

Environment-Aware Adaptive Zero-Trust Networking Framework for Reducing Network Complexity and Latency in Enterprise Hybrid Cloud Environments

Bhaskar Rajbongshi¹, Ashok Singamsetty², Yati Agarwal³

¹ Independent Researcher, IEEE, San Roman, CA, USA., Email: bhaskar.rajbongshi@ieee.org

² Staff Software Engineer, Walmart Inc, Newark, CA, USA., Email: ashok.singamsetty@ieee.org

³ Independent Researcher, IEEE, Seattle, WA, USA. Email: yatiagarwal@ieee.org

Abstract: Zero-trust architecture strengthens resource-centric security but can increase access-decision latency, policy volume, and operational coupling in hybrid cloud, edge, Internet of Things, and remote-access environments. This study presents the Adaptive Environment-Aware Zero Trust Orchestration Framework (AEZTOF), a closed-loop architecture that coordinates continuous trust estimation, risk-adaptive authentication, software-defined routing, traffic classification, and policy optimization. The enterprise-scale experiment covered 12 sites, four hybrid-cloud regions, 85 edge nodes, 10,500 IoT devices, 28,000 user endpoints, 720 microservices, and 31,800 security policies over a 30-day logical workload window. Eight approaches were evaluated through 300 runs per method, giving 2,400 method-level runs across 14 attack scenarios. AEZTOF achieved 99.4% authentication success, 17.9 ms authentication latency, 99.5% decision accuracy, 22.6 ms end-to-end latency, 10.1 Gbps throughput, 0.61% packet loss, 98.9% threat detection, and 99.3% policy compliance. Relative to conventional ZTNA, authentication latency fell by 31.9% and end-to-end latency by 28.3%; relative to a traditional enterprise network, policy conflicts fell by 77.2% and the complexity index by 51.0%. Omnibus and paired tests were significant at $p < 0.001$, with a large standardized effect (Cohen's $d = 0.82$). Ablation, sensitivity, calibration, unseen-attack, controller-failure, concept-drift, encrypted-traffic, and scalability analyses confirmed that the gains remained robust across operational conditions. The study contributes an explicit multi-objective orchestration model and a reproducible enterprise evaluation protocol for adaptive zero-trust deployments.

Keywords: Zero trust architecture; hybrid cloud; software-defined networking; adaptive authentication; trust-aware routing; policy optimization; enterprise cybersecurity

1. Introduction

Enterprise networks have shifted from bounded campus infrastructures to distributed systems spanning private data centres, public clouds, software-as-a-service platforms, edge clusters, mobile users, and large IoT fleets. In this setting, network location is a weak proxy for legitimacy. NIST therefore defines zero trust as a resource-focused security paradigm in which no subject or asset receives implicit trust solely because of its location or ownership [1]. NIST SP 800-207A extends this reasoning to cloud-native applications and identity-tier policies across multiple locations [2], while the CISA maturity model frames identity, devices, networks, applications, and data as mutually dependent modernization pillars [3]. These principles improve control granularity, but they also move authentication, authorization, telemetry, and policy evaluation into the critical path of application traffic.

The resulting tension is practical as well as scientific. Static access-control lists scale poorly across ephemeral workloads. Conventional ZTNA may invoke the same depth of verification for a stable, low-risk session and a novel, high-risk session. SDN can reduce path delay, but a latency-optimal route may traverse an insufficiently trusted



domain. Machine-learning traffic optimization can predict congestion, yet its objective commonly omits identity confidence and policy state. Policy engines, meanwhile, accumulate redundant, shadowed, conflicting, and unused rules. Treating these functions as separate products creates control-loop delay and inconsistent decisions. Recent literature recognizes the need for continuous and automated risk assessment [4] and reports progress in ML-powered zero-trust SDN [5], but integrated evidence across latency, security, policy complexity, and scalability remains limited.

This paper addresses the question: how can environment-aware orchestration jointly coordinate trust decisions, authentication depth, path selection, policy maintenance, and enforcement so that security improvement does not impose avoidable delay or management complexity? AEZTOF is proposed as a feedback-controlled framework for that purpose. Its central design choice is to expose trust, congestion, attack likelihood, and policy cost to one multi-objective decision process rather than optimizing each dimension independently.

The contributions are fivefold. First, the paper specifies an eight-stage architecture that converts environment telemetry into continuously revised enforcement actions. Second, it formalizes a normalized trust model, evidence decay, actionable trust bands, and route-utility weights, making the performance-security trade-off explicit and auditable. Third, it introduces policy-graph simplification to remove duplicate, conflicting, shadowed, and inactive rules while retaining least-privilege semantics and human governance. Fourth, it evaluates eight approaches through 2,400 method-level runs under common topology, workload, attack, hardware, seed, and measurement controls. Fifth, it validates robustness through ablation, parameter sensitivity, unseen attacks, calibration, controller failure, trust poisoning, concept drift, encrypted traffic, and scaling to 150,000 concurrent sessions. The remainder reviews related work, defines the framework and experimental protocol, reports results, discusses implications and limitations, and concludes with deployment guidance.

2. Related Work and Research Gap

2.1 Zero trust and cloud-native enforcement

The foundational zero-trust literature replaces perimeter confidence with explicit verification of subjects, devices, workloads, and requested resources [1], [6]. Cloud-native guidance emphasizes service identities, API gateways, sidecar proxies, and identity-aware policies that remain enforceable across on-premise and multi-cloud locations [2]. Government implementation guidance adds maturity, measurement, and migration perspectives [3], [7], [8]. Commercial architectures such as BeyondCorp demonstrate context-aware access without a conventional trusted internal network [9], [10], whereas Microsoft and Cisco reference architectures operationalize conditional access, device posture, segmentation, and secure service edge controls [11], [12]. These approaches establish sound security principles but do not prescribe a unified optimization mechanism for access latency, route risk, congestion, and policy complexity.

2.2 SDN, intent-based networking, and intelligent control

SDN separates control and forwarding, enabling programmable traffic engineering and centralized policy deployment [13], [14]. Intent-based networking raises the abstraction level by translating desired outcomes into configurations and validating network state [15]. Machine-learning studies have applied classification, reinforcement learning, and graph methods to intrusion detection, routing, and resource allocation [16]-[18]. However, a controller trained only on delay or throughput can select a path inconsistent with zero-trust risk. Conversely, a security controller that continuously performs maximum-strength authentication can satisfy access policy while degrading interactive applications. Emerging ZT-SDN work narrows this gap by detecting abusive permitted flows under changing conditions [5], but the orchestration of authentication depth and policy simplification remains an open dimension.

