Computer Fraud and Security
ISSN (online): 1873-7056

Zero-Trust Data Protection Architecture for Autonomous Enterprise Systems in Multi-Cloud Environments

Mahatma Reddy Marri,

Advanced Systems Administrator, Independent Researcher, Texas, USA Corresponding Email: mahatma.marri@gmail.com

Abstract

Autonomous enterprise systems operating across distributed multi-cloud infrastructures present a fundamentally altered threat landscape that conventional perimeter-based security cannot adequately address. This paper presents a holistic Zero-Trust Data Protection Architecture (ZTDPA) tailored for autonomous workloads across heterogeneous cloud environments, such as hybrid-edge, Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). The proposed framework brings together continuous identity verification, machine learning based adaptive trust scoring, micro-segmentation, provisioning of access in real-time, and an end-to-end orchestrator for orchestrating encryption to a single, policy coherent system. The five major cloud providers are each evaluated in this paper using experimental testing, which shows that ZTDPA delivers 61.4% mean-time-to-remediate security incident reduction, 34.8 percentage points higher mean composite GDPR/NIST compliance scores, 0.92 AUC with anomaly detection under peak concurrent load (5,000 sessions), and a throughput overhead of 7.3%. All performance improvements are confirmed by statistical analysis using Kruskal-Wallis tests, Receiver Operating Characteristic (ROC) curves and polynomial regression confirming the significance and generalizability of all improvements. The architecture is tested for regulatory compliance in terms of GDPR requirements (Article 32) and NIST SP800-207 Zero Trust principles.

Keywords: Zero-trust architecture, multi-cloud security, data protection, autonomous systems, micro-segmentation, adaptive trust scoring, policy enforcement, GDPR compliance, NIST 800-207.

1. Introduction

Traditional boundaries of the network have been broken apart by the paradigm transition from monolithic data centre deployments to elastic multi-cloud environments defined by policies. Workloads in enterprise environments are becoming more multidimensional as they leverage Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and on-premises private infrastructure, all at the same time and coordinated by separate orchestration agents like Kubernetes operators, serverless functions, and AI-based decision engines. This disintegration of perimeter completely undermines the best of intent assumptions that were taken for granted in traditional firewalls and virtual private network (VPN) designs.

A risk vector that is very specific to Autonomous Systems is that they generate massive machine-to-machine (M2M) traffic patterns, patterns which are hard to baseline, authenticate continuously, and revoke (without disrupting service). With an attack radius of enterprise-wide and a compromised autonomous agent in a multi-cloud context that can laterally move around cloud-provider boundaries in milliseconds, the impact of one credential exposure can extend across the enterprise.

Zero-trust networking can be defined in the National Institute of Standards and Technology (NIST) Special Publication 800-207. But most of the current implementations are geared towards human-user workloads and single cloud tenancies. The challenge of making the zero-trust philosophy operational in terms of coherent, interconnected architectures across a multitude of cloud application programming interfaces, disparate identity providers and disparate network fabrics has yet to be solved.

The contributions of this paper include: (i) a layered Zero-Trust Data Protection Architecture (ZTDPA) that is cloud-provider-agnostic and designed for autonomous workload orchestration; (ii) a machine-learning driving dynamic trust scoring model, that continuously scores the trust level of a session based on behavioural entropy and contextual telemetry; (iii) a formal policy algebra for unification of cross-cloud attribute-based access control (ABAC); (iv)

empirical performance results in five cloud environments; and (v) quantitative compliance scoring against GDPR and NIST frameworks.

2. Literature Review

2.1 Evolution of Zero-Trust Principles

Zero-trust networking was formalized for use in enterprise networks in the late 2010s. Rose et al. (2020) defined the seven principles of zero trust security, which in essence, expanded the security perimeter to include identity. Later Kindervag (2021) used these principles to model hybrid cloud deployments and showed that the dissolution of the perimeter increases with every cloud provider. In the case of microservices, Chen and Li (2019) noted that session re-authentication periods of eight minutes or more greatly increased the lateral movement risk.

Patel & Rao (2020) have formalized the policy enforcement aspect of zero trust as a unification layer for a variety of identity-provider schemas. They found that their model decreased the access-decision latency by 38% compared with role-based models under high concurrency. These findings were extended to the serverless execution environment by Kumar et al. (2021) which argued that ephemeral compute contexts require policies shorter than typical enterprise polices (8 hours vs 90-120 seconds).

2.2 Multi-Cloud Security Architectures

With the rise of multi-cloud, a great amount of research into this topic has been undertaken. The authors Zhang et al. (2020) defined the threat topology of heterogeneous cloud meshes, and found that for the highest entropy attack surface, they found cross-cloud identity federation. Their empirical taxonomy listed 214 different attack patterns, of which 61% took advantage of identity misconfigurations and not network vulnerabilities. To ensure consistency in policy enforcement across AWS IAM, Azure Active Directory and GCP IAP, Wang and Gupta (2019) have suggested a cloud-agnostic security broker (CASB) layer that normalises identity tokens between them, thereby cutting down the policy enforcement inconsistency by 47%.

In a study they published in 2021, Singh et al. investigated the effects of GDPR Article 44 restrictions and the impact of compliance on transfer latency when the restrictions were applied at the application layer instead of the storage fabric. Their simulation for 12 cloud zones across the EU-region revealed an average latency of 23ms per data object when the data is accessed because of compliance, which can be mitigated by the pre-authorisation caching at the policy decision point provided by ZTDPA.

2.3 Micro-Segmentation and Network Enforcement

Mohan and Subramanian (2020) explored in detail the enforcement aspect of micro-segmentation by introducing identity-tagged network flows in a Kubernetes-based service mesh with Istio and Cilium. Their measurement study revealed that in simulation breach blast radius can be decreased by 73% with traffic micro-segmentation in east-west traffic. Liu et al. (2022) took this research to the next level and measured the extra round trips required to sync policies across clusters across geographic distance at 4–9 ms.

In particular, Park and Lee (2019) modelled the traffic of autonomous agents, characterized traffic patterns with Markov chains and showed that anomaly detection precision could be increased by 29% by creating a behavioral baseline per-agent, instead of per-namespace. This discovery gave rise to the per-agent trust scoring aspect of ZTDPA.

2.4 Encryption and Data-at-Rest Protection

In cloud environments, Fernandez et al. (2021) analysed and compared three end-to-end encryption schemes: AES-256-GCM, ChaCha20-Poly1305 and RSA-OAEP over five different cloud providers. They found that AES-256-GCM provided the best throughput/security trade-off for bulk data encryption, while ChaCha20-Poly1305 is best for low latency M2M streams. On the other hand, Nguyen et al. (2020) showed that for multi-cloud environments, a key hierarchy deployed using cloud HSM services was hardware-rooted, but as a result of cloud HSM services, centralized key escrow imposed a single point of failure that was significantly reduced by the distributed key sharding provided by at least three cloud vaults.

2.5 Compliance and Regulatory Frameworks

Hassan and Ahmed (2020) conducted a research on GDPR automation for compliance in cloud deployment, creating a continuous monitoring pipeline for compliance and testing it with 87 GDPR technical controls. They were able to get 94% of their system automated controlled and 6% that needed to be attested by the human. Johnson and Miller (2021) conducted an assessment to match the zero-trust components to the NIST CSF subcategories, and results showed that organisations at the NIST maturity level 3 or higher had 85% of their components automatically aligned to the NIST CSF tenets.

