IoT

MDPI

Systematic Review

Performance Trade-Offs in Multi-Tenant IoT–Cloud Security: A Systematic Review of Emerging Technologies

Bader Alobaywi [1,2,*], Mohammed G. Almutairi [1,2] and Frederick T. Sheldon [1,*]

[1] Department of Computer Science, College of Engineering, University of Idaho, Moscow, ID 83844, USA; almu9701@vandals.uidaho.edu
[2] Applied College, University of Hafr Al Batin, Hafar Al Batin 39923, Saudi Arabia
* Correspondence: alob8017@vandals.uidaho.edu (B.A.); sheldon@uidaho.edu (F.T.S.)

Abstract

Multi-tenancy is essential for scalable IoT–Cloud systems; however, it introduces complex security vulnerabilities at the intersection of shared cloud infrastructures and resource-constrained IoT environments. This systematic review evaluates next-generation security frameworks designed to enforce tenant isolation without violating the strict latency (<10 ms) and energy bounds of lightweight sensors. Adhering to PRISMA guidelines, we analyze selected high-quality studies to categorize intersectional threats, including cross-tenant data leakage, side-channel attacks, and privilege escalation. Our analysis identifies a critical, unresolved conflict: existing mitigation strategies often incur a 12% computational and communication overhead, creating a significant barrier for real-time applications. Furthermore, we critically analyze emerging technologies, including Zero Trust Architectures (ZTA), adaptive Artificial Intelligence (AI), blockchain, and Post-Quantum Cryptography (PQC). We find that direct PQC deployment is currently infeasible for LPWAN protocols due to key-size constraints (1.6 KB) that exceed typical payload limits. To address these challenges, we propose a novel multi-layer security design principle that offloads heavy isolation and cryptographic workloads to hardware-accelerated edge gateways, thereby maintaining tenant isolation without compromising real-time performance. Finally, this review serves as a roadmap for future research, highlighting federated learning and hardware enclaves as essential pathways for securing next-generation multi-tenant IoT ecosystems.

Keywords: IoT–cloud security; multi-tenancy; Post-Quantum Cryptography (PQC); Zero Trust Architecture (ZTA); performance-security trade-off; IoT–cloud isolation

Check for updates

Academic Editor: Amiya Nayak

Received: 16 January 2026
Revised: 18 February 2026
Accepted: 20 February 2026
Published: 22 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The Internet of Things (IoT) and cloud integration, also referred to as the Cloud-of-Things (CoT) [1], has become an increasingly central component. This architecture allows billions of resource-constrained devices to leverage the elastic computation and storage of the cloud to support applications in smart cities and industrial automation. At its core lies multi-tenancy, an architectural model that enables multiple independent users or tenants to share the same physical resources. Multi-tenancy is a critical driver for large-scale IoT due to its cost-effectiveness and resource efficiency. The financial momentum behind this model is clear, as the global market for multi-tenant data centers is expected to more than triple to USD 189.59 billion by 2034. This rapid expansion proves that multi-tenancy is no longer optional but a foundational element of the digital age [2].

While multi-tenancy offers clear benefits, the shared environment introduces unique security and privacy vulnerabilities absent in single-tenant systems. Previous research emphasizes significant flaws in data partitioning, resource isolation, and access control within IoT–cloud ecosystems [3,4]. These vulnerabilities are worsened by the inherent constraints of IoT hardware, specifically limited memory, low processing power, and the need for real-time responsiveness. This constraint hinders the implementation of resource-intensive security protocols.

Although general IoT and cloud security have been studied extensively [5], the literature lacks a focus on their intersection. Most existing reviews treat IoT and cloud computing as isolated sectors, failing to account for how their integration reshapes the threat landscape. The interaction between resource-constrained IoT hardware and the complex multi-tenant structures of the cloud introduces unique attack vectors, such as cross-VM exploits, side-channel leakage, and inter-tenant privilege escalation, that remain underexplored. Addressing this lack of specialized analysis is vital for advancing both theoretical research and practical industry standards.

This review examines the security challenges associated with multi-tenancy in integrated IoT–cloud environments. It presents a structured threat of taxonomy covering cross-tenant data leakage, misconfigured identity boundaries, side-channel exploitation, and noisy-neighbor resource interference. In particular, the review highlights a central unresolved challenge in the literature: achieving strong tenant isolation without violating the strict performance, latency, and energy constraints imposed by IoT systems. This trade-off remains an open problem, especially in large-scale, diverse deployments.

IoT systems consist of highly heterogeneous and resource-constrained devices that continuously generate large volumes of data. However, they show significant constraints in the processing of power, storage, and energy capacity required for advanced analytics, AI inference, and long-term data management. Cloud computing addresses these limitations by offering scalable computation, virtually unlimited storage, and flexible resources [6]. By offloading workloads to the cloud, IoT systems enable real-time analytics, predictive insights, and efficient decision-making that are impractical at the device level. Additionally, cloud platforms provide global accessibility, standardized interfaces, and service orchestration. That ensures interoperability across diverse IoT ecosystems [7]. Although this integration facilitates centralized monitoring, diagnostics, and optimization across various domains, it presents a structural security conflict between tenant isolation and IoT performance constraints that current models do not sufficiently resolve.

A central and unresolved challenge in IoT–Cloud research is the Security-Performance Trade-off. While traditional cloud-based security mechanisms offer robust protection, they are often too heavy for the IoT edge. Empirical data indicates that standard encryption and isolation protocols can additionally impose a 12% resource, which can lead to system instability in critical latency environments that require responses in under 10 ms. This review specifically explores how emerging technologies attempt to bridge this protection-overhead gap.

To mitigate this fragmentation, the review evaluates existing mitigation approaches that include advanced access control mechanisms, strengthened virtualization-level protections, and AI-driven anomaly detection techniques. We critically evaluate their practical suitability for IoT environments, moving beyond simple descriptions to analyze the friction between security overhead and device constraints. Additionally, we highlight emerging trends that represent promising future movements, including Zero Trust Architecture (ZTA), adaptive AI-driven analytics, blockchain-based auditing, and the integration of lightweight Post-Quantum Cryptography (PQC).

https://doi.org/10.3390/iot7010021

This review is structured around the following research questions, which define the scope and guide the analysis:

1. What are the dominant multi-tenancy security threats at the intersection of IoT and shared cloud layers?
2. How do next-generation mitigations (ZTA, AI, Blockchain, PQC) perform when measured against strict IoT resource constraints?
3. What architectural shifts, such as edge offloading or hardware isolation, are required to enable future-proof, quantum-resilient multi-tenant deployments?

Contributions

While existing studies have separately addressed IoT security or cloud virtualization risks, a significant need remains to analyze the specific conflict points where these two domains intersect. This paper distinguishes itself from prior work through the following contributions.