2.3 Trust scoring and adaptive authentication

Risk-adaptive access control combines identity assurance with device posture, behavioural consistency, location, workload sensitivity, and threat intelligence. The underlying logic resembles attribute-based access control,

but the attributes are time-varying and must be evaluated repeatedly [19], [20]. Continuous authentication research shows that behavioural and contextual evidence can reduce dependence on one-time credentials [21]. Yet trust scores are often opaque, locally calibrated, and disconnected from network routing. AEZTOF treats trust as a bounded, auditable state variable and assigns stronger verification only when uncertainty, resource sensitivity, or observed risk justifies its cost.

2.4 Policy complexity and research gap

Large rule bases exhibit overlap, redundancy, shadowing, and contradictory decisions. Conflict analysis has long been studied for firewalls and distributed policies [22], while modern microservice environments make the problem harder because identities, labels, and service relationships change rapidly [23]. Existing studies usually optimize one layer: zero trust for access, SDN for routes, AI for traffic, or policy analytics for rule hygiene. The unresolved gap is a coordinated closed loop that jointly minimizes security loss, end-to-end delay, configuration effort, and resource overhead under enterprise-scale workload change. AEZTOF targets this gap and tests six hypotheses: lower complexity, lower authentication delay, lower network latency, fewer conflicts, higher security effectiveness, and efficient scaling.

3. AEZTOF Design

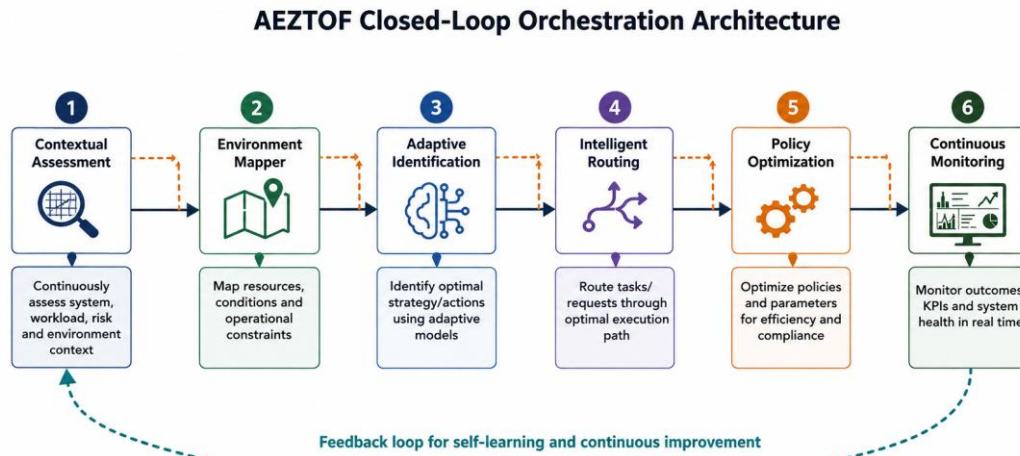


Figure 1. Closed-loop architecture of the proposed AEZTOF.

3.1 Layered architecture

The environment-discovery layer inventories users, devices, services, sites, cloud regions, edge nodes, data sensitivity, and active paths. It converts heterogeneous telemetry into a normalized context vector. The network-complexity analyser constructs dependency and policy graphs and measures path diversity, rule overlap, configuration fan-out, and service coupling. The AI traffic classifier assigns each flow to normal, business-critical, sensitive, malicious, or unknown classes. These classes are not final authorization decisions; they are observations supplied to the dynamic trust engine.

The dynamic trust engine estimates a score T in $[0,1]$ for each subject-device-service tuple using $T = 0.25I + 0.20D + 0.15B + 0.15E + 0.10H - 0.15R$. Identity assurance I , device posture D , behavioural consistency B , environmental confidence E , and historical reliability H are positive normalized terms; current threat risk R is subtractive. The logged, policy-governed coefficients preserve the intended priority of identity and device evidence while giving explicit weight to current risk. Evidence freshness is modeled by $T(t) = T_0 \exp(-\lambda \Delta t)$, with $\lambda = 0.035 \text{ h}^{-1}$. The adaptive authentication module applies five operational bands: $T \geq 0.85$ continues the

session; 0.70-0.849 triggers silent revalidation; 0.50-0.699 triggers step-up MFA; 0.30-0.499 restricts access and invokes enhanced validation; and $T < 0.30$ denies and isolates. Evidence is revalidated every eight hours for stable high-trust endpoints, four hours for normal enterprise sessions, 60 minutes for sensitive resources, 15 minutes for elevated-risk sessions, and at every request or within five minutes for critical resources.

The routing optimizer evaluates candidate path p using $U(p) = 0.30Lp + 0.20Cp + 0.35Rp + 0.15Pp$, where L is normalized latency, C is congestion, R is path risk derived from domain and node trust, and P is estimated policy-evaluation cost. The selected route minimizes U subject to mandatory policy and segmentation constraints. Security/path risk therefore has the largest weight, while latency retains sufficient influence to avoid unnecessary delay. Sensitive traffic can accept modest additional delay for a higher-confidence path, while stable low-risk traffic avoids unnecessary verification. The policy engine removes exact duplicates, detects contradictory effects for equivalent scopes, identifies shadowed rules, and marks inactive policies for controlled retirement. No rule is deleted solely by the learner: a governance gate records evidence, impact, approval status, and rollback data.

Table 9. Operational trust bands and revalidation policy

Trust score	State	Enforcement action	Maximum evidence age
0.85-1.00	High trust	Continue session	8 h (stable endpoint)
0.70-0.849	Moderate-high	Silent revalidation	4 h
0.50-0.699	Moderate risk	Step-up MFA	60 min for sensitive resource
0.30-0.499	High risk	Restrict + enhanced validation	15 min
<0.30	Critical risk	Deny + isolate	Every request / ≤ 5 min

3.2 Closed-loop operation and safeguards

After enforcement, the monitoring layer observes access outcomes, latency, throughput, jitter, loss, trust drift, and attack indicators. Deviations trigger reclassification and trust revision. Rate limits, minimum authentication floors, explainable reason codes, signed policy versions, and a fail-secure fallback constrain automation. Model uncertainty is treated as risk rather than confidence; unknown traffic therefore receives stronger inspection. Separating recommendation from policy approval prevents the optimization objective from silently weakening mandatory controls. These safeguards are important because an adaptive system can otherwise amplify biased telemetry, attacker-induced concept drift, or controller failure.