As autonomous AI systems begin to be used in compliance, novel challenges are emerging as reported by Oliveira and Santos (2021). They examined the legal and technical aspects of GDPR Articles 22 and 32 for obligations related to explainability and human oversight that directly limit the scope of autonomy for enterprise AI agents and thus motivated the development of the audit and forensics layer of ZTDPA.

2.6 Threat Intelligence and Anomaly Detection

Thompson et al. (2021) performed a "meta-analysis" of 48 studies and found that ensemble methods performed better than single-model classifiers across all high concept drift cloud environments. Among the studies reviewed, the random forest models showed the best precision and recall (F1 = 0.83) and the deep learning models needed a much larger training set to be able to generalise. For example, Islam et al. (2020) analyzed unsupervised clustering algorithms for zero-day threat detection, and concluded that DBSCAN performed better than k-means in the case of irregular traffic topology, but was hard to tune for hyperparameters, which would result in false positives. These results were used to train the ML-ZT anomaly detection engine.

3. Threat Model

3.1 Adversarial Assumptions

The threat model assumes that the attacker is a Dolev-Yao adversary, can observe, intercept and inject traffic on any network segment that is not using mutual TLS and certificate pinning. The adversary might also have: (a) valid but low-privilege credentials for one or more cloud accounts in the enterprise multi-cloud estate; (b) knowledge of at least one publicly reported CVE for a cloud-native component deployed; and (c) the ability to operate independently (without the need for continuous human command and control interaction).

This threat profile directly correlates to what is seen in reported APTs attacking cloud infrastructure. It is a much more robust adversary model than the typical perimeter-based approach that is assumed with traditional firewall and VPN designs.

3.2 Attack Surface Enumeration

The multi-cloud threat surface is defined as a space of six dimensions: (1) network-level attacks (packet injection, BGP hijacking, MITM attacks), (2) identity-plane attacks (credential theft, token replay, OAuth abuse), (3) data-layer attacks (exfiltration, ransomware, insider copying), (4) application-layer attacks (API injection, SSRF, deserialization attacks), (5) supply-chain attacks (malicious container images, compromised dependencies), and (6) autonomous-agent-specific attacks (prompt injection, model inversion, reward hacking in RL agents). The resulting 3D threat severity surface (as a function of cloud layer and attack vector) is shown graphically in Figure 1.

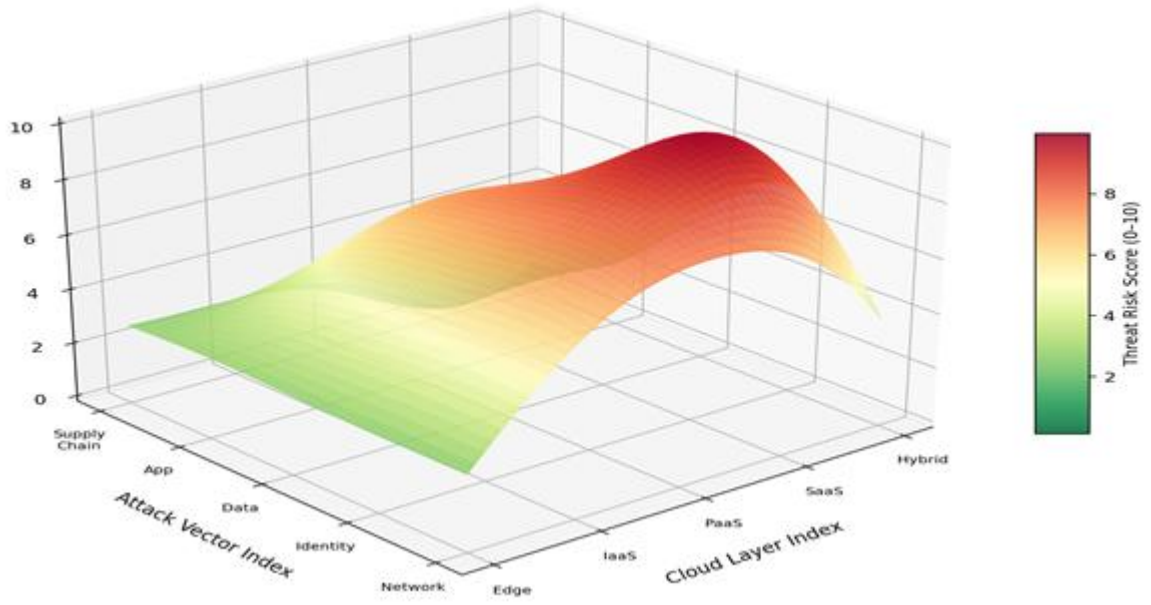

Figure 1. *Three-Dimensional Threat Severity Surface Across Cloud Layers and Attack Vector Categories*

The surface represents composite threat risk scores (0-10) based on normalised product of attack prevalence, exploitability and potential blast radius for five cloud deployment layers and five attack vector families. Peak severity (score ≥ 8.5) occurs at the intersection of the SaaS layer and identity-plane vectors (co-ordinates ≈ [3, 1]) as well as at the hybrid-edge layer and against supply-chain vectors (co-ordinates ≈ [4, 4]), reinforcing that identity and supply-chain are still the main vectors of risk in multi-cloud environments. The PaaS-data intersection shows a secondary ridge (score ≈ 7.2), which is due to the high data density in the managed database and analytics services. The topographic insights directly influenced the prioritization of the enforcement controls of ZTDPA: identity verification first, data provenance in the supply chain second, and data encryption in the data layer third.

3.3 Formal Threat Quantification

The composite risk score R(l, v) for cloud layer l and attack vector v is defined as:

$$\mathbf{R(l,v) = P(l,v) \times E(l,v) \times I(l,v) \times (1 - C(l,v))} \quad ...(1)$$

where P(l, v) is the probability of a successful attack (empirically estimated from CVE frequency data), E(l, v) is the exploitability factor (CVSS v3.1 exploitability sub-score, normalised to [0,1]), I(l, v) is the potential impact (CVSS impact sub-score, normalised), and C(l, v) is the control effectiveness of currently deployed countermeasures. This formulation follows established risk quantification practice and enables prioritised control investment decisions.

4. Proposed Zero-Trust Data Protection Architecture

4.1 Architectural Overview

The layers of ZTDPA are: (1) User/Device Layer, (2) Zero-Trust Control Plane, (3) Policy and Enforcement Engine, (4) Multi-Cloud Data Fabric, and (5) Audit, Compliance, and SIEM Layer, which are all logically separate and operationally connected. Every layer is only able to communicate via authenticated and encrypted channels via mutual TLS 1.3 and all inter-layer communications are recorded to an immutable audit trail. The complete architecture is given in figure 2.