- A Tenant-Centric Threat Taxonomy: We classify threats based on their specific impact on tenant isolation rather than on generic attack vectors such as cross-tenant data leakage or noisy neighbor resource exhaustion in shared environments.
- Intersection Analysis: Unlike general surveys, we evaluate mitigation strategies specifically against IoT constraints (latency, energy, and limited processing power), identifying why standard cloud defenses often fail at the IoT edge.
- Architectural Synthesis: We propose an integrated architectural approach that combines Zero Trust for logical isolation, AI for dynamic threat detection, and PQC for long-term data resilience, mapping these solutions to specific layers of the IoT–Cloud stack (e.g., Edge Gateways vs. End-Devices).

The rest of this paper is organized as follows: Section 2 details the systematic review methodology and PRISMA selection process. Section 3 provides the foundational architecture of IoT–Cloud systems and defines multi-tenancy models. Section 4 presents a critical analysis of multi-tenancy security risks and a synthesis of primary studies. Section 5 evaluates emerging technologies and their performance in terms of trade-offs. Also, it discusses feasibility and implementation challenges, followed by a future research roadmap in Section 6. Finally, Section 7 concludes the paper.

2. Methodology

To ensure a comprehensive examination of security challenges in multi-tenant IoT cloud environments, this study employs a systematic review methodology. This systematic review was conducted in accordance with the PRISMA 2020 guidelines to ensure scientific rigor and transparency (see Supplementary File S1). As this is a literature-based review of existing peer-reviewed studies, protocol registration was not required. By adhering to these systematic principles, we ensured that the selection of the 13 primary studies was objective and directly emerged from our predefined search strategy and eligibility criteria.

2.1. Search Strategy and Data Sources

We conducted a targeted search across three primary academic databases: IEEE Xplore, ACM Digital Library, and MDPI. These sources were selected for their high concentration of peer-reviewed technical literature in computer science and engineering. To capture the specific intersection of domains, we employed the following Boolean search strings: (“Multi-tenancy” OR “Tenant Isolation” OR “Shared Cloud”) AND (“IoT” OR “Internet of Things” OR “Cloud-of-Things” OR “Edge” OR “Fog”) AND (“Security” OR “Privacy” OR “Access Control” OR “Attack” OR “Performance”).

2.2. Inclusion and Exclusion Criteria

All records were imported into EndNote and screened using Rayyan 1.12. After automated deduplication and preliminary eligibility filtering, 74 records underwent independent title and abstract screening by two reviewers (B.A. and M.A.). Full-text assessment was subsequently performed for 41 articles, resulting in the inclusion of 13 primary studies that met all eligibility criteria (Figure 1).

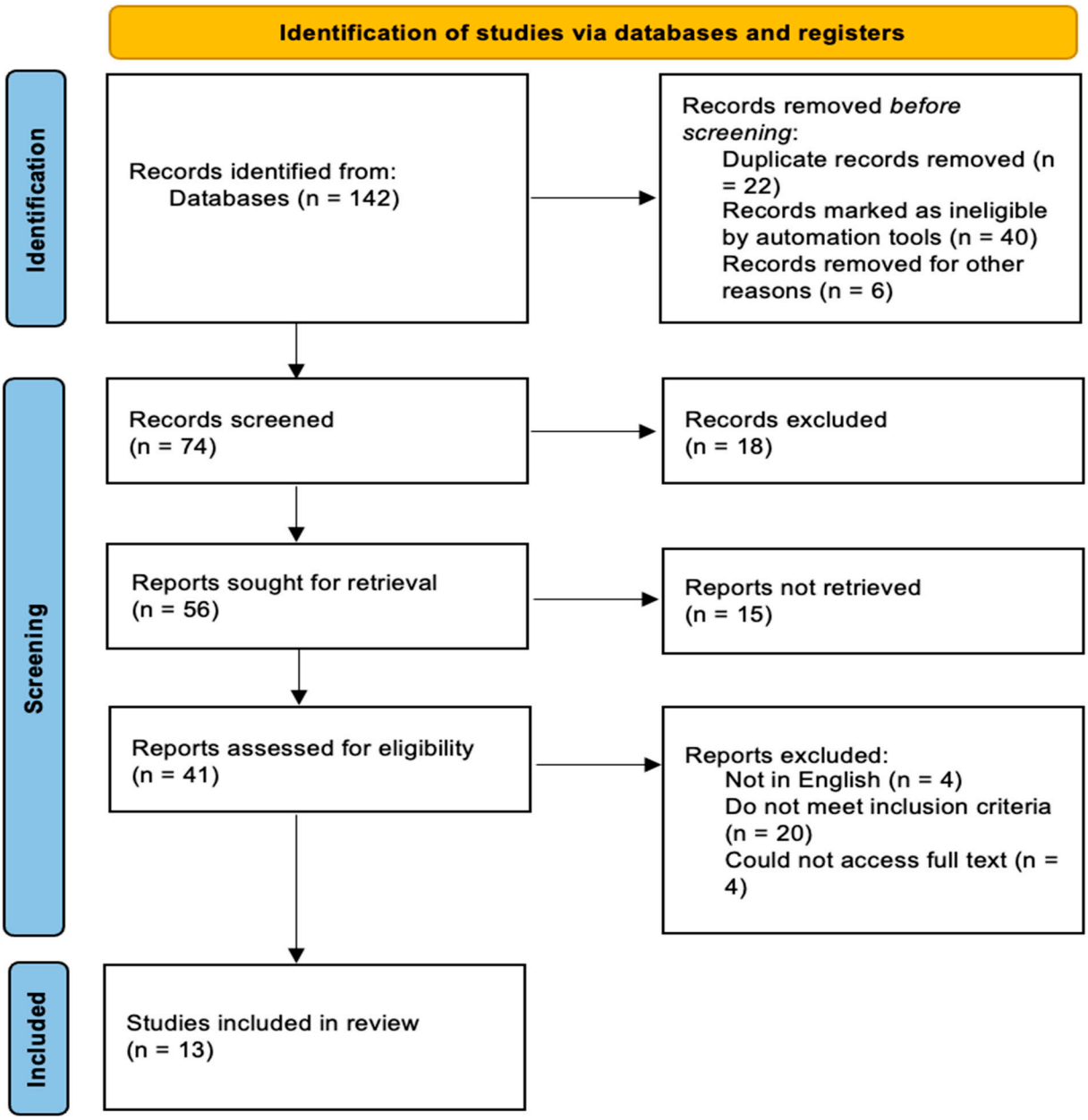


Figure 1. PRISMA 2020 flow diagram for the systematic review, including searches of databases and selection criteria.

- Inclusion: Peer-reviewed journal articles and conference proceedings published between 2020 and 2025 that explicitly address multi-tenancy in IoT–cloud environments and propose concrete security mitigation mechanisms.
- Exclusion: included generic cloud studies without an IoT component, non-peer-reviewed or non-technical works, non-English publications, and studies predating modern containerization paradigms.

2.3. Scope and Limitations of the Review

This review intentionally focuses on security challenges introduced by multi-tenancy in IoT–cloud integration. This excludes single-tenant IoT systems and generic cloud security studies by design. Repeated keyword refinement was applied to ensure coverage; however, broader terms (e.g., resource sharing) consistently returned cloud-centric studies that lacked IoT performance constraints.