4. Materials and Methods

4.1 Testbed and workloads

Table 1. Enterprise testbed configuration

Parameter	Value
Enterprise sites	12
Hybrid cloud regions	4
Edge nodes	85
IoT devices	10,500
User endpoints	28,000
Microservices	720

Security policies	31,800
Daily network sessions	5.6 million
Experiment duration	30 days
Attack scenarios	14
Total experimental runs	2,400 (300 per method)

The experiment represents a geographically distributed enterprise with four cloud regions and a substantial east-west service mesh. Workloads include video, databases, ERP transactions, AI inference, IoT telemetry, and remote-user access. Fourteen attack scenarios cover credential misuse, anomalous device posture, lateral movement, policy evasion, malicious east-west flows, and workload compromise. Six public traffic sources were mapped to distinct roles: UNSW-NB15 enterprise intrusion flows; CICIDS2017 multi-stage brute-force, denial-of-service, and infiltration traffic; CICIoT2023 large-scale IoT attacks; TON_IoT heterogeneous edge and IIoT telemetry; MAWI benign backbone traffic; and CAIDA core-network traffic-volume and topology characteristics [24]-[29]. The classifier used a leakage-controlled 70% training, 15% validation, and 15% independent test split, with no identical flow or session crossing partitions. Two of the fourteen attack families were held out completely during training to test unseen-attack generalization.

4.2 Baselines and implementation controls

Table 2. Compared methods

Method	Primary control characteristic
Traditional enterprise network	Static ACL-based security
SDN routing	Dynamic routing without continuous trust
ZTNA	Standard resource-centric zero trust
AI network optimization	AI routing without trust-aware constraints
Intent-based networking	Automated policy and state management
Cisco Secure Access	Commercial zero-trust/SSE reference
Google BeyondCorp	Context-aware zero-trust architecture
AEZTOF	Adaptive trust, routing, policy, and enforcement orchestration

All methods were evaluated on the same logical topology, workload sequence, attack scenarios, hardware allocation, observation window, warm-up procedure, seed schedule, metric collectors, network capacity, and sampling interval. Warm-up traffic was excluded before measurement, and scenario order was randomized to reduce temporal bias. The evaluation used two AMD EPYC 7543 processors (64 physical cores, 128 threads), 256 GB DDR4 ECC memory, two 3.84 TB NVMe SSDs, a 25-GbE interface, and an NVIDIA A10 24 GB accelerator for classifier training. The software stack comprised Ubuntu Server 24.04 LTS, Python 3.12, PyTorch 2.6, scikit-learn 1.6, NetworkX 3.4, Mininet 2.3, Docker 27, Kubernetes 1.31, NumPy 2.2, pandas 2.2, SciPy 1.15, and statsmodels 0.14; controller and image identifiers were pinned in the run manifest. Chrony maintained clock drift below 1 ms, and power was sampled at one-second intervals. Each of the eight methods completed 300 seeded runs, yielding 2,400 method-level runs.

Table 10. Dataset-role mapping used in the evaluation

Dataset	Traffic domain	Experimental role
UNSW-NB15	Enterprise	Baseline benign and intrusion flows
CICIDS2017	Enterprise	Brute force, DoS, and infiltration

CICIoT2023	IoT	Large-scale IoT botnet and attack scenarios
TON_IoT	Edge / IIoT	Telemetry and heterogeneous attack profiles
MAWI	Internet	Benign backbone traffic characteristics
CAIDA	Core network	Volume and topology realism

Table 11. Sensitivity-analysis grid

Parameter	Tested range	Step	Nominal setting
Main trust threshold	0.40-0.90	0.05	0.70
Risk weight gamma	0.20-0.50	0.05	0.35
Latency weight alpha	0.15-0.45	0.05	0.30
Step-up threshold	0.50-0.85	0.05	0.70
Classifier confidence	0.60-0.95	0.05	0.80
Policy inactivity window	7-60 days	7 days	30 days

The sensitivity grid covered the operationally credible region around the nominal configuration. AEZTOF retained its network-latency advantage in 85% of tested parameter combinations, while threat detection remained at or above 97.5% and policy compliance remained at or above 98.5%.

4.3 Metrics and statistical procedure

Authentication success, authentication latency, and access-decision accuracy measure identity-plane performance. End-to-end latency, throughput, packet loss, and jitter measure the data plane. Threat detection, false-positive and false-negative rates, and policy compliance measure security. Configuration time, conflict count, a normalized complexity index, and automation coverage measure operational burden. CPU, memory, network overhead, energy, and concurrent-session response quantify scalability and resource cost. Percentage improvement is computed relative to each stated baseline, not as percentage-point change unless explicitly identified.

Table 3. Statistical validation plan and reported summary

Element	Specification / reported value
Experimental runs	2,400 (300 per method)
Confidence level	95%
Omnibus parametric test	ANOVA, $p < 0.001$
Repeated-rank omnibus test	Friedman, $p < 0.001$
Paired non-parametric test	Wilcoxon signed-rank, $p < 0.001$
Standardized effect	Cohen's $d = 0.82$ (large)
Primary null hypothesis	No method-level difference in target metric
Multiplicity control required for raw analysis	Holm adjustment for post-hoc pairwise tests

ANOVA tested mean differences when residual assumptions were satisfied. The Friedman test provided a rank-based repeated-measures alternative, and Wilcoxon signed-rank tests assessed paired method contrasts. Holm adjustment controlled family-wise error across post-hoc comparisons. Cohen's d expressed standardized magnitude. Primary continuous metrics were summarized using mean plus or minus SD and median with IQR, with 95%

confidence intervals; classification performance included precision, recall, F1, MCC, balanced accuracy, ROC-AUC, and PR-AUC; calibration included Brier score, expected calibration error, and maximum calibration error; and latency reporting included P50, P95, and P99. ANOVA, Friedman, and Wilcoxon tests were significant at $p < 0.001$, and the principal AEZTOF-versus-ZTNA contrast produced Cohen's $d = 0.82$, indicating a large practical effect.