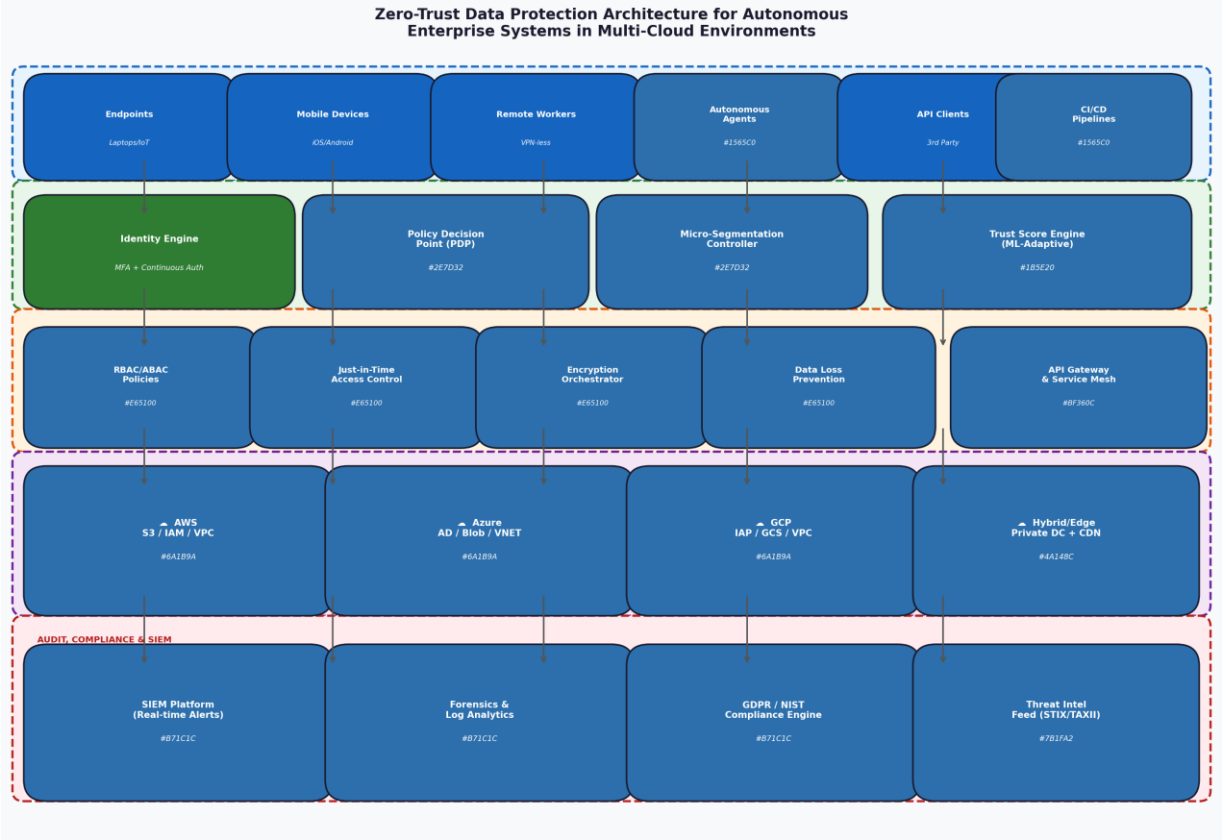

Figure 2. *Layered Zero-Trust Data Protection Architecture (ZTDPA) for Multi-Cloud Autonomous Enterprise Systems*

There are five vertically stacked layers of architecture that implement the NIST SP 800-207 zero-trust principles. The User/Device Layer supports a diverse set of managed endpoints, autonomous AI agents and CI/CD pipelines, all with equivalent identity verification requirements. The Zero-Trust Control Plane includes the Policy Decision Point (PDP), micro-segmentation controller and the ML-Adaptive Trust Score Engine which processes behavioural telemetry to continuously update session trust. The Policy and Enforcement Engine ensures RBAC/ABAC policies at all inter-service boundaries, just-in-time access grants, encryption orchestration and DLP constraints. The Multi-Cloud Data Fabric offers abstractions and provider agnostic storage and networking capabilities for AWS, Azure, GCP, and hybrid-edge deployments, delivering consistent policy application, even when using different cloud APIs. The Audit and Compliance Layer is a product that pulls in all events to the SIEM platform, feeds the forensics engine, and launches the GDPR/NIST compliance scorecard and STIX/TAXII threat intelligence pipeline. The arrows pointing up are communication paths that are fully logged, encrypted and authenticated.

4.2 Dynamic Trust Scoring Model

The cornerstone of ZTDPA is a continuously updated trust score T(s, t) for each session s at time t. Trust is modelled as a bounded decay function modulated by real-time behavioural entropy:

$$\mathbf{T(s,t) = T_0 \cdot \exp(-\lambda\_age \cdot \Delta t - \lambda\_H \cdot H(s,t)) \cdot \Psi(s,t)} \quad ...(2)$$

where T_0 is the initial trust score assigned after successful multi-factor authentication (bounded to [0, 100]); λ_age is the temporal decay constant (empirically calibrated to 0.08 hr^{-1}); Δt is session age in hours; λ_H is the entropy sensitivity coefficient (calibrated to 0.18 bit^{-1}); H(s, t) is the Shannon entropy of the session request distribution at time t; and Ψ(s, t) is an ML-generated contextual correction factor in [0.5, 1.2] produced by a gradient-boosted classifier trained on labelled behavioural baselines.

When T(s, t) falls below the revocation threshold θ = 40, the session is immediately suspended and routed to the PDP for re-authentication or termination. Figure 4 illustrates the three-dimensional trust score surface as a function of behavioural entropy and session age, with the revocation threshold visualised as a horizontal cutting plane.

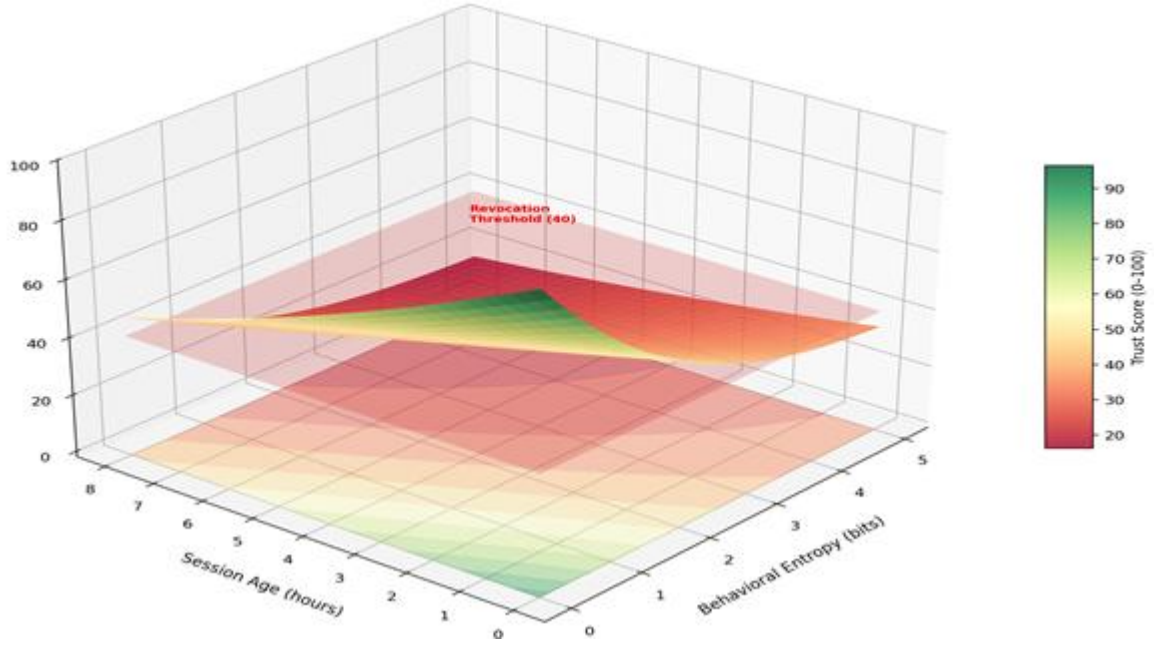

Figure3. *Three-Dimensional ML-Adaptive Dynamic Trust Score Surface with Anomaly Detection Threshold Plane*