To further substantiate the robustness of the selected dataset (n = 13), a bibliographic dispersion analysis was conducted. The studies originate from multiple independent research environments, including 5 from the USA, 3 from India, 1 from Canada, 1 from Japan, 1 from Bahrain, 1 from Malaysia/Iraq, and 1 from Romania. No recurring authorship patterns were identified, and no institutional affiliations were shared across the selected works. This confirms that the dataset does not reflect repeated contributions from a single research lineage but rather independent investigations. Despite this dispersion, thematic convergence is evident. For example, a performance overhead of approximately 12% was reported in two of the analyzed studies, indicating that the observed bottlenecks reflect systemic architectural constraints rather than isolated experimental artifacts (see Section 4.1).

Accordingly, the analysis emphasizes next-generation adaptations of established frameworks, such as ZTA, AI-assisted detection, blockchain-based trust models, and PQC, which are applicable under strict IoT resource constraints and represent next-generation mitigation strategies for IoT–cloud environments. Due to various methodologies across studies, results are synthesized qualitatively and comparatively rather than through statistical analysis.

3. Foundational Architecture and Multi-Tenancy Models

This section provides the necessary technical foundation and architectural definitions required to categorize the threats identified in the literature search.

3.1. IoT–Cloud Systems Architecture

The IoT and cloud integration form a layered, service-oriented architecture designed to handle large-scale data streams, enable real-time analytics, and support scalable applications. This architecture capitalizes on the broad connectivity of IoT devices combined with the cloud's processing and storage capacity to enable scalable, intelligent, and adaptive services. The IoT–cloud stack is organized into five tightly connected layers, as shown in Figure 2 [8]. Each layer has a clearly defined role, working together to support reliable communication, efficient resource management, and end-to-end service provisioning.

3.1.1. Perception Layer

The perception layer supports the IoT–cloud architecture and includes heterogeneous smart devices, such as sensors, actuators, RFID tags, cameras, and wearables, which directly interact with the physical world. These components capture environmental parameters (e.g., temperature, humidity, motion, and location) and translate them into digital data streams. Due to their constrained computer and storage capabilities, most IoT devices rely on higher layers for data processing and storage. From a risk standpoint, this layer is a primary attack surface, as devices are frequently exposed and vulnerable to physical compromise, spoofing, and malicious data injection [8].

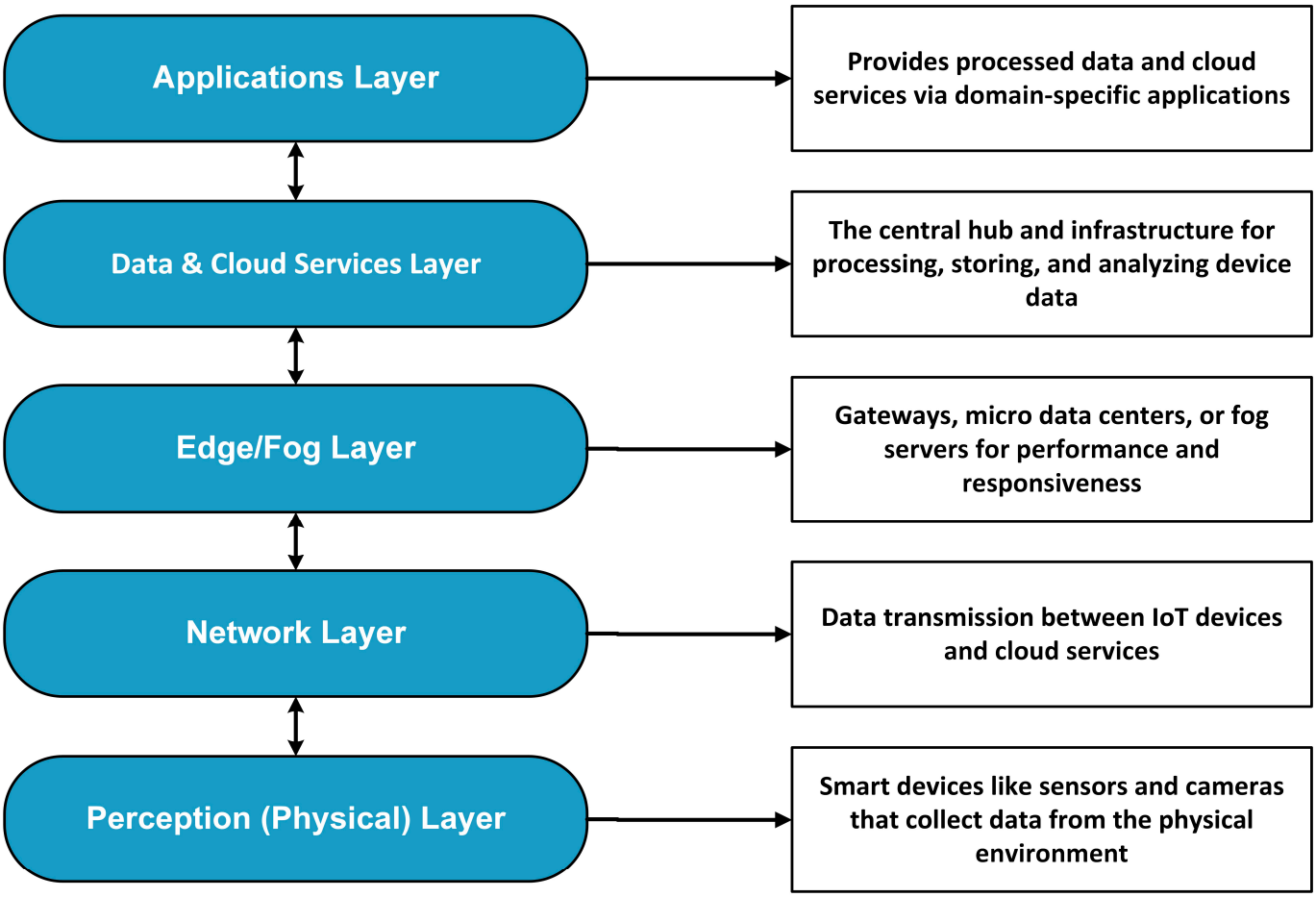


Figure 2. IoT–Cloud Architecture.

3.1.2. Network Layer

The network layer is responsible for dependable and efficient data transport between IoT endpoints and cloud services. It leverages a mix of communication technologies (e.g., Wi-Fi, Bluetooth, Zigbee, 4G/5G, LoRaWAN, and NB-IoT) to support various connectivity and performance requirements [9]. This layer acts as the transport backbone, carrying large data volumes produced at the perception layer to downstream processing components. To meet performance and trust requirements, it relies on network management protocols, Quality of Service (QoS) controls, and encryption mechanisms to preserve data availability, reliability, and confidentiality. From a security perspective, the network layer is a valuable target. It handles sensitive traffic and is exposed to threats such as man-in-the-middle attacks, where communications are intercepted or modified. Also, it is disclosed to denial-of-service (DoS) attacks that disrupt data delivery. These risks are expanded by the scale and variety of IoT deployments, which significantly expand the network's attack surface [10].

