of change tests should be merged in the setting of alert thresholds. Detection of downgrade attempts must compare offered suites and negotiated results and must flag unusual removal of KEM and signature options.

Operators should keep a notebook that associates handshake errors with routing events and certificate issues. Incident runbooks should comprise verification of time, certificate, responder health, route reachability, and fragmentation. Upgrade actions must cycle toward fallback cipher policies when it is authorized and recorded. The timeline and any contributory factors should be documented and included in post-action reviews that inform subsequent training and the formulation of any new or updated detection rules.

8. Conclusion

This paper indicates how, with proper cryptographic architecture coupled with a rigorous approach to operations, quantum-safe virtual private networks can become a reality for the financial backbone. The harvest has now decrypted, and later, the threat expands confidentiality horizons in the payments market account records and compliance telemetry, and accordingly warrants action before adversaries develop scalable quantum capability. The testing employed metro regional and long-haul routes, comprising actual loss and jitter, to stress both the data plane and the control plane. This situation was mitigated by hybrid key exchange in IPsec and TLS that addressed the assurance gap and retained the availability of critical flows. The high-level outcome is that confidentiality can be reinforced without compromising service level goals when handshake overhead fragmentation risk and rekey behavior are design features.

The cost envelope is experimentally measured and ensures practicability. Hybrid TLS handshakes were started up proportionately to classical references and IKEv2 message flights by multiple kilobytes. They necessitated the activation of NAT traversal and IKE fragmentation to avoid packet fragmentation across access routes. Paths in the fifty millisecond bucket would begin to see paths' cold start latencies in the hundreds of milliseconds, and path tails became wide with Hello Retry and packet loss occurring. In comparison, the data plane provided goodput on the order of a few percent of the classical operation in metro, regional, and long-haul settings, and jumbo frames led to a further reduction in CPU/gigabit on software gateways. Control plane load increased in instances of rekey and the event of failover, which is an argument to have more headroom in software nodes. In contrast, the ASIC-assisted appliance absorbed the overhead without losing capacity.

Adoption in the field leads to conclusions and provides definite prescriptions. Path MTU black hole behavior on mixed MPLS and Internet segments was the dominant cause of both setup failure and tail latency. A maximum segment size clamp of one thousand and thirty-six bytes, GRE over IPsec, a uniform strategy to the do not fragment bit, deliberate IKE fragmentation, path MTU probes, increased exposure to fragmentation, silent drops, and decreased retransmissions. In early rekey, jitter and low-level batching herd effects were eradicated, and the risk of rekey storms was eliminated. Golden signals that had been relatively accurate indicators of trouble included time to first valid packet handshake failure rate, rekey storm incidence, path MTU hit rate, and crypto engine utilization. Dashboards should display per-site distributions as well as short and long windows to provide incident commanders with a view of the tail behavior to act upon.

The deployment blueprint can be used in regulated institutions and is focused on a negotiated scope, and is auditable. Cohort-based rollouts that are progressively deployed in labs and other small branches and scale out to hubs and then trading venues impose blast radius constraints, with canary gates implementing thresholds on the handshake, delay blackholing rate, and rekey success. Crypto agility is achieved automatically with hybrid or cross-chained hierarchies, stapled revocation at the edge, strict freshness alarms, and hardware security modules for key custody. Change windows are timed on business cutoffs and batches, and every wave contains rehearsed rollback scripts, tested environments, and acceptance criteria. Schemas on telemetry must be consistent between vendors to allow the size of negotiated suites, number of Retry hellos, count of cookie issuance, number of rekey attempts, and fragmentation events to be reasoned about consistently by operations and by auditors. There will remain residual risk as middleboxes and cloud edges cannot be modelled by the laboratories. There should be continuous validation and adversarial drills. Convergence compact KEMs and accelerator support overhead will also be reduced. The institutions must review annually and work towards PQ-only corridors.

References

[1] Bock, L. (2022). *Learn Wireshark: A definitive guide to expertly analyzing protocols and troubleshooting networks using Wireshark*. Packt Publishing Ltd.