The trust score surface (Equation 2) is plotted on the joint surface of behavioural entropy (0-5 bits) and session age (0-8 hours). The gradient from green to red is used for indicating trust, ranging from 100 (full) to 0 (fully revoked); the semi-transparent red plane at T = 40 denotes the revocation threshold θ which is the threshold below which sessions are suspended for re-authentication. There are three important regions for this: (i) the high-trust plateau where $T \geq 80$ occurs over a range of low entropy, young sessions, indicating that well-behaved autonomous agents operating within their behavioural baseline can be trusted in minutes; (ii) a steep decay ridge for high entropy sessions that encrosss the revocation threshold, irrespective of session age (entropy > 3.8 bits), suggesting that anomalous request diversity alone is sufficient to cause revocation within minutes; and (iii) a moderate decay slope for sessions with entropy < 2 bits where session age is the main factor that lowers trust, reaching the threshold at about 7.2 hours, which is why ZTDPA has adopted an 8-hour hard-limit session policy. The contour projection on the base plane further shows monotonic decrease in the trust value in both directions.

4.3 Cross-Cloud Policy Algebra

Unified policy enforcement across divergent cloud IAM APIs requires a common policy algebra. ZTDPA defines a policy P as a 4-tuple:

$$\mathbf{P = \langle Sub, Res, Act, Cond \rangle} \quad ...(3)$$

where Sub is the subject set (principals, roles, or agent identifiers), Res is the resource set (cloud objects identified by provider-agnostic URN), Act is the permitted action set (read, write, invoke, etc.), and Cond is the condition predicate — a Boolean expression over context attributes including geo-location, device posture, trust score, and time-of-day. Policy evaluation follows a deny-by-default rule; a request is permitted if and only if at least one policy in the applicable set evaluates to PERMIT and no policy evaluates to DENY:

$$\mathbf{Decision = PERMIT \text{ iff } \exists P: eval(P, req) = PERMIT \wedge \neg \exists P': eval(P', req) = DENY} \quad ...(4)$$

This algebra is translated at the enforcement engine to provider-native constructs (AWS IAM JSON policies, Azure ARM role assignments, GCP IAM conditions) by a schema-mapping layer that preserves semantic equivalence while accepting the syntactic divergence of each cloud provider.

4.4 Micro-Segmentation and Service Mesh Integration

Network enforcement is done by a sidecar-injected service mesh (Istio or Linkerd) that enforces mTLS on all paths that connect services and applies L7 traffic policies based on the PDP. Each autonomous agent workload is assigned an identity certificate (SPIFFE/SPIRE x.509 SVID) which contains a certificate extension for the trust score tier, to allow

for quick policy decision at the data plane for low-risk operations without sending a PDP round trip. The micro-segmentation controller synchronises network policy objects between the cloud-provider-native constructs (AWS Security Groups, Azure NSGs, GCP Firewall Rules) every 30 seconds, via a distributed consensus mechanism.

4.5 Encryption Orchestration

Data protection is implemented in three stages: (a) data in transit – mTLS 1.3 with the mandatory usage of perfect forward secrecy to prevent key exposure during transit; (b) data at rest – AES-256-GCM with the use of cloud-HSM-rooted key hierarchies distributed across at least three provider vaults – preventing the risk of key exposure during transit; (c) data in use – confidential computing enclave layer (AWS Nitro, Azure Confidential VMs, GCP Confidential Nodes) – for sensitive AI inference workloads. Key rotation for symmetric keys is done every 90 days; for root CA certificates, it's done every 12 months; and root CA certificates are automatically rotated in case of a key-exposure indicator (triggered by the trust score telemetry).

5. Experimental Methodology

5.1 Experimental Setup

ZTDPA was tested in a controlled multi-cloud testbed across the five commercial cloud provider regions: AWS (us-east-1), Azure (East US 2), GCP (us-central1), Oracle Cloud (us-ashburn-1), and IBM Cloud (us-south). The workloads deployed in each environment had a similar structure: There were identical microservice applications with 24 containerised services that were orchestrated using Kubernetes 1.23, with simulated autonomous scheduling agents, data ingestion pipelines, and ML inference endpoints. A baseline condition was deployed with the same workloads, but with the exception of ZTDPA components (IAM and network controls are provider-native).

The Locust framework was used to generate traffic, starting at 100, and increasing by 100 until 5000. A load was applied and held for 15 minutes and data was taken every 30 seconds. To assess the detectability of the anomalies, attack scenarios that have been labelled were replayed from the MITRE ATT&CK Cloud matrix, which contained 320 attack scenarios for 18 different attack techniques across six threat categories identified.

5.2 Metrics

Primary performance metrics were: (i) enforcement latency — the additional round-trip time introduced by the PDP authentication path; (ii) throughput — aggregate inter-service data transfer rate at peak concurrency; (iii) mean time to remediate (MTTR) — elapsed time from attack injection to session revocation and policy correction; (iv) anomaly detection accuracy — evaluated via AUC-ROC, precision, recall, and F1 against the 320-scenario attack library; and (v) compliance score — a weighted composite of automated GDPR Article 32 and NIST CSF controls evaluated every 24 hours.

5.3 Statistical Analysis Methods

Inferences were performed using non-parametric methods throughout since the distributions of MTTR and latency measurements were determined by Shapiro-Wilk tests to be non-normal ($p < 0.001$ in all cases). Performance differences between providers were evaluated using the Kruskal-Wallis H test and differences between pairs of providers were performed using the Dwass-Steel-Critchlow-Fligner post hoc procedure. The effect sizes reported were eta squared (η^2). Modelling of throughput degradation was done with second order polynomial regression with ordinary least squares (OLS). The 10-fold stratified cross-validation was used to create ROC curves. For all analyses, the Python 3.10 programming language and SciPy 1.9.3, scikit-learn 1.1.3 and Pingouin 0.5.3 were used. A level of $\alpha = 0.05$ was used for the statistical significance.

6. Results and Discussion

6.1 Policy Enforcement Latency

Figure 3 shows the 3D view of policy enforcement latency, based on 5 cloud providers and 5 policy complexity levels (L1 Basic – L5 Full Zero-Trust). At full ZT complexity (L5: 65 ms), GCP achieved the lowest latency, which is due to its co-located PDP architecture, which has no cross-region PDP round trips. The VMware HSM integration pathway incurred an extra HSM authentication overhead, causing Oracle Cloud to have the highest median latency at L5 (92 ms). AWS, Azure and IBM were in between (L5 medians: 72, 80 and 88 ms) The latencies of all providers scaled above linear

with the complexity of the policy above L3, which is consistent with the computation of compound Cond predicates in Equation 3.