3.1.3. Edge/Fog Layer

Placed between the network and cloud layers, the edge/fog layer enhances system performance and responsiveness by introducing intermediate compute resources such as gateways, micro data centers, and fog servers. These nodes handle local data aggregation, filtering, and preliminary analytics before forwarding only relevant information to the cloud. Executing computations close to the data source can reduce latency, optimize bandwidth utilization, and support real-time decision-making. Those capabilities are critical for sensitive latency use cases such as autonomous driving and industrial automation [11]. In addition, fog computing supports privacy preservation by keeping sensitive data at or

near the source when feasible. This reduces unnecessary exposure to centralized cloud environments [12].

3.1.4. Data and Cloud Services Layer

The cloud layer functions as the central hub for IoT ecosystems, utilizing virtualization techniques to deliver elastic, multi-tenant resources for data processing and storage. Compute resources within cloud systems encompass virtualized servers, containers, and serverless architectures; however, containerization requires specialized runtimes to mitigate isolation risks [13–16]. These resources are supported by scalable storage solutions, categorized into object, block, and file storage [17], and by robust networking components such as Virtual Private Clouds (VPCs), load balancers, and Content Delivery Networks (CDNs) that ensure logical isolation and low latency [18,19]. Furthermore, management and monitoring tools are essential for optimizing resource provisioning and operational health throughout the system's lifecycle [20].

Security within this layer is paramount and is established through a combination of contractual Service Level Agreements (SLAs) and technical controls, including encryption, strict authentication, and regular audits [21,22]. To address the specific risks of multi-tenancy, advanced strategies are increasingly utilized, such as data segmentation combined with integrity checks and the two-tier WAY framework, which employs virtual machine monitoring to assess user trust and applies security policies across network, infrastructure, and data levels [23,24].

Cloud delivery relies on three service models: Infrastructure as a Service (IaaS) for virtualized resources, Platform as a Service (PaaS) for development environments, and Software as a Service (SaaS) for end-user applications [25,26]. These are implemented through various deployment strategies, including public clouds for cost-effective scalability, Private clouds for strict control, and Community clouds for shared regulatory needs. Hybrid clouds are particularly prevalent in mature IoT deployments, balancing the scalability of public infrastructure with the data sensitivity requirements of private environments [27].

3.1.5. Application Layer

This layer delivers data processing and cloud services to the end users through specific domains, including smart cities, healthcare, transportation, industrial control, agriculture, and energy management. Also, it enables streamlined interaction between IoT devices and cloud platforms using application-layer protocols such as MQTT, CoAP, XMPP, and AMQP [27,28]. The classification of protocols within the application layer remains an ongoing subject of discussion in the field of IoT research. For instance, while MQTT is widely categorized as an application-layer protocol due to its data-handling capabilities, some interpretations consider it a lower-level messaging or session protocol because it operates over TCP/IP to provide data transport services. In multi-tenant environments, this distinction is of critical importance: MQTT functions as a connectivity intermediary that must enforce isolation through tenant-specific namespaces and broker-level Access Control Lists (ACLs). This review treats MQTT, CoAP, and AMQP as application-layer components within the IoT–Cloud stack, acknowledging their dual role in both high-level data orchestration and low-level message routing. Also, in this layer, applications rely on APIs and service interfaces to interact with the infrastructure. These APIs expose core capabilities, including data visualization, decision support, and automated control functions [29].

At the application layer, security is a critical priority because it directly interacts with end users and often handles sensitive information. Delivering a trusted service at this level depends on robust access controls, strong user authentication, and reliable mechanisms

for verifying data integrity [30]. At this phase, Main security concerns focus on user data privacy and application security. These include access control breaches, malicious code injection, and network sniffing attacks. Such threats allow adversaries to bypass authorization controls, inject malicious code, and intercept network traffic. As a result, IoT data confidentiality and system integrity are compromised [31].

Overall, the IoT–Cloud architecture transforms raw sensor data into actionable insights by tightly coupling physical devices with cloud services. However, this intersection introduces major challenges, including latency, data privacy, and interoperability. Also, the most critical issue in multi-tenant environments is security, which remains a central focus of IoT–cloud research.

The layered IoT–cloud architecture presented in this work is not a direct implementation of the ISO/OSI reference model, but a functional abstraction tailored to multi-tenant IoT systems. The perception and network layers broadly map to the lower OSI layers (physical through transport), handling data collection and communication. Higher layers differ from the OSI model by incorporating edge and cloud computing capabilities that span the session and application layers, while addressing resource management and isolation that fall outside the OSI scope. The application layer aligns with the OSI application layer, which relies on APIs and service interfaces to interact with the infrastructure. That makes this abstraction more suitable for analyzing IoT–cloud security challenges.

3.2. Multi-Tenant Cloud Systems

Multi-tenancy is a fundamental principle of cloud computing, representing a single software model and shared infrastructure to provision multiple independent tenants. As an architectural paradigm, multi-tenancy supports the as-a-service model (SaaS, PaaS, IaaS), enabling resource efficiency, scalability, and reduced operational costs by sharing compute, storage, and networking resources [32]. In IoT–Cloud environments, enforcing multi-tenancy is more complex, as isolation must transit the entire data lifecycle, including device connectivity, messaging, processing, and storage. However, shared resources introduce additional attack surfaces, such as co-residency, shared hypervisors, API management, and multi-tenant storage, which may lead to cross-tenant interference or data leakage if isolation fails. These risks are well documented in cloud security threat models [33]. In IoT–cloud systems, multitenancy is typically implemented using one of three models. In the single database model, each tenant has its own dedicated application and database instance. The shared database with isolated schemas model uses a single database while providing each tenant with a separate schema to maintain logical data separation. In the shared database with a shared schema model, all tenants operate within the same database schema. This relies on additional mechanisms to determine and isolate specific tenant data [34].

3.2.1. Tenant Isolation Mechanisms

The intersection of IoT and cloud computing grows multi-tenancy security challenges by combining shared cloud infrastructure with large-scale, heterogeneous IoT devices. In IoT–Cloud architectures, many devices access shared resources for storage, analytics, and orchestration, significantly expanding the attack surface and complicating tenant isolation. The core objective is to enforce strict logical separation across the platform stack, ensuring that each tenant's data, devices, and configurations remain isolated from those of others. Achieving effective multi-tenancy in IoT–Cloud environments requires isolation mechanisms across multiple architectural layers. This ensures confidentiality, integrity, and availability at physical, logical, and application levels.

Device Isolation

At this layer, connectivity and device isolation ensure that each IoT device is securely connected to and authenticated with its tenant. To prevent cross-tenant data exposure and interference, communication protocols such as MQTT, AMQP, and CoAP must be carefully configured. This is commonly accomplished by using specific tenant namespaces and ACLs at the message broker level, which block unauthorized tenants from publishing or subscribing to other tenants' data streams. Furthermore, secure device connectivity, unique credentials, and mutual authentication are essential for ensuring device-level isolation. While security remains a critical priority at the application layer, the enforcement of tenant isolation at the lower architectural levels is equally essential to prevent unauthorized access at the network entry point. However, implementing robust security at the perception layer is complicated by the inherent constraints of IoT hardware, which often lacks the processing power and energy required for standard encryption protocols. For instance, traditional Public Key Infrastructure (PKI) based on RSA-2048 involves complex mathematical operations and large certificate exchanges that exceed the memory and power bounds of lightweight sensors. Consequently, multi-tenant IoT systems must shift toward lightweight cryptographic primitives and elliptic curve-based PKI (ECC), which provide equivalent security with significantly smaller key sizes and lower computational overhead, thereby protecting the message broker level without compromising device longevity.