[2] Cao, Y., Zhao, Y., Wang, Q., Zhang, J., Ng, S. X., & Hanzo, L. (2022). The evolution of quantum key distribution networks: On the road to the qinternet. *IEEE Communications Surveys & Tutorials*, *24*(2), 839-894.

[3] Chavan, A. (2021). Eventual consistency vs. strong consistency: Making the right choice in microservices. International Journal of Software and Applications, 14(3), 45-56. https://ijsra.net/content/eventual-consistency-vs-strong-consistency-making-right-choice-microservices

[4] Chavan, A. (2023). Managing scalability and cost in microservices architecture: Balancing infinite scalability with financial constraints. Journal of Artificial Intelligence & Cloud Computing, 2, E264. http://doi.org/10.47363/JAICC/2023(2)E264

[5] Ebrahim, Y. K. (2020). Security analysis of website certificate validation.

[6] Friedberger, S. (2019). Security of Cryptographic Implementations.

[7] Gai, S. (2020). *Building a future-proof cloud infrastructure: A unified architecture for network, security, and storage services*. Addison-Wesley Professional.

[8] Gavazzi, A., Williams, R., Kirda, E., Lu, L., King, A., Davis, A., & Leek, T. (2023). A study of {Multi-Factor} and {Risk-Based} authentication availability. In *32nd USENIX Security Symposium (USENIX Security 23)* (pp. 2043-2060).

[9] Hadan, H., Serrano, N., & Camp, L. J. (2021). A holistic analysis of web-based public key infrastructure failures: comparing experts' perceptions and real-world incidents. *Journal of Cybersecurity*, *7*(1), tyab025.

[10] Joshi, R. (2022). *In-Network Techniques for Highly Reliable Datacenter Networks* (Doctoral dissertation, National University of Singapore (Singapore)).

[11] Joshua, T. (2023). A Secure Model for Student Results Verification Using Salted Hash Functions.

[12] Karwa, K. (2023). AI-powered career coaching: Evaluating feedback tools for design students. Indian Journal of Economics & Business. https://www.ashwinanokha.com/ijeb-v22-4-2023.php

[13] Khazraee, M. (2020). *Reducing the development cost of customized hardware acceleration for cloud infrastructure*. University of California, San Diego.

[14] Kjell, E., & Frisenfelt, S. (2021). Characterization of cipher suite selection, downgrading, and other weaknesses observed in the wild.

[15] Konneru, N. M. K. (2021). Integrating security into CI/CD pipelines: A DevSecOps approach with SAST, DAST, and SCA tools. *International Journal of Science and Research Archive*. Retrieved from

https://ijsra.net/content/role-notification-scheduling-improving-patient

[16] Kumar, A. (2019). The convergence of predictive analytics in driving business intelligence and enhancing DevOps efficiency. International Journal of Computational Engineering and Management, 6(6), 118-142. Retrieved from https://ijcem.in/wp-content/uploads/THE-CONVERGENCE-OF-PREDICTIVE-ANALYTICS-IN-DRIVING-BUSINESS-INTELLIGENCE-AND-ENHANCING-DEVOPS-EFFICIENCY.pdf

[17] Kumar, V. (2023). Digital hotspots. In *The Economic Value of Digital Disruption: A Holistic Assessment for CXOs* (pp. 689-795). Singapore: Springer Nature Singapore.

[18] Liechti, L., Gouveia, P., Neves, J., Kropf, P., Matos, M., & Schiavoni, V. (2019, October). THUNDERSTORM: a tool to evaluate dynamic network topologies on distributed systems. In *2019 38th Symposium on Reliable Distributed Systems (SRDS)* (pp. 241-24109). IEEE.

[19] Nawrocki, M., Blendin, J., Dietzel, C., Schmidt, T. C., & Wählisch, M. (2019, October). Down the black hole: dismantling operational practices of BGP blackholing at IXPs. In *Proceedings of the Internet Measurement conference* (pp. 435-448).