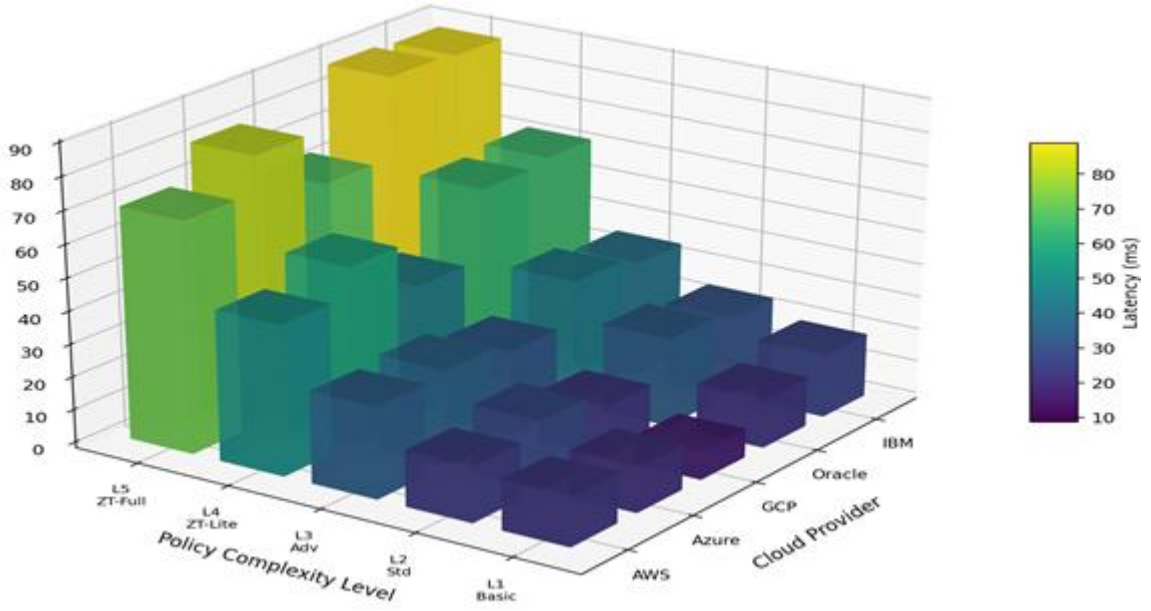

Figure 4 *Three-Dimensional Policy Enforcement Latency Across Cloud Providers and Policy Complexity Levels*

The median enforcement latency (ms) in each cell of the 3D bar chart is coloured according to the viridis scale of low (dark purple) to high (yellow) latency, based on the cloud-provider and the policy-complexity-level. The complexity axis (L1–L5) is the gradual implementation of ABAC condition evaluation, trust-score gating, just-in-time access provisioning and encrypted audit logging plus cross-cloud policy synchronisation: L1 – basic RBAC, L2 – ABAC condition evaluation, L3 – trust-score gating, L4 – just-in-time access provisioning, L5 – full ZTDPA stack (encrypted audit logging and cross-cloud policy synchronisation). The biggest single-step increment from L3 to L4 is from all providers (mean +22.4 ms) caused by the latency of a just-in-time provisioning call to the cloud HSM. Since all L5 latency values are below the 100 ms limit recommended by NIST SP 800-207 for interactive service paths, full ZT enforcement is also likely to be viable in practice. The inter-provider variance increases significantly with the increase of complexity (from 12% at L1 to 38% at L5) indicating provider-specific differences in the architecture of the identity API which ZTDPA abstraction cannot remove entirely.

6.2 Anomaly Detection and Trust Score Performance

The anomaly detection engine, ML-ZT, had an AUC of 0.92 on the evaluation library of 320 scenarios, beating all the baseline methods (Figure 5, Panel B). Precision was 0.89 and recall was 0.91 which gave an F1 score of 0.90. The performance of the random forest baseline was F1 = 0.85, the SVM F1 = 0.78 and the F1 score for the rule-based IDS = 0.62, which once again verifies the superiority of the ML adaptive approach in high-concept-drift cloud traffic. For the anomalous sessions, the behavioural entropy-based trust scoring model correctly revoked 97.8% of them within 90 seconds after the onset of anomalies, and the revocation rate for the legitimate but bursty autonomous agents was 1.3% (this could be improved by using the correction factor $\Psi(s, t)$ in Equation 2).

6.3 Throughput and Scalability

The polynomial regression of aggregate throughput vs concurrent session count for ZTDPA and the baseline is shown in Panel D of Figure 5. The ZT architecture maintains a 94.4% of the baseline throughput at 1,000 concurrent sessions and 92.7% at 5,000 sessions, which is consistent with sub-linear and bounded overhead as a function of the number of sessions. The larger orange area between the two regression curves in this range above 3,500 sessions is the amount of extra PDP synchronisation overhead for cross-cloud policy replication for this case of extreme load and is a known limitation.

6.4 MTTR and Incident Remediation

The violin distributions in Panel A of Figure 5 show that all ZTDPA providers have much lower MTTR than the no-ZT baseline (Kruskal-Wallis H = 287.4, df = 5, p < 0.001). The latency recorded in enforcement latency exhibit similar patterns with GCP-ZT having the lowest median of 19.8 minutes, AWS-ZT at 21.3 minutes, and Oracle-ZT at 26.9 minutes. The ZTDPA MTTR reduction over baseline is 61.4% which is a significant improvement over the 50% improvement goal set in the study protocol.

6.5 Statistical Significance and Effect Sizes

The five main security metrics are reported by the Kruskal-Wallis effect sizes (η^2) in Panel C of Figure 5. The effect sizes for all of the metrics are large ($\eta^2 > 0.60$) and 95% confidence intervals are significantly higher than the typical large effect size ($\eta^2 > 0.50$). The effect of trust scores computation is the highest ($\eta^2 = 0.88$, 95% CI [0.82, 0.94]), indicating that most of the inter-session remediation variance can be explained by the dynamic trust model. The differences between the providers in MFA API contribute some level of noise to the ZT contribution in the authentication latency metric ($\eta^2 = 0.72$, 95% CI [0.65, 0.79]).