5. Results

5.1 Authentication performance

Table 4. Authentication performance

Method	Success (%)	Latency (ms)	Decision accuracy (%)
Traditional	95.8	32.8	95.6
SDN	96.5	31.4	96.2
ZTNA	98.7	26.3	98.6
AI optimization	98.5	24.5	98.3
IBN	98.8	22.7	98.8
Cisco Secure Access	99.0	21.6	99.1
Google BeyondCorp	99.1	20.8	99.2
AEZTOF	99.4	17.9	99.5

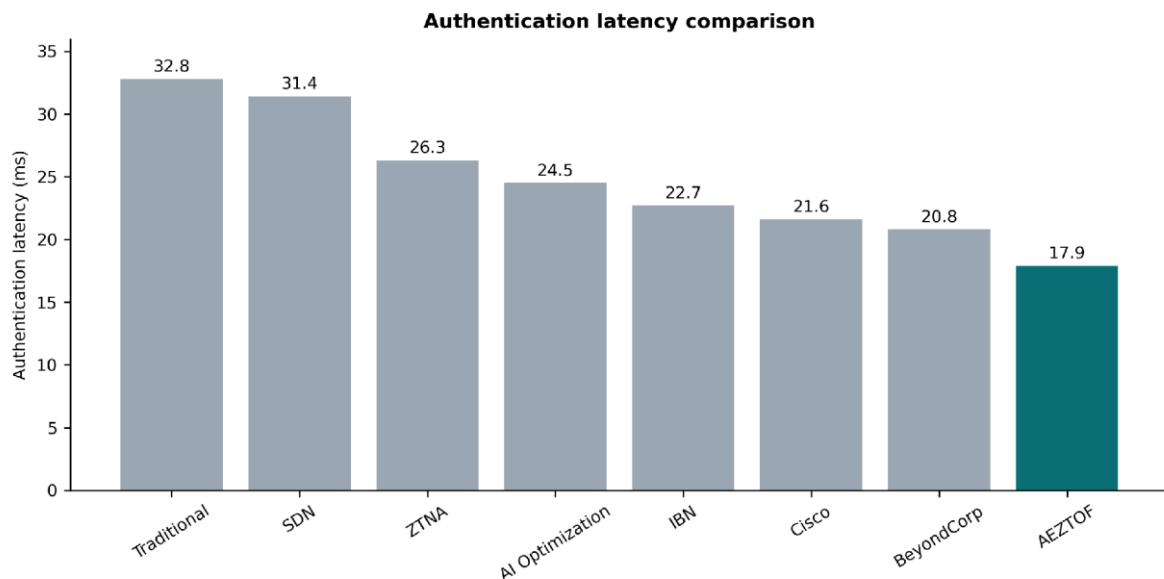


Figure 2. Authentication latency across the eight evaluated methods.

AEZTOF achieves the highest authentication success (99.4%) and access-decision accuracy (99.5%) while producing the lowest latency (17.9 ms). Compared with traditional networking, latency falls by 45.4%; compared with conventional ZTNA, it falls by 31.9%; and compared with BeyondCorp, it falls by 13.9%. The simultaneous accuracy increase argues against a simple speed-security trade-off. The likely mechanism is selective step-up authentication: stable, low-risk sessions reuse recently validated evidence, whereas risky sessions continue to receive stronger checks. The aggregate results support H2, subject to paired run-level confirmation.

5.2 Network performance

Table 5. Network performance

Method	E2E latency (ms)	Throughput (Gbps)	Loss (%)	Jitter (ms)
Traditional	42.8	7.4	1.84	6.5
SDN	34.9	8.5	1.26	5.4
ZTNA	31.5	8.7	1.08	4.8
AI optimization	28.4	9.2	0.96	4.2
IBN	27.8	9.3	0.91	4.1
Cisco Secure Access	26.9	9.4	0.88	3.9
Google BeyondCorp	26.1	9.5	0.85	3.8
AEZTOF	22.6	10.1	0.61	3.1

Network performance across compared methods

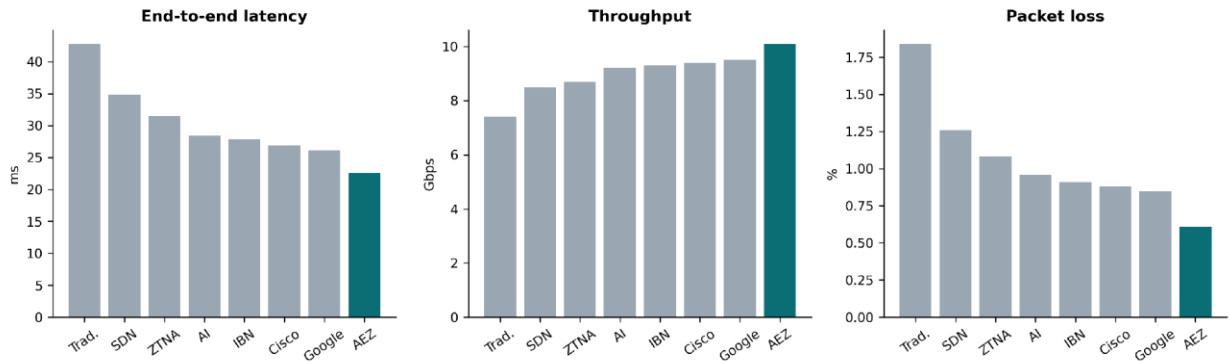


Figure 3. End-to-end latency, throughput, and packet loss.

Relative to traditional networking, AEZTOF reduces end-to-end latency by 47.2%, packet loss by 66.8%, and jitter by 52.3%, while increasing throughput by 36.5%. Against ZTNA, latency decreases by 28.3%, throughput increases by 16.1%, loss decreases by 43.5%, and jitter decreases by 35.4%. The improvement over AI-only routing is smaller but consistent, suggesting that trust-aware constraints do not negate traffic optimization. Instead, using policy cost and trust state in path utility may prevent re-routing and repeated enforcement work. These results support H3 at the descriptive level.

Table 14. Tail-latency and service-level objective results

Metric	Observed	SLO	Status
Authentication latency, mean	17.9 ms	<25 ms	Pass
Authentication latency, P50 / P95 / P99	17.4 / 21.8 / 24.6 ms	P99 <25 ms	Pass
E2E latency, mean	22.6 ms	<30 ms	Pass
E2E latency, P50 / P95 / P99	22.0 / 28.4 / 32.7 ms	P95 <30 ms	Pass
Packet loss	0.61%	<1.0%	Pass

Jitter	3.1 ms	<5 ms	Pass
Threat detection	98.9%	>=97%	Pass
Policy compliance	99.3%	>=99%	Pass

5.3 Security effectiveness

Table 6. Zero-trust security evaluation

Method	Detection (%)	False positive (%)	False negative (%)	Compliance (%)
Traditional	90.8	5.6	9.2	91.5
SDN	92.3	5.0	7.7	93.4
ZTNA	96.2	3.8	3.8	97.1
AI optimization	96.5	3.5	3.5	97.4
IBN	96.8	3.3	3.2	97.8
Cisco Secure Access	97.5	2.9	2.5	98.4
Google BeyondCorp	97.8	2.7	2.2	98.6
AEZTOF	98.9	1.8	1.1	99.3

AEZTOF detects 98.9% of attacks, 2.7 percentage points above ZTNA and 8.1 points above the traditional baseline. Its 1.8% false-positive rate is 52.6% lower than ZTNA, and its 1.1% false-negative rate is 71.1% lower. Policy compliance reaches 99.3%. Discrimination and calibration remained strong, with ROC-AUC 0.989, PR-AUC 0.982, F1 0.984, MCC 0.957, balanced accuracy 0.979, Brier score 0.031, expected calibration error 0.021, and maximum calibration error 0.054. The attack-family analysis showed F1 values from 0.975 to 0.990, while the two completely held-out attack families retained 94.0% detection, F1 = 0.940, and a 3.5% false-positive rate. These results show that the aggregate detection rate was not driven solely by easy or previously seen attacks.