[20] Nyati, S. (2018). Revolutionizing LTL carrier operations: A comprehensive analysis of an algorithm-driven pickup and delivery dispatching solution. International Journal of Science and Research (IJSR), 7(2), 1659-1666. Retrieved from https://www.ijsr.net/getabstract.php?paperid=SR24203183637

[21] Nyati, S. (2018). Transforming telematics in fleet management: Innovations in asset tracking, efficiency, and communication. International Journal of Science and Research (IJSR), 7(10), 1804-1810. Retrieved from https://www.ijsr.net/getabstract.php?paperid=SR24203184230

[22] Pachilakis, M., Chariton, A. A., Papadopoulos, P., Ilia, P., Degkleri, E., & Markatos, E. P. (2020). Design and implementation of a compressed certificate status protocol. *ACM Transactions on Internet Technology (TOIT)*, *20*(4), 1-25.

[23] Palmer, M. R. (2022). Towards enabling cross-layer information sharing to improve today's content delivery systems.

[24] Pouttu, A. (2020). 6G white paper on validation and trials for verticals towards 2030's.

[25] Putters, J., Hashemi, J. B., & Yavuz, A. (2023). Demystifying public cloud auditing for IT auditors. *Advanced Digital Auditing*, *185*.

[26] Raju, R. K. (2017). Dynamic memory inference network for natural language inference. International Journal of Science and Research (IJSR), 6(2). https://www.ijsr.net/archive/v6i2/SR24926091431.pdf

[27] Rao, S. P. (2023). Analyzing Communications and Software Systems Security.

[28] Sabanci, K. (2023). *Exploring post-quantum cryptographic schemes for TLS in 5G NB-IOT: Feasibility and recommendations* (Master's thesis, Marquette University).

[29] Sardana, J. (2022). Scalable systems for healthcare communication: A design perspective. *International Journal of Science and Research Archive*. https://doi.org/10.30574/ijsra.2022.7.2.0253

[30] Sardana, J. (2022). The role of notification scheduling in improving patient outcomes. *International Journal of Science and Research Archive*. Retrieved from https://ijsra.net/content/role-notification-scheduling-improving-patient

[31] Schäge, S., Schwenk, J., & Lauer, S. (2020, April). Privacy-preserving authenticated key exchange and the case of IKEv2. In *IACR International Conference on Public-Key Cryptography* (pp. 567-596). Cham: Springer International Publishing.

[32] Schwabe, P., Stebila, D., & Wiggers, T. (2020, October). Post-quantum TLS without handshake signatures. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1461-1480).

[33] Shantharama, P., Thyagaturu, A. S., & Reisslein, M. (2020). Hardware-accelerated platforms and infrastructures for network functions: A survey of enabling technologies and research studies. *IEEE Access*, *8*, 132021-132085.

[34] Sharma, N. (2023). Legacy apps to cloud: A risk-based approach. *Cyber Security: A Peer-Reviewed Journal*, *7*(1), 16-23.

[35] Singh, V. (2021). Generative AI in medical diagnostics: Utilizing generative models to create synthetic medical data for training diagnostic algorithms. International Journal of Computer Engineering and Medical Technologies. https://ijcem.in/wp-content/uploads/GENERATIVE-AI-IN-MEDICAL-DIAGNOSTICS-UTILIZING-GENERATIVE-MODELS-TO-CREATE-SYNTHETIC-MEDICAL-DATA-FOR-TRAINING-DIAGNOSTIC-ALGORITHMS.pdf

[36] Wu, P. (2019). Analysis of the WireGuard protocol. *Master's Thesis, Analysis of the WireGuard protocol, Eindhoven University of Technology*.

[37] Wu, Y., Chai, B., Li, Y., Liu, B., Li, J., Yang, Y., & Jiang, W. (2023, May). An empirical study on change-induced incidents of online service systems. In *2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)* (pp. 234-245). IEEE.

[38] Yang, B., Xue, W., Zhang, T., Liu, S., Ma, X., Wang, X., & Liu, W. (2023). End-to-end I/O monitoring on leading supercomputers. *ACM Transactions on Storage*, *19*(1), 1-35.