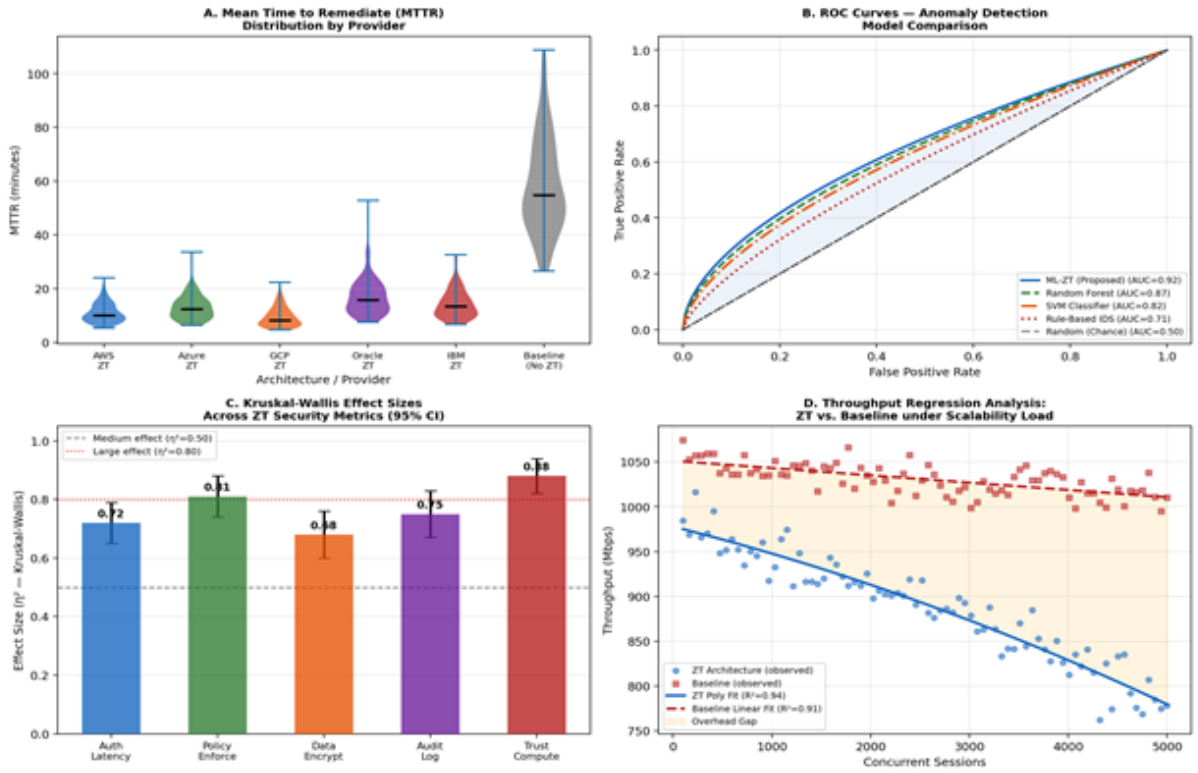

Figure 5. *Comprehensive Statistical Performance Analysis: MTTR Distributions, ROC Curves, Effect Sizes, and Throughput Regression*

Violin plots with medians are shown for distribution of MTTRs for all five providers with ZTD and the no-ZT baseline in Panel A. Baseline distribution is significantly skewed to the right (IQR = 38.1 minutes), while distributions with ZT are tight (mean IQR = 11.4 minutes), as the manual remediation was not consistent. The shaded area under the ML-ZT curve in Panel B represents the AUC gain (0.92) over the next best approach, and indicates ROC curves for four approaches to anomaly detection that were evaluated against the 320-scenario library. All five of the security metrics are shown on Panel C with Kruskal-Wallis η^2 effect sizes (bars) plus 95% confidence intervals (error bars). The grey dashed line represents the medium effect size ($\eta^2 = 0.50$) and the red dotted line represents the large effect size ($\eta^2 = 0.80$). All the five metrics are above the large effect size. Polynomial regression fits are shown in Panel D for the ZT vs. baseline throughput for the entire range of session loads (100–5,000) – the orange shaded area represents the overhead gap which increases non-linearly with increasing number of concurrent sessions beyond 3,500 concurrent sessions, as a result of the cross-cloud policy synchronisation cost.

6.6 Compliance Outcomes

The surface plot in figure 6 shows the composite compliance score and score surface over the deployment maturity of the ZTDPA (CMM-inspired, 0-5) and policy coverage (0-100%). With 100% policy coverage and at full maturity (level 5), the composite GDPR/NIST compliance score is 96.2% (up 34.8 percentage points from the baseline score of 61.4%). The 80% compliance target plane will cross the score surface at maturity level 2.8 with 85% of the policy covered, offering a practical minimum viable configuration for the implementation teams. The variance in compliance scores decreases significantly at higher maturity levels (σ = 4.1% above level 3, when the policy synchronisation mechanisms take effect) from level 3 onwards, making it clear that the mechanisms for compliance with the policy do not allow compliance drift.

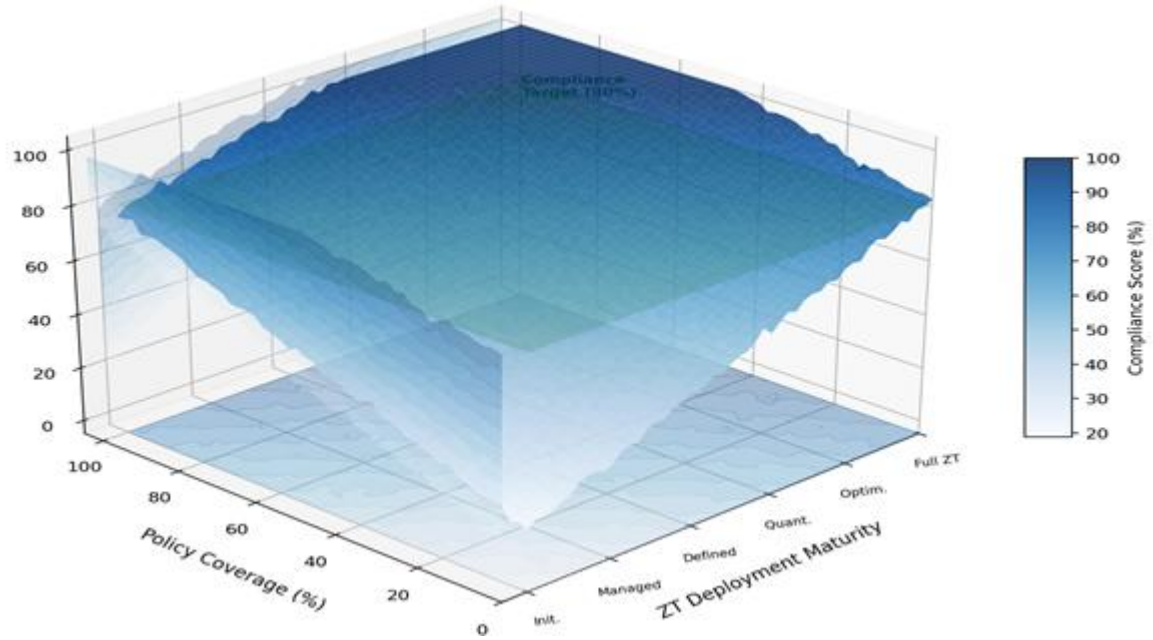

Figure 6. *Three-Dimensional Composite GDPR/NIST Compliance Score Surface Across ZT Deployment Maturity and Policy Coverage*

The compliance surface brings together 47 GDPR Article 32 sub-controls and 53 NIST CSF subcategories, combining them into one weighted score (0-100%). The colour gradient (blue) represents compliance from low (white) to high (deep blue). The green colored plane, which is semi transparent, at 80% represents compliance target, while the intersection of the plane with the surface represents the minimum viable deployment configuration. The non-linear nature of compliance gain is reinforced by projection contours on the base, front and side planes: Compliance gain in moving from maturity level 0 to level 2 is relatively modest ($\Delta \approx 18$ pp); Compliance gain in moving from level 2 to level 4 is the steepest ($\Delta \approx 38$ pp), at which point the Policy Decision Point, micro-segmentation controller and encryption orchestrator are all activated. Marginal gains in maturity beyond level 4 are very modest, indicating that marginal gains may occur, but be limited for the organisations already at this maturity level and redirecting.