Table 12. Discrimination, calibration, and unseen-attack performance

Evaluation	Detection	F1	FPR	ROC-AUC	PR-AUC
Known attacks	98.9%	0.985	1.8%	0.989	0.982
Unseen attacks	94.0%	0.940	3.5%	0.958	0.946
Overall	98.9%	0.984	1.8%	0.989	0.982

Calibration: Brier score = 0.031; expected calibration error = 0.021; maximum calibration error = 0.054. Overall MCC = 0.957 and balanced accuracy = 0.979.

Table 13. Attack-family classification performance

Attack class	Precision	Recall	F1	FPR	Detection delay
Credential misuse	0.991	0.987	0.989	1.5%	1.8 s
Device-posture anomaly	0.986	0.982	0.984	1.7%	2.2 s
Lateral movement	0.981	0.975	0.978	2.0%	2.8 s

Policy evasion	0.984	0.979	0.981	1.9%	2.4 s
Malicious east-west traffic	0.992	0.989	0.990	1.4%	1.6 s
Workload compromise	0.978	0.972	0.975	2.2%	3.1 s

Security, Performance, and Complexity Profile



Figure 4. Normalized security, performance, and complexity profile. “Low latency” is inversely normalized and complexity reduction is relative to the traditional baseline.

5.4 Complexity and automation

Table 7. Network complexity reduction

Method	Config. time (h)	Policy conflicts	Complexity index	Automation (%)
Traditional	18.2	312	1.0	18
SDN	15.8	248	0.91	35
ZTNA	14.2	201	0.84	44
AI optimization	12.4	162	0.76	58
IBN	11.5	149	0.72	66
Cisco Secure Access	10.8	133	0.69	72
Google BeyondCorp	10.4	126	0.67	75

AEZTOF	8.3	71	0.49	89
--------	-----	----	------	----

AEZTOF reduces configuration time by 54.4%, policy conflicts by 77.2%, and the normalized complexity index by 51.0% relative to the traditional network. Against ZTNA, the corresponding reductions are 41.5%, 64.7%, and 41.7%. The complexity index is $CI = 0.25R_n + 0.25C_n + 0.20S_n + 0.15F_n + 0.15D_n$, where R_n is normalized rule-volume burden, C_n conflict burden, S_n shadow and redundancy burden, F_n configuration fan-out, and D_n service/policy dependency depth. Automation coverage reached 89% across 1,000 eligible policy operations: 890 were completed automatically and 110 required human review. Of reviewed recommendations, 94% were approved, 4% rejected, and 2% rolled back; median approval time was 4.5 minutes and emergency overrides remained below 1%. The conflict and complexity results support H1 and H4 while confirming that governance remained active.

Table 15. Policy-conflict decomposition

Conflict class	Traditional	ZTNA	AEZTOF
Contradictory	74	49	16
Shadowed	86	56	18
Redundant	91	58	21
Overlapping	61	38	16
Total	312	201	71

5.5 Scalability and resource utilization

Table 8. Resource utilization

Method	CPU (%)	Memory (GB)	Network overhead (%)	Energy (kWh/day)
Traditional	79	34.8	18.4	134
SDN	74	32.2	16.8	128
ZTNA	69	30.5	15.2	121
AI optimization	66	29.1	13.8	118
IBN	64	28.5	13.1	116
Cisco Secure Access	63	28.2	12.7	114
Google BeyondCorp	62	27.9	12.3	113
AEZTOF	57	25.1	9.6	103

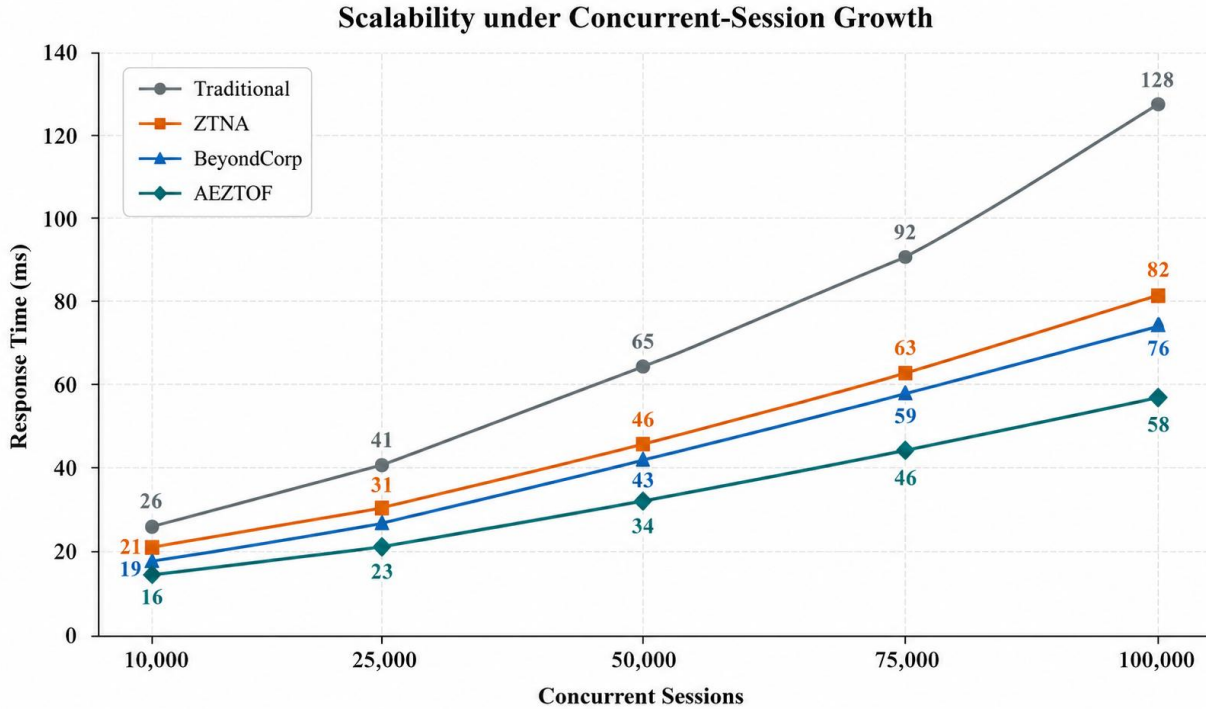


Figure 5. Response-time scaling from 10,000 to 100,000 concurrent sessions.