Network Isolation

It is a standard approach for separating tenant traffic within shared IoT–cloud infrastructures. Its objective is to ensure that data flows associated with one tenant remain logically distinct from those of others. Network isolation strategies are generally categorized into physical and logical isolation models [35].

Physical isolation assigns dedicated network hardware or infrastructure to individual tenants, thereby eliminating resource sharing at the network level. While this approach provides strong separation, it is typically limited to environments with strict regulatory or compliance requirements due to cost and scalability constraints.

Logical isolation is more commonly adopted in large-scale IoT–cloud systems [36]. It relies on software mechanisms to divide shared physical networks. Common approaches include Virtual LANs (VLANs), which segment network traffic by assigning tenants to isolated logical domains, and Virtual Private Networks (VPNs), which establish encrypted tenant-specific tunnels over shared communication channels. Software-Defined Networking (SDN) further enables programmable and dynamic traffic separation at the tenant level. At the same time, micro segmentation and policy-based controls use firewalls, ACLs, and policy-driven routing to enforce fine-grained network isolation [35]. While logical isolation often focuses on wired backends, contemporary radio protocols, such as LoRaWAN and NB-IoT, introduce unique multi-tenancy challenges due to their limited payload sizes (e.g., 51–222 bytes for LoRaWAN) and shared radio spectrum. In a multi-tenant environment, the shared nature of the radio medium makes these protocols vulnerable to inter-tenant interference and jamming. Furthermore, the strict bandwidth limits of Lo-RaWAN make the transmission of large security artifacts, such as PQC public keys (typically >1.6 KB), a significant bottleneck that requires specialized fragmentation or offloading strategies.

Resource Isolation

In multi-tenant architecture, this isolation ensures that each tenant's computer, storage, and networking resources function as self-contained units. That means each tenant is logically and operationally separated from the other tenants. The strategic goal is to eliminate cross-tenant interference that could degrade performance, compromise security,

or disrupt service availability, without damaging overall system efficiency or scalability. Isolation mechanisms are implemented across multiple tiers, spanning physical hardware, virtualization layers, container-based platforms, and application-level controls. Commonly utilized technologies include virtual machines (VMs) and container frameworks such as Docker and Kubernetes. These divide tenant workloads to contain faults and security incidents, thereby preventing their propagation across shared infrastructure.

To support fair resource allocation in multi-tenant platforms, orchestration frameworks typically enforce per-tenant quotas and limits on CPU, memory, storage, and network bandwidth. These mechanisms help maintain predictable performance and service availability while enabling efficient utilization of shared infrastructure, as demonstrated in prior studies on resource isolation in containerized multi-tenant PaaS architectures [37]

Data Isolation

It ensures that the large volumes of time-series telemetry, metadata, and analytics are stored within shared databases. Ensuring logical separation of tenant data in these repositories is critical to prevent data leakage or inference attacks. This can be achieved through different database isolation models, such as dedicated databases per tenant, separate schemas for each tenant, or a shared schema. Choosing the appropriate model typically depends on factors like scalability, performance, and regulatory compliance [38].

Application Isolation

This software-level isolation ensures that each tenant's user interfaces, business logic, and rule engine configurations are securely separated. This is often achieved by running tenants in separate runtime environments and enforcing tenant-aware Role-Based Access Control (RBAC), which restricts users and APIs to their own resources. This isolation is essential in IoT–Cloud environments, where shared analytics engines and dashboards could lead to accidental or malicious cross-tenant data exposure [39].

4. Critical Analysis of Multi-Tenancy Security

The IoT–cloud intersection has produced several studies on data privacy, virtualization security, and cross-domain trust management. However, research on multi-tenancy, a core aspect of modern cloud infrastructures in IoT–cloud contexts, remains fragmented. While progress has been made in related areas, few studies treat multi-tenancy as a comprehensive security challenge. This section synthesizes existing findings to highlight unresolved gaps in isolation, performance, and architectural design.

The architectures of multi-tenancy inherently expand the attack surface, as multiple tenants share virtualized environments. The threat model in Figure 3 illustrates how these shared resources make tenants vulnerable to cross-tenant attacks. Research indicates that such infrastructures double risks to data confidentiality, access control, and resource isolation. Studies [3,4] highlight that shared physical resources in multi-tenant deployments can allow malicious tenants to access or compromise sensitive data. Moreover, the distributed and heterogeneous nature of IoT systems makes maintaining effective tenant isolation a persistent challenge.

4.1. Multi-Tenancy Security Risks Classification

Figure 3 illustrates the expansion of risks; recent studies categorize multi-tenancy threats into several key areas. Among these, cross-tenant data leakage is particularly critical in shared IoT–cloud platforms [40], often arising from weak isolation of virtualized resources, insecure APIs, or misconfigured access controls. Hashim et al. found that flawed permission models and poorly configured authorization mechanisms frequently lead to tenant-level data exposure [3]. Kumar et al. showed that inadequate separation between

virtual machines or containers can facilitate cross-tenant inference attacks and unauthorized side-channel activity [39].