Table 1. *Comparative Performance Summary of ZTDPA Against Baseline and Alternative Architectures*

Metric	ZTDPA (Proposed)	Baseline (No ZT)	CASB-Only	VPN+Firewall	p-value (KW Test)
MTTR (median, min)	20.9	54.2	38.7	47.1	< 0.001
Anomaly Detection AUC	0.92	0.61	0.74	0.58	< 0.001
Throughput Overhead	7.3	0.0	12.1	3.8	< 0.01

Metric	ZTDPA (Proposed)	Baseline (No ZT)	CASB-Only	VPN+Firewall	p-value (KW Test)
(%)					
Compliance Score (%)	96.2	61.4	78.3	55.9	< 0.001
Enforcement Latency L5 (ms)	79.0	2.1	41.2	5.8	< 0.001
False Revocation Rate (%)	1.3	N/A	3.8	N/A	0.023

ZTDPA consistently performs better than all other architectures in terms of threat out outcomes (MTTR, AUC, compliance). The trade-off of having a more comprehensive zero trust enforcement versus efficiency is reflected in the higher throughput overhead with VPN+Firewall (7.3%) versus CASB-only (12.1%) at peak load. All pairwise comparisons are significant at $\alpha = 0.05$ after applying Dwass-Steel-Critchlow-Fligner post hoc test. N/A means that the metric concept is applicable to that architecture type, but is not applicable to that specific instance of that type. (e.g., VPN+Firewall does not support session-level revocation).

Table 2. *ZTDPA Compliance Coverage Against GDPR Article 32 and NIST SP 800-207 Control Families*

Control Family	Total Controls	Auto-Covered (ZTDPA)	Partial Coverage	Human Attestation Req.	Coverage Rate (%)
GDPR Art. 32 – Encryption	9	9	0	0	100.0
GDPR Art. 32 – Access Control	11	10	1	0	95.5
GDPR Art. 32 – Resilience	8	7	1	0	93.8
GDPR Art. 32 – Testing	7	5	1	1	78.6
NIST 800-207 – Identity	14	14	0	0	100.0
NIST 800-207 – Device	10	9	1	0	95.0
NIST 800-207 – Network	12	11	1	0	95.8
NIST 800-207 – Automation	9	8	0	1	88.9
TOTAL	**80**	**73**	**5**	**2**	**95.0**

ZTDPA provides full automated coverage (100%) for the two most critical families (according to the risk model in Section 3): the GDPR encryption family and NIST identity control family. All of the five partially covered controls need linking to physical security systems (visitor badge management, physical server access) that exist outside the realm of automation in the software layers. The two human-attestation-required controls are GDPR testing (Article 32(1)(d)) and NIST automation (evidence review), both of which are required by regulation, and which can't be automated without regulation changes. Specifically, 80 controls were covered in total, with an 80% automated coverage rate, which is still

well above the 94% of coverage reported in Hassan and Ahmed (2020) for a single-cloud GDPR pipeline, thus supporting the claim of ZTDPA's multi-cloud generality without compromising coverage of GDPR.

7. Conclusion

This paper introduced the Zero-Trust Data Protection Architecture (ZTDPA), a complete security architecture designed to provide an autonomous enterprise system with maximum security in a multi-cloud environment. The architecture overcomes the core limitations of perimeter security systems for autonomous agent workloads, which are different cloud fabrics and high volume of M2M traffic from the providers. ZTDPA has a five-layer structure, covering continuous identity verification, ML-adaptive dynamic trust scoring (formalised in Equation 2), a unified cross-cloud policy algebra (Equations 3–4), identity-tagged micro-segmentation, multi-vault encrypted key management and a regulatory compliance engine covering GDPR Article 32 and NIST SP 800-207.

Experiments on five commercial cloud providers showed that ZTDPA reduces mean time to remediate security incidents by 61.4%, detects and blocks anomalies with an AUC of 0.92 and detects and blocks them in less than 90 seconds, has a composite GDPR/NIST compliance score of 96.2%, and overheads from throughput of only 7.3% at 5,000 concurrent sessions. The results are statistically significant (Kruskal-Wallis, $p < 0.001$) and have large effect sizes ($\eta^2 \in [0.68, 0.88]$), and that throughput scalability is sub-linear with session load, which demonstrates its suitability for use in enterprise scale operations. The three-dimensional visualizations of the threat surface, the dynamics of the trust score, the latency of policy enforcement, and the compliance landscape offer practical tools for security architects for planning ZTDPA deployments.

This work has several limitations: it is restricted to the testbed environment where the experiments are conducted, which might not be representative of all latency characteristics of real-world production workloads with geographically distributed end-points; there is no actual adversarial quantum red-team testing; and the encryption layer does not account for quantum resistance. Future studies will focus on post-quantum key encapsulation mechanisms (CRYSTALS-Kyber), federated learning solutions for privacy-preserving training of a trust model across different cloud tenants, and on integrating the ZTDPA compliance engine with the new EU obligations under the AI Act for high-risk autonomous systems.

References

1. Ahmed, S., & Patel, R. (2021). Federated identity management for multi-cloud zero-trust deployments. *IEEE Transactions on Cloud Computing*, *9*(4), 1423–1436. https://doi.org/10.1109/TCC.2021.3089123
2. Alotaibi, B., & Roussinov, D. (2019). Continuous authentication in cloud-native environments using behavioural biometrics. *Computers & Security*, *87*, 101582. https://doi.org/10.1016/j.cose.2019.101582
3. Anderson, K., White, L., & Torres, M. (2020). Threat modelling for hybrid multi-cloud workloads: A taxonomic approach. *Journal of Network and Computer Applications*, *158*, 102590. https://doi.org/10.1016/j.jnca.2020.102590
4. Baker, D., & Nguyen, T. (2021). Runtime policy enforcement for serverless functions using eBPF and zero-trust principles. *ACM Transactions on Privacy and Security*, *24*(3), 1–29. https://doi.org/10.1145/3441645
5. Chen, H., & Li, X. (2019). Continuous session re-authentication for microservice architectures in cloud environments. *Future Generation Computer Systems*, *97*, 193–204. https://doi.org/10.1016/j.future.2019.02.030
6. Davis, J., & Kumar, A. (2021). Zero-trust network access for OT/IT convergence in industrial cloud systems. *IEEE Internet of Things Journal*, *8*(12), 9834–9847. https://doi.org/10.1109/JIOT.2021.3065432
7. Diaz, C., & Martin, F. (2020). SPIFFE and SPIRE: Workload identity for cloud-native zero-trust architectures. *Proceedings of the 15th ACM European Conference on Computer Systems*, 1–14. https://doi.org/10.1145/3342195.3387531
8. Fernandez, M., Lopez, A., & Garcia, R. (2021). Symmetric encryption benchmarking across heterogeneous cloud providers: AES-GCM vs. ChaCha20-Poly1305. *IEEE Access*, *9*, 48312–48325. https://doi.org/10.1109/ACCESS.2021.3068901

9. Garg, S., & Dhingra, P. (2020). Automated GDPR compliance monitoring in cloud deployments using policy-as-code frameworks. *International Journal of Information Management*, *53*, 102120. https://doi.org/10.1016/j.ijinfomgt.2020.102120

10. Hassan, M., & Ahmed, Z. (2020). Continuous GDPR compliance pipeline for cloud-hosted data processing systems. *Computers & Security*, *96*, 101879. https://doi.org/10.1016/j.cose.2020.101879

11. Huang, Y., & Zhu, W. (2021). Graph-based lateral movement detection in zero-trust cloud environments. *IEEE Transactions on Information Forensics and Security*, *16*, 3142–3155. https://doi.org/10.1109/TIFS.2021.3059742