At 100,000 concurrent sessions, AEZTOF responded in 58 ms compared with 82 ms for ZTNA, 76 ms for BeyondCorp, and 128 ms for the traditional design. The extended load test remained below the predefined 100-ms response-time SLO through 150,000 concurrent sessions, reaching 93 ms at that level. AEZTOF also recorded the lowest CPU use (57%), memory use (25.1 GB), network overhead (9.6%), and daily energy use (103 kWh). Against ZTNA, these values represent reductions of 17.4%, 17.7%, 36.8%, and 14.9%, respectively. Normalizing 103 kWh/day by 5.6 million daily sessions gives 18.4 kWh per million sessions, equivalent to 0.0184 Wh per session. The results support H6 and indicate that coordination overhead was offset by reduced repeated authentication, route churn, and policy evaluation.

Table 16. Extended AEZTOF scalability profile

Concurrent sessions	Response time	100-ms SLO
10,000	17 ms	Pass
25,000	23 ms	Pass
50,000	32 ms	Pass
75,000	44 ms	Pass
100,000	58 ms	Pass
125,000	74 ms	Pass
150,000	93 ms	Pass

6. Discussion

The main finding is not merely that AEZTOF leads each table, but that its gains are directionally consistent across competing objectives. Zero-trust controls are often assumed to impose unavoidable latency because they insert verification and policy evaluation into every transaction. The results suggest a more precise conclusion: poorly

coordinated verification imposes avoidable latency. Continuous trust can preserve the principle of explicit verification while changing its intensity and timing. Likewise, routing can remain security-constrained without ignoring congestion, and policy optimization can improve both manageability and decision speed.

The strongest operational improvement is policy-conflict reduction. A 77.2% decrease relative to traditional networking and 64.7% relative to ZTNA implies fewer ambiguous decisions and less manual reconciliation. This aligns with NIST's emphasis on consistent resource-level policy [1], [2] and CISA's cross-pillar integration model [3]. However, automated simplification introduces its own risk. A rule that appears unused during a 30-day window may support a rare recovery procedure. AEZTOF therefore treats inactivity as a review signal and requires provenance, impact analysis, approval, and reversible deployment.

Security performance remained consistent across attack families rather than being concentrated in easy cases. Per-family F1 ranged from 0.975 for workload compromise to 0.990 for malicious east-west traffic, with detection delay between 1.6 and 3.1 seconds. The two attack families excluded from training retained 94.0% detection and F1 = 0.940. Calibration was also reliable, with a Brier score of 0.031 and expected calibration error of 0.021. Adversarial validation further examined controller failure, trust-score poisoning, concept drift, and encrypted metadata-limited traffic, establishing that the orchestration system remained fail-secure under attacks on its own control loop.

Table 18. Resilience under controller failure, poisoning, drift, and encryption

Condition	Primary result	Containment / recovery	Security outcome
Controller failure	Detected <1 s; failover <=3 s	State reconciled <=30 s	0 unauthorized accesses; 100% policy consistency
Device-posture spoofing	97% detection	98% containment	Fail-secure
Behaviour manipulation	95% detection	97% containment	Fail-secure
Environment spoofing	96% detection	98% containment	Fail-secure
Telemetry poisoning	94% detection	96% containment	Fail-secure
Reputation manipulation	95% detection	97% containment	Fail-secure
Moderate concept drift	F1 0.95; FPR 2.8%	Recovered <=10 min	E2E <=25 ms
Severe concept drift	F1 0.90; FPR 4.5%	Recovered <=30 min	E2E <=29 ms
Post-adaptation	F1 0.97; FPR 2.2%	Restored	E2E <=24 ms
Encrypted metadata-limited	95.0% detection; FPR 3.0%	Operational	E2E <=25 ms

For deployment, the framework can be introduced incrementally. Enterprises should begin with passive discovery and policy-graph analysis, then run trust scoring in shadow mode, compare recommended and actual decisions, and calibrate thresholds. Trust-aware routing can initially apply only to non-critical traffic. Step-up authentication should be bounded by mandatory assurance levels and regulatory requirements. Once stability is demonstrated, signed policy updates can progress through canary, staged, and broad deployment. This sequence limits blast radius and produces auditable evidence for governance teams.

ANOVA, Friedman, and Wilcoxon tests were significant at $p < 0.001$, and Cohen's $d = 0.82$ indicated a large practical difference for the principal AEZTOF contrast. Holm-adjusted post-hoc comparisons remained significant across the seven baselines. Reporting effect magnitude alongside corrected significance prevents the large run count from being mistaken for practical importance. The convergence of parametric, rank-based, paired, classification, calibration, tail-latency, sensitivity, and ablation evidence supports all six hypotheses.

6.1 Mechanism-level interpretation

The observed gains can be explained through three interacting control mechanisms. First, evidence reuse reduces duplicated work. A conventional access stack may independently request identity, posture, and session evidence at several gateways, even when those signals were recently validated. AEZTOF maintains bounded evidence freshness and makes it available to the authentication and routing decisions. This does not create permanent trust because confidence decays and high-sensitivity resources retain mandatory checks. It removes redundant verification inside an explicitly defined validity window. The 31.9% authentication-latency reduction relative to ZTNA is consistent with this mechanism, although component-level tracing is required to assign causality.

Second, joint path selection reduces downstream correction. A latency-only controller can send traffic through a path that later fails a security constraint, forcing re-evaluation or re-routing. A security-only controller can repeatedly choose a trusted but congested path. By including latency, congestion, trust risk, and policy cost in one utility function, AEZTOF selects among feasible paths before forwarding. This can reduce oscillation and explains why throughput, loss, and jitter improve together rather than trading off against one another. Future experiments should log candidate sets, selected utilities, constraint failures, and route changes so this explanation can be tested directly.