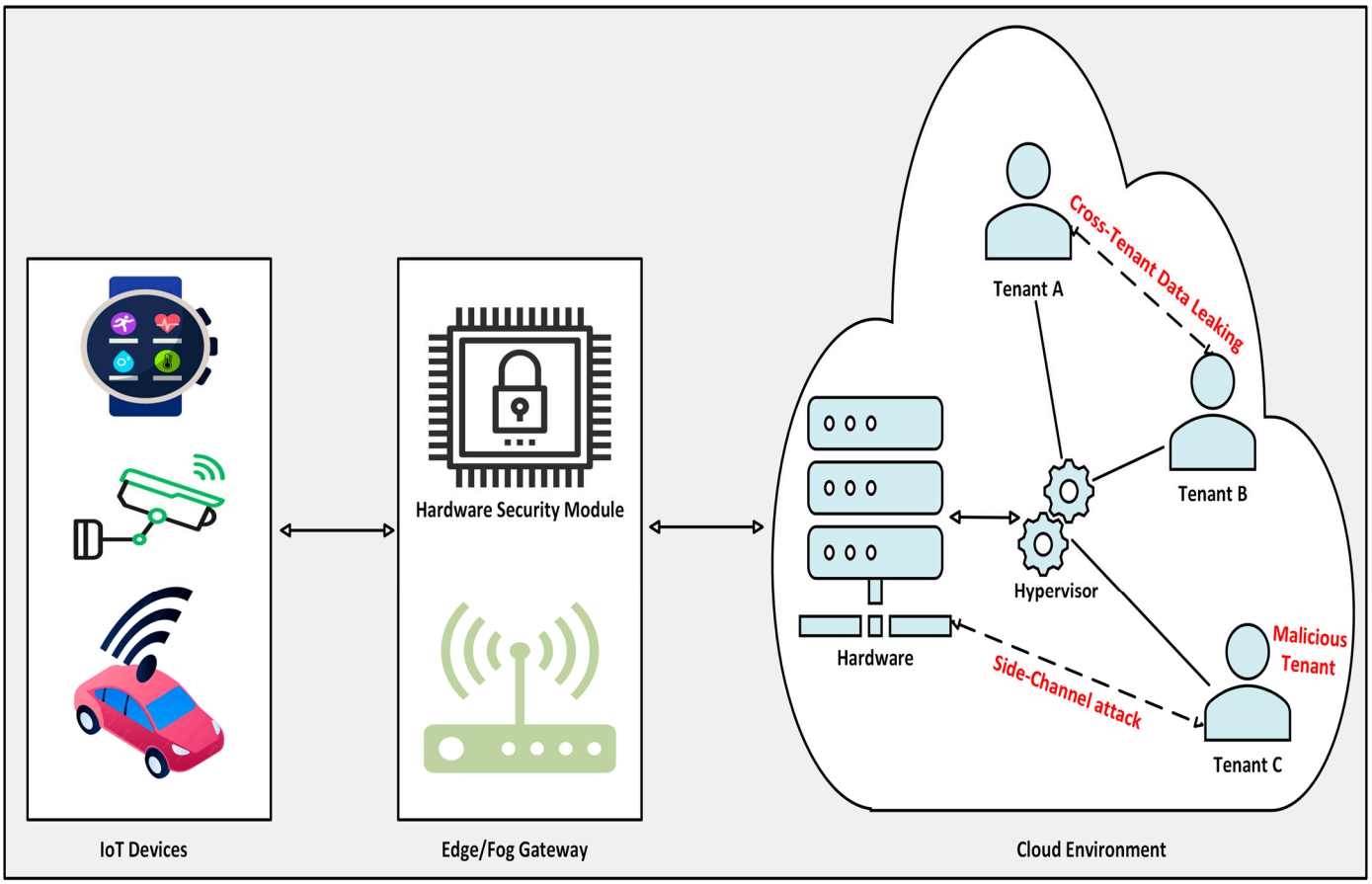


Figure 3. Multi-Tenant Threat Model in IoT–Cloud Environment.

Beyond logical threats, research also highlights the risk of physical and timing-based attacks. Side-channel exploits, DoS events, and API-level vulnerabilities are common in multi-tenant environments, and these security-efficiency trade-offs are challenging IoT workloads that demand low latency and minimal resource use.

In this analysis, Hashim et al. and Al-Balasmeh are categorized together as a single primary research theme focused on tenant-aware access control [3,41]. While Hashim et al. established the foundational benchmark that traditional cloud-based encryption and isolation mechanisms incur a performance overhead of approximately 12%, Al-Balasmeh provides a contemporary validation of this metric. Specifically, the latter study demonstrates that a lightweight Zero Trust framework can achieve a 95% reduction in unauthorized access while maintaining a consistent 12% RAM overhead. This grouping addresses previous limitations regarding the representativeness of isolated studies by providing empirical evidence of scalability across 1000 heterogeneous nodes. This transitions the analysis from theoretical vulnerability to a validated, scalable isolation model. It is important to note that these overheads are highly dependent on experimental conditions, such as the virtualization layer and hardware capabilities, and should not be interpreted as universal penalties. It is important to note that all studies selected for this review, as outlined in Table 1, were filtered to ensure they specifically address security within an IoT–Cloud context. The selected studies originate from geographically diverse institutions across Asia (Malaysia/Iraq [3], Bahrain [41], Japan [42], India [4,43,44]), North America (USA [5,39,40,45,46], Canada [47]), and Europe (Romania [48]). While some studies, such

as Kumar et al. and Yadav et al., utilize virtualization techniques common in general cloud computing, they were included only because their models were evaluated against specific IoT metrics like resource overhead or edge-layer latency.

Table 1. Summary of Selected Key Studies on IoT–Cloud Multi-Tenancy Isolation.

Study	Focus/Approach	Key Contribution	Observations
Hashim et al., 2024 [3]	Tenant isolation & access control	95% reduction in unauthorized access and eliminated insider privilege escalation.	12% resource overhead
Al-Balasmeh, H., 2025 [41]	Tenant isolation & Zero Trust validation	Confirmed Hashim et al.'s overhead; scalable across 1000 nodes	Lightweight framework
Surianarayanan et al., 2023 [4]	Data confidentiality	Exposed leakage risks in shared setups	No mitigation modeling
Panguraj et al., 2025 [40]	Resource sharing	Identified inter-tenant leakage paths	No integrated isolation model
Kumar et al., 2020 [39]	VM isolation	Simulated weak boundaries	Lacked IoT performance context
Kyriakidou et al., 2024 [42]	ABAC + Verifiable credentials	Tenant-level privacy authentication	Added computation overhead
Yadav et al., 2025 [43]	Zero Trust analytics	Continuous access validation	Not tested at scale
Pandit et al., 2025 [44]	AI anomaly detection	97.3% accuracy for tenant attacks	Training bias; scalability issue
Neto et al., 2022 [47]	Federated DDoS detection	84.2% accuracy; privacy preserved	Energy overhead
Almutairi et al., 2025 [5]	AI & PQC survey	Highlighted framework gaps	No unified integration
Malikireddy et al., 2024 [45]	Elasticity testing	Found inadequate dynamic security	No adaptive model proposed
Hariharan et al., 2025 [46]	Zero Trust	Continuous verification model	Conceptual; no IoT validation
Sebestyen et al., 2025 [48]	Blockchain audit	Immutable audit trails	Scalability constraints

4.2. Current Mitigation Strategies

Initial strategies for securing multi-tenancy were largely shaped by leading providers like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud. These companies established key safeguards like strong virtualization boundaries, tenant-scoped Identity and Access Management (IAM), Virtual Private Cloud (VPC) segmentation, hardware-assisted encryption, and continuous compliance monitoring. While effective as baseline protections, these solutions are primarily tailored to traditional cloud workloads and often do not account for the resource limitations and low-latency requirements typical of IoT systems [49].

Studies have proposed several defense strategies to mitigate these risks. For instance, Kyriakidou examined Attribute-Based Access Control (ABAC) combined with verifiable credentials, offering a privacy-preserving approach that supports tenant-specific identity verification [50]. Multi-layered security frameworks based on Zero Trust principles have also gained traction. This highlights ongoing authentication and stringent access control,

which are frequently supported by behavioral analytics to monitor user activity and identify real-time anomalies.

Artificial Intelligence (AI) and Machine Learning (ML) are increasingly applied to strengthen threat detection in multi-tenant environments. For example, Pandit developed an ML anomaly detection system, achieving 97.3% accuracy in identifying malicious tenant activity [44], while Neto proposed a federated learning approach for collaborative DDoS detection, reaching 84.2% accuracy without compromising tenant data privacy [47]. Despite ongoing optimizations, existing security mechanisms continue to impose significant overhead. This means the high resource demands of standard protocols like AES-128 contribute to this sustained 12% overhead [3,41]. Therefore, it underscores a primary design challenge, which entails balancing rigorous security protocols with the lightweight performance demands for IoT–cloud environments.