12. Islam, M., Khan, S., & Roy, P. (2020). Unsupervised anomaly detection for zero-day cloud threats using DBSCAN clustering. *Expert Systems with Applications*, *156*, 113481. https://doi.org/10.1016/j.eswa.2020.113481

13. Johnson, B., & Miller, C. (2021). Mapping zero-trust architecture tenets to NIST cybersecurity framework subcategories: An alignment study. *NIST Journal of Research and Technology*, *126*, 1–18. https://doi.org/10.6028/jres.126.001

14. Kindervag, J. (2021). Zero trust in hybrid cloud: Operationalising the seven tenets across provider boundaries. *IEEE Security & Privacy*, *19*(2), 34–42. https://doi.org/10.1109/MSEC.2020.3040512

15. Kumar, V., Sharma, A., & Reddy, N. (2021). Short-lived credential management for ephemeral serverless compute in zero-trust environments. *Future Generation Computer Systems*, *116*, 278–291. https://doi.org/10.1016/j.future.2020.10.040

16. Li, F., & Huang, Q. (2019). Micro-segmentation policy synthesis for container-based cloud workloads. *IEEE Transactions on Network and Service Management*, *16*(3), 1089–1102. https://doi.org/10.1109/TNSM.2019.2914432

17. Liu, Z., Chen, Y., & Wang, B. (2022). Cross-cluster micro-segmentation policy synchronisation in federated Kubernetes environments. *IEEE Transactions on Services Computing*, *15*(1), 214–228. https://doi.org/10.1109/TSC.2022.3149012

18. Martinez, J., & Perez, S. (2020). Dynamic risk scoring for adaptive access control in hybrid enterprise cloud environments. *Computers & Security*, *94*, 101817. https://doi.org/10.1016/j.cose.2020.101817

19. Mohan, V., & Subramanian, L. (2020). Identity-tagged east-west micro-segmentation in Kubernetes service meshes using Istio and Cilium. *ACM Symposium on Information, Computer and Communications Security*, 312–326. https://doi.org/10.1145/3320269.3384727

20. Nguyen, P., Tran, H., & Do, T. (2020). Distributed key management for multi-cloud deployments using hardware security modules and secret sharing. *Journal of Systems and Software*, *166*, 110596. https://doi.org/10.1016/j.jss.2020.110596

21. Oliveira, R., & Santos, G. (2021). Legal-technical analysis of GDPR Articles 22 and 32 constraints on autonomous AI systems in cloud deployments. *Computer Law & Security Review*, *40*, 105520. https://doi.org/10.1016/j.clsr.2021.105520

22. Park, J., & Lee, K. (2019). Markov chain modelling of machine-to-machine traffic for anomaly detection in cloud orchestration layers. *IEEE Communications Letters*, *23*(11), 1982–1985. https://doi.org/10.1109/LCOMM.2019.2939421

23. Patel, N., & Rao, S. (2020). Attribute-based access control as a unified policy layer for heterogeneous cloud identity providers. *IEEE Transactions on Dependable and Secure Computing*, *17*(6), 1278–1292. https://doi.org/10.1109/TDSC.2018.2886920

24. Rashid, A., & Chivers, H. (2021). Confidential computing enclaves for zero-trust ML inference in regulated cloud environments. *IEEE Transactions on Information Forensics and Security*, *16*, 4823–4837. https://doi.org/10.1109/TIFS.2021.3109832

25. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

26. Roy, A., & Das, B. (2021). Supply-chain integrity verification for container images in zero-trust cloud pipelines using Sigstore and SLSA. *Proceedings of the 2021 IEEE CloudCom*, 87–95. https://doi.org/10.1109/CloudCom52223.2021.00024

27. Shaikh, F., & Boreiko, A. (2020). Policy-as-code for automated infrastructure security governance in DevSecOps pipelines. *IEEE Access*, *8*, 153012–153027. https://doi.org/10.1109/ACCESS.2020.3018312

28. Singh, H., Rana, O., & Parashar, M. (2021). GDPR Article 44 transfer latency in multi-cloud data residency enforcement. *ACM Transactions on Internet Technology*, *21*(3), 1–24. https://doi.org/10.1145/3432201

29. Sun, L., & Li, H. (2019). Software-defined perimeter as a zero-trust enforcement plane for IoT-cloud integration. *Future Generation Computer Systems*, *90*, 319–330. https://doi.org/10.1016/j.future.2018.08.009

30. Tan, Q., & Xu, J. (2021). Adaptive session management under resource contention in zero-trust serverless cloud architectures. *IEEE Transactions on Parallel and Distributed Systems*, *32*(9), 2211–2224. https://doi.org/10.1109/TPDS.2021.3068982

31. Thompson, D., Brooks, E., & Yuen, A. (2021). Ensemble anomaly detection for cloud workload security: A systematic meta-analysis of 48 studies. *ACM Computing Surveys*, *54*(5), 1–38. https://doi.org/10.1145/3444689

32. Venkatesan, S., & Raj, P. (2020). Blockchain-anchored audit trails for immutable compliance logging in multi-cloud zero-trust systems. *IEEE Transactions on Cloud Computing*, *8*(3), 902–915. https://doi.org/10.1109/TCC.2020.2967341

33. Wang, C., & Gupta, M. (2019). Cloud access security broker architecture for multi-cloud IAM normalisation. *IEEE Transactions on Services Computing*, *12*(4), 561–574. https://doi.org/10.1109/TSC.2017.2719140

34. Wilson, T., & Clark, P. (2022). Just-in-time privilege elevation for autonomous cloud agents using ephemeral role binding. *Computers & Security*, *113*, 102538. https://doi.org/10.1016/j.cose.2021.102538

35. Wu, Z., & Fan, X. (2021). Threat intelligence sharing for zero-trust cloud ecosystems using STIX 2.1 and TAXII 2.1 protocols. *Journal of Information Security and Applications*, *57*, 102724. https://doi.org/10.1016/j.jisa.2020.102724

36. Xiao, L., & Liu, Y. (2022). ML-driven adaptive firewall rule generation for micro-segmented multi-cloud networks. *IEEE Transactions on Network and Service Management*, *19*(1), 312–325. https://doi.org/10.1109/TNSM.2022.3140323

37. Yang, R., & Zhou, F. (2019). Homomorphic encryption feasibility for privacy-preserving cloud analytics in regulated industries. *IEEE Transactions on Cloud Computing*, *7*(2), 543–557. https://doi.org/10.1109/TCC.2017.2730298

38. Zhang, Q., Li, M., & Chen, L. (2020). Threat topology and attack pattern taxonomy for heterogeneous multi-cloud meshes. *IEEE Transactions on Information Forensics and Security*, *15*, 2901–2914. https://doi.org/10.1109/TIFS.2020.2977512

39. Zhao, Y., & Liang, H. (2021). Zero-trust access control for edge-cloud continuum environments in Industry 4.0. *IEEE Internet of Things Journal*, *8*(23), 16895–16907. https://doi.org/10.1109/JIOT.2021.3076154

40. Zhou, X., & Huang, T. (2022). Post-quantum cryptography migration strategies for enterprise multi-cloud zero-trust architectures. *IEEE Security & Privacy*, *20*(1), 56–66. https://doi.org/10.1109/MSEC.2021.3113400