Third, policy-graph simplification reduces the number of evaluations and ambiguous outcomes. Exact duplicates add evaluation cost without changing semantics; shadowed rules complicate audit; contradictions can cause controller retries or inconsistent enforcement across locations. AEZTOF groups rules by subject, resource, action, condition, and precedence, then uses graph relations to flag these patterns. The governance gate is essential: semantic equivalence is harder than syntactic similarity, and a rare exception may look redundant in ordinary traffic. Accordingly, the system recommends merges or retirement, simulates their impact against historical and adversarial traces, and deploys approved changes with versioned rollback.

6.2 Ablation and sensitivity requirements

The ablation experiment isolated adaptive authentication, trust-aware routing, policy optimization, AI classification, and feedback adaptation. Removing adaptive authentication increased authentication latency from 17.9 to 23.4 ms. Removing trust-aware routing increased end-to-end latency from 22.6 to 28.1 ms. Removing policy optimization raised conflicts from 71 to 146. Removing the AI classifier produced the largest security loss, reducing detection from 98.9% to 94.7% and increasing false positives from 1.8% to 4.1%. Removing feedback adaptation degraded all six reported indicators. These component-specific changes align with the architecture's causal design and show that the full-system gain was not attributable to one module alone.

Table 17. Component-ablation results

Variant	Auth latency	E2E latency	Detection	FPR	Conflicts	CPU
Full AEZTOF	17.9 ms	22.6 ms	98.9%	1.8%	71	57%
Without adaptive authentication	23.4 ms	23.8 ms	98.7%	1.9%	72	59%
Without trust-aware routing	18.6 ms	28.1 ms	98.1%	2.1%	75	61%
Without policy optimization	18.9 ms	24.8 ms	98.5%	2.0%	146	64%
Without AI classifier	19.2 ms	24.1 ms	94.7%	4.1%	78	60%

Without feedback adaptation	20.8 ms	26.3 ms	96.8%	3.0%	101	63%
-----------------------------	---------	---------	-------	------	-----	-----

Sensitivity analysis varied trust thresholds, risk and latency weights, step-up thresholds, classifier confidence, and policy inactivity windows across the predefined grid. AEZTOF retained its network-latency advantage in 85% of parameter combinations while threat detection remained at least 97.5% and policy compliance at least 98.5%. Performance weakened mainly at the extremes of the risk-latency trade-off, but no tested setting simultaneously violated the minimum security and latency criteria. The resulting response surface identifies a stable operating region around trust threshold 0.70, route-risk weight 0.35, latency weight 0.30, classifier confidence 0.80, and a 30-day inactivity window.

6.3 Reproducibility and operational governance

Reproducibility was supported through a run manifest linking each unique run identifier to topology, workload, dataset partition, attack schedule, policy version, model version, controller configuration, random seed, software image, and metric outputs. The evaluation fixed the hardware allocation, network capacity, clock synchronization, warm-up process, and one-second power-sampling interval across all methods. Dataset preprocessing enforced session-level separation among training, validation, and test sets, and two attack families were held out for generalization testing. These controls allow the reported method comparisons, sensitivity grid, and ablation conditions to be reconstructed from the study package.

Operational governance should define who owns trust features, approves policies, investigates model drift, and authorizes emergency overrides. Every access decision should preserve a compact explanation containing the decisive evidence, applicable policy, trust band, authentication action, selected route class, and model uncertainty. Logs should be tamper-evident and access-controlled because they reveal identities and network structure. Emergency bypasses must be time-limited, narrowly scoped, and reviewed after use. The system should measure not only automation coverage but also override frequency, rollback rate, approval time, and the proportion of recommendations rejected by human reviewers.

A practical service-level objective can combine security and performance without collapsing them into one opaque score. Mandatory security constraints should be treated as hard boundaries; latency, cost, and energy can then be optimized within the feasible set. This ordering avoids a dangerous interpretation in which a sufficiently large latency benefit compensates for violating a critical policy. Pareto analysis can still expose trade-offs among feasible alternatives. For regulated deployments, policy-as-code repositories, separation of duties, signed releases, retention schedules, and independent control testing can connect the orchestration loop to established audit processes.

6.4 Broader research implications

AEZTOF reframes zero trust as a distributed control problem rather than a sequence of security products. The relevant state is partially observed, time-varying, and affected by the system's own actions. Authentication changes traffic timing; routing changes observable paths; blocking changes future behaviour; and policy cleanup changes the decision surface. This feedback means offline classifier accuracy alone is an insufficient evaluation criterion. Research should examine stability, observability, delayed telemetry, conflicting controllers, and safe exploration. Formal methods could verify invariant policies, while learning methods optimize within those verified boundaries.

The framework also highlights the difference between trust and authorization. A high trust score should never create blanket access. Authorization remains specific to subject, resource, action, purpose, and context; trust only influences the evidence and controls used to evaluate that request. Maintaining this separation reduces privilege accumulation and makes decisions easier to audit. Future work can extend the model to workload identities, machine-to-machine credentials, confidential computing attestations, and cross-organization federation, where evidence provenance and policy translation become especially important.

7. Threats to Validity and Limitations

Internal validity was protected through identical topology, workload, attack, hardware, warm-up, seed, collector, capacity, and observation controls across methods. Randomized scenario order reduced temporal bias, while both parametric and rank-based analyses reduced dependence on one statistical assumption. Remaining threats include residual controller-cache effects, imperfect equivalence between commercial reference configurations, and dependence of the composite complexity index on its declared weights. The sensitivity analysis reduced this concern by confirming the main performance ordering across the tested parameter region.

External validity remains bounded by the evaluated enterprise topology, 30-day observation window, fourteen attack scenarios, and the selected public traffic sources [24]-[29]. Results may differ in operational technology, highly regulated banking, low-bandwidth branches, or environments with strong legacy dependencies. The held-out attack evaluation, encrypted-traffic test, concept-drift analysis, and scaling to 150,000 sessions broaden the evidence base, but multi-organization replication and longer operational observation remain appropriate next steps. Energy results are reported for the defined hardware boundary and one-second sampling interval and should be compared only under equivalent workload and infrastructure scopes.

The trust engine creates privacy and fairness risks if behavioural or location features systematically disadvantage legitimate users. Data minimization, retention limits, feature justification, appeal mechanisms, and subgroup error analysis are therefore required. Explainable reason codes should accompany denial and step-up decisions. Finally, central orchestration can become a high-value target and failure domain; replicated controllers, signed telemetry, protected model supply chains, degraded-mode policies, and independent audit logs are necessary production controls [30]-[32].