4.3. IoT Device Authentication Limitations in Multi-Tenant Cloud Environments

Although authentication and access control are generally regarded as establishing a fundamental layer for tenant isolation, empirical evidence indicates that vulnerabilities in device authentication often limit the efficacy of higher-level isolation mechanisms in practice. Analyses of measurement data from real-world IoT deployments show that these vulnerabilities are significant. Paracha et al. report that 11 out of 32 consumer IoT devices tested were susceptible to Transport Layer Security (TLS) interception due to improper certificate validation or trust-store issues, even though they supported standard secure-channel protocols [51]. These results show that the theoretical assurances provided by certificate-based mutual authentication can be significantly impacted by implementation issues and lifecycle vulnerabilities.

Black-box assessments further confirm that authentication failures are systemic rather than incidental. Lazzaro et al. identify recurring authentication bypass and replay vulnerabilities across a broad set of consumers' IoT devices. This is attributable to weak credential processing, acceptance of invalid trust anchors, and inconsistent validation behavior [52]. These validated observations indicate that authentication weaknesses persist even in deployed systems that follow established security models.

In multi-tenant IoT–cloud environments, these failures can have amplified consequences. Misconfigured or compromised device credentials may be reused across shared gateways, brokers, or identity services, thereby weakening tenant boundaries even when network segmentation and access-control policies are properly enforced. Therefore, authentication should be treated as a limiting factor in tenant isolation rather than as a standalone mitigation. The cited empirical studies confirm that without disciplined provisioning, validation, and trust-store management, authentication weaknesses can damage robust multi-tenant security architectures.

5. Discussion

Although significant progress has been made in threat classification and mitigation, contemporary research remains constrained by several persistent limitations. Table 2 shows previous methods that offered broad studies of IoT–Cloud security [5] or common cloud vulnerabilities [4]. Nevertheless, a comprehensive review specifically targeting the intersection of multi-tenancy challenges remains absent from current literature. This work bridges that gap by cataloging tenant-centric threats, such as isolation breaches and inter-tenant leakage, and assessing mitigation strategies tailored to the stringent performance limitations of CoT.

Table 2. Comparison of Past Studies and This Review.

Aspect	Previous Works	This Review
Scope	Earlier studies have largely examined IoT security or cloud security in isolation, with limited attention to the challenges introduced by shared tenancy [3,40].	Provides a review of security challenges in IoT–cloud multi-tenancy environments.
Target Environment	Most works addressed single-tenant or hybrid edge models [39,50].	Focuses on multi-tenant resource sharing and isolation.
Depth of Threat Analysis	Broader surveys have addressed general cloud security threats while offering limited analysis of tenant-specific risks [4,5].	Classifies tenant-level threats, including data leakage, privilege escalation, and cross-VM attacks.
Mitigation Techniques	Emphasized traditional encryption and access control [43,45].	Introduces adaptive models integrating ZTA, AI-driven detection, blockchain, and PQC.
Evaluation Focus	Prior work offered qualitative insights only [44,47].	Provides comparative evaluation based on scalability, latency, and isolation effectiveness.
Gap Analysis	Often lacked systematic categorization [46].	Delivers structured taxonomy of unresolved issues and testable metrics.
Post-Quantum Readiness	PQC rarely examined [3,53].	Places PQC as a key enabler for quantum-resistant multi-tenant communication.
Contribution Type	Mostly descriptive surveys [4,5].	Provides comparative synthesis and a roadmap for future research directions.

5.1. Fragmentation of Security Architectures

Our review identifies a critical lack of multi-tenant standardized security architectures, a deficit that has resulted in disjointed, specific platform implementations [5]. Current approaches fail to provide sufficient resource isolation, thereby exposing systems to risks of data leakage and privilege escalation [40]. Ultimately, the absence of a unified framework results in an unresolved difficult trade-off between security robustness and operational performance.

5.2. The Security-Performance Trade-Off

A critical barrier disclosed in this study is the inherent trade-off between implementing strong security and maintaining the performance necessary for IoT operations. While standard encryption and access controls are effective, they generate an unacceptable 12% overhead in IoT environments [3,41]. Specific examples, such as the 3.6 KB Flash memory requirement and 1.24 µJ/bit energy consumption of AES-128 [54], illustrate why these protocols fail in low-latency scenarios. Consequently, these demands create a bottleneck for low-latency applications, forcing a compromise between security and speed. This dynamic establishes a 'security-constraint feedback loop' wherein economic limitations compel the utilization of low-resource hardware, which in turn fails to support robust defense mechanisms. This consequently elevates the risk of vulnerabilities. As a result, this cycle serves as a fundamental obstacle to the scalability of protected multi-tenant architectures, representing a bottleneck that remains insurmountable without the aid of

specialized hardware acceleration. In addition to cryptographic and monitoring overheads, the review shows that multi-tenant isolation also depends on reliable device authentication. Implementation flaws and lifecycle weaknesses can still compromise higher-level security controls, even when architectural safeguards are in place.

5.3. Comparative Analysis of Emerging Technologies

Although multiple mitigation strategies have been proposed, they vary considerably in their appropriateness for restricted environments. Unlike standard cloud servers, IoT nodes often lack computational capability for heavy encryption or continuous monitoring. We identify and analyze four main trends that seek to address this core issue of conflict. A central challenge identified in this review is the inverse relationship between the strictness of isolation and system performance.

5.3.1. Zero Trust Architectures (ZTA)

In response to the ongoing challenge of the isolation gap in multi-tenant environments, ZTA is increasingly recognized as a fundamental security framework. Based on the principle of "never trust, always verify," ZTA regards every entity, including users, devices, or workloads, as untrusted until proper verification has been conducted [46]. Unlike traditional perimeter models, ZTA ensures ongoing authentication, authorization, and context validation for each access, regardless of its origin.

ZTA addresses the inherent risks of shared infrastructure by mitigating cross-tenant threats and preventing lateral movement. By prioritizing dynamic isolation mechanisms such as micro segmentation and real-time trust evaluation over static virtualization, this model offers a more robust defense. Furthermore, studies suggest that ZTA's unified approach to identity and telemetry management reduces policy implications and prevents unauthorized data disclosure [55].

Nevertheless, widespread implementation is obstructed by operational complexity, a fragmented tooling ecosystem, and significant performance overhead, particularly in large-scale or resource-limited environments. The latency introduced by continuous verification, coupled with the difficulty of enforcing uniform procedures throughout various tenants, presents a substantial barrier. Consequently, current scholarship advocates for the development of lightweight ZTA frameworks. These models must incorporate automated trust assessment, adaptive policy enforcement, and privacy-preserving monitoring to ensure both scalability and compliance with regulations.