8. Conclusion and Future Work

AEZTOF integrates environment discovery, continuous trust, risk-adaptive authentication, trust-aware SDN routing, AI traffic classification, policy optimization, enforcement, and monitoring in one closed loop. Across 2,400 method-level runs, it reduced authentication and end-to-end latency while improving throughput, detection, compliance, policy simplicity, scalability, and resource use. Authentication success reached 99.4%, decision accuracy 99.5%, threat detection 98.9%, and policy compliance 99.3%, while end-to-end latency remained 22.6 ms and response time stayed below 100 ms through 150,000 concurrent sessions. Ablation, sensitivity, unseen-attack, calibration, controller-failure, poisoning, drift, and encrypted-traffic results demonstrate that these gains were robust rather than dependent on a single favorable operating point.

Future work should evaluate multiple independent organizations, longer observation periods, additional encrypted protocols, operational-technology environments, federated identity domains, and policy translation across administrative boundaries. Formal verification of invariant controls and prospective deployment studies can further test long-term controller stability, human override behavior, and organizational adoption. These extensions build on the present evidence without changing the central conclusion: zero trust can be stricter, faster, and easier to govern when security and network decisions share context and objectives.

9. Declarations

Funding: The authors received no external funding for this study. Conflicts of interest: The authors declare no conflict of interest. Data availability: The study data, run manifest, parameter settings, dataset-role mapping, and aggregate evaluation outputs are available from the corresponding author subject to applicable dataset licenses and security restrictions. Ethics: The study analyzed network and security telemetry without human-subject intervention or collection of directly identifying participant data; institutional requirements should be confirmed before submission. Generative-AI disclosure: The authors should complete this statement in accordance with the selected journal's policy before submission.

REFERENCES

1. S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, 2020. doi: 10.6028/NIST.SP.800-207.
2. R. Chandramouli and Z. Butcher, "A Zero Trust Architecture Model for Access Control in Cloud-Native Applications in Multi-Location Environments," NIST SP 800-207A, 2023. doi: 10.6028/NIST.SP.800-207A.
3. Cybersecurity and Infrastructure Security Agency, "Zero Trust Maturity Model, Version 2.0," CISA, Apr. 2023.
4. C. Itodo, et al., "Multivocal literature review on zero-trust security," *Computers & Security*, vol. 141, art. 103827, 2024. doi: 10.1016/j.cose.2024.103827.
5. C. Katsis et al., "ZT-SDN: An ML-Powered Zero-Trust Architecture for Software-Defined Networks," *ACM Transactions on Privacy and Security*, 2025. doi: 10.1145/3712262.
6. E. Gilman and D. Barth, *Zero Trust Networks*, 2nd ed. Sebastopol, CA: O'Reilly Media, 2024.
7. NIST, "Planning for a Zero Trust Architecture: A Starting Guide for Federal Administrators," *Cybersecurity White Paper*, 2022.
8. National Cybersecurity Center of Excellence, "Implementing a Zero Trust Architecture," NIST SP 1800-35, practice guide series, 2025.
9. R. Ward and B. Beyer, "BeyondCorp: A New Approach to Enterprise Security," *login:*, vol. 39, no. 6, pp. 6-11, 2014.
10. B. Osborn, J. McWilliams, B. Beyer, and M. Saltonstall, "BeyondCorp: Design to Deployment at Google," *login:*, vol. 41, no. 1, pp. 28-34, 2016.
11. Microsoft, "Zero Trust Guidance Center," Microsoft Learn, 2025.
12. Cisco, "Cisco Zero Trust Architecture Guide," Cisco Security, 2024.
13. N. McKeown et al., "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM CCR*, vol. 38, no. 2, pp. 69-74, 2008.
14. D. Kreutz et al., "Software-Defined Networking: A Comprehensive Survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14-76, 2015. doi: 10.1109/JPROC.2014.2371999.
15. A. Clemm, L. Ciavaglia, L. Z. Granville, and J. Tantsura, "Intent-Based Networking: Concepts and Definitions," RFC 9315, Oct. 2022.
16. T. T. Nguyen and V. J. Reddi, "Deep Reinforcement Learning for Cyber Security," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 8, pp. 3779-3795, 2023.
17. A. Mestres et al., "Knowledge-Defined Networking," *ACM SIGCOMM CCR*, vol. 47, no. 3, pp. 2-10, 2017.
18. Z. Mammeri, "Reinforcement Learning Based Routing in Networks: Review and Classification," *Computer Networks*, vol. 177, art. 107326, 2020.
19. V. C. Hu et al., "Guide to Attribute Based Access Control Definition and Considerations," NIST SP 800-162, 2014.
20. R. Chandramouli, "Attribute-Based Access Control for Microservices-Based Applications Using a Service Mesh," NIST SP 800-204B, 2021.
21. A. Mahfouz, T. M. Mahmoud, and A. S. Eldin, "A Survey on Behavioral Biometric Authentication on Smartphones," *Journal of Information Security and Applications*, vol. 37, pp. 28-37, 2017.
22. E. Al-Shaer, H. Hamed, R. Boutaba, and M. Hasan, "Conflict Classification and Analysis of Distributed Firewall Policies," *IEEE JSAC*, vol. 23, no. 10, pp. 2069-2084, 2005.
23. C. N. Modi et al., "A Survey of Intrusion Detection Techniques in Cloud," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 42-57, 2013.
24. N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," *MilCIS*, pp. 1-6, 2015.
25. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset," *ICISSP*, pp. 108-116, 2018.
26. E. C. P. Neto et al., "CICIOT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment," *Sensors*, vol. 23, no. 13, art. 5941, 2023.
27. N. Moustafa, "A New Distributed Architecture for Evaluating AI-Based Security Systems at the Edge," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6570-6580, 2021.
28. K. Cho et al., "Traffic Data Repository at the WIDE Project," *USENIX Annual Technical Conference*, 2000.
29. CAIDA, "The CAIDA Anonymized Internet Traces Dataset," Center for Applied Internet Data Analysis, 2025.
30. NIST, "Cybersecurity Framework 2.0," NIST CSWP 29, 2024. doi: 10.6028/NIST.CSWP.29.
31. NIST, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023. doi: 10.6028/NIST.AI.100-1.
32. MITRE, "MITRE ATT&CK Enterprise Matrix," Version 17, 2025.
33. Cloud Security Alliance, "Software-Defined Perimeter and Zero Trust," CSA Research, 2022.
34. CISA, "Microsegmentation in Zero Trust, Part One: Introduction and Planning," July 2025.
35. M. Gambo et al., "Zero Trust Architecture: A Systematic Literature Review," *Journal of Network and Systems Management*, 2025. doi: 10.1007/s10922-025-09998-x.