Trade-off Analysis: Despite these benefits, ZTA offers the highest level of logical isolation at the cost of performance. As noted in recent implementations, the "never trust, always verify" model introduces significant latency overhead [56,57]. For real-time IoT applications (e.g., industrial automation), the continuous re-authentication required by ZTA can violate strict timing requirements (jitter). Consequently, current ZTA models are best suited for cloud backends or robust Edge gateways rather than constrained end-nodes where latency is critical.

5.3.2. AI-Driven Threat Detection

Standard testing frameworks often fail to align with the volatility of dynamic cloud provisioning [45]. In scenarios where multi-tenant containers spin up and down rapidly, static security policies cannot adapt with sufficient speed. This temporal misalignment creates a vulnerability gap, allowing malicious workloads to run during the delay between instantiation and rule updates. This limitation highlights the necessity for autonomous AI-based systems that detect anomalies in real-time, effectively decoupling security from the latency of manual configuration.

To resolve the friction between security rigor and system efficiency, contemporary research favors AI-driven, adaptive detection over static, resource-heavy mechanisms. Pandit et al. demonstrate the efficacy of this approach by using machine learning to analyze telemetry and inter-tenant traffic in real-time [44]. This capability is critical for differentiating between benign tenant workloads and cross-tenant attacks within shared infrastructures. Furthermore, Saxena et al. extend this adaptive paradigm by proposing a predictive model specifically designed to anticipate threats emerging from shared virtualization layers [42].

Trade-off Analysis: While AI models offer proactive defense against 'noisy neighbor' attacks, they introduce significant resource demands. As noted in recent studies on Deep Learning (DL) optimization, the substantial computational requirements of modern AI make them vulnerable to resource exhaustion attacks if not properly provisioned [58]. AI-driven threat detection, therefore, necessitates clear trade-offs across the IoT–cloud continuum. On-device deployment minimizes inference latency and limits data exposure but is constrained to lightweight models, incurs significant energy overhead, and is more vulnerable to physical tampering. Cloud-based deployment supports large-scale model training but introduces higher end-to-end latency and increases privacy leakage through the transmission of raw telemetry across trust boundaries. Edge/Fog deployment provides a pragmatic balance by enabling more expressive models with lower latency than cloud execution while reducing data exposure relative to centralized processing [59]. However, this approach requires trusted and well-isolated gateways to prevent cross-tenant interference. Considering these trade-offs, AI-based threat detection is best deployed at the Fog/Edge layer in multi-tenant IoT–cloud systems.

5.3.3. Blockchain Integration

Blockchain technology addresses the current deficit in verifiable auditing within multi-tenant architecture by establishing a decentralized, transparent trust model. As noted by Sebestyen et al., the implementation of a distributed ledger creates an immutable record of transactions that tenants and service providers can access [48]. This shared visibility guarantees transparency and maintains data integrity, thereby effectively reducing the risk of insider threats and unauthorized access.

Trade-off Analysis: While blockchain strengthens accountability and forensic readiness, its integration into IoT–cloud systems must be selective due to latency, storage, and scalability constraints. Ledger interaction can be confined to the edge and cloud layers, where computational resources and network stability permit reliable operation. IoT end devices do not directly interact with the ledger. Instead, edge gateways and cloud control services record low-frequency, security-critical events, such as access-control policy updates, firmware integrity attestations, and cross-tenant isolation decisions. To maintain scalability, bulk telemetry and application data are stored off-chain, with only compact cryptographic hashes or metadata anchored on-chain to preserve auditability. Ledger transactions are triggered asynchronously and event-driven, rather than continuously, ensuring that blockchain overhead does not interfere with latency-sensitive workloads. This hybrid on-chain/off-chain model preserves audit trail integrity while avoiding prohibitive computational and operational costs [48].

5.3.4. Post-Quantum Cryptography (PQC)

Most existing multi-tenant frameworks largely overlook the 'harvest now, decrypt later' threat associated with the advent of quantum computing. The continued reliance on classical encryption standards (RSA, ECC) leaves long-lifespan IoT data exposed to future decoding techniques [53]. This omission creates a significant resilience gap, as the current lack of lightweight PQC integration threatens long-term confidentiality.

To mitigate this cryptographic elasticity gap, PQC is identified as the necessary forward-looking solution [3]. As noted by Peng et al., integrating PQC is essential to mitigate the risk of cross-tenant data breaches once classical cryptographic primitives are inevitably compromised [60].

Trade-off Analysis: Implementing PQC introduces significant new risks related to bandwidth and side-channels. Algorithms like CRYSTALS-Kyber have larger key sizes and ciphertext overheads compared to traditional ECC [53]. This bandwidth consumption makes them challenging for Low-Power Wide-Area Networks (LPWANs) like LoRaWAN, which have payload limits. Additionally, complex PQC operations can increase the risk of physical side-channel attacks (SCA). Therefore, future models should emphasize hardware-accelerated PQC to reduce leakage.

5.3.5. Strategic Synthesis

Collectively, these trends signal a paradigm shift toward intelligent, distributed, and inherently future-proof security architectures. Moving away from monolithic, one-size-fits-all approaches, the next generation of multi-tenant security will likely be defined by a defense-in-depth strategy. This composite model synergizes ZTA for logical boundary enforcement, AI for real-time anomaly detection, Blockchain for decentralized auditing, and PQC for long-term cryptographic durability. Far from being mutually exclusive, these technologies function as a cohesive suite within IoT–Cloud ecosystems. Table 3 synthesizes this strategic roadmap, mapping each mechanism to the specific security gap it resolves while accounting for associated implementation constraints.

Table 3. Comparative Evaluation of Multi-Tenancy Defense Mechanisms.

Mechanism	Primary Strength	Key IoT Constraint	Best-Fit Deployment
ZTA	Granular access enforcement	Latency and network jitter	Cloud backends, edge gateways
AI/ML-Based Detection	Proactive threat identification.	High compute and energy demand	Edge servers (fog layer)
Blockchain	Tamper-proof auditability	Storage overhead, scalability	Private or permissioned cloud layers
PQC (e.g., CRYSTALS-Kyber)	Long-term data confidentiality.	Increased bandwidth from key sizes	High-value, long-lifecycle data

5.4. Feasibility and Implementation Challenges

While ZTA and PQC offer robust security, integrating them into IoT environments poses significant feasibility challenges.

- Protocol Overhead and Bandwidth Constraints: PQC algorithms, such as CRYSTALS-Kyber, introduce substantial communication overhead due to their larger key and ciphertext sizes (approximately 1.6 KB) [61]. This creates a critical feasibility gap when measured against contemporary radio communication protocols. For example, LoRaWAN typically supports a physical payload maximum transmission unit (MTU) ranging from only 51 to 222 bytes, depending on the data rate. Direct transmission of a 1.6 KB PQC key would require extensive packet fragmentation and multiple transmission windows, leading to significant battery depletion and increased collision risks in the shared radio spectrum. Furthermore, while Bluetooth Low Energy (BLE) and Zigbee offer slightly higher throughput, they still struggle with the "harvest now, decrypt later" resilience gap without hardware acceleration. As a result, the direct deployment of PQC at the IoT end-node layer remains infeasible for many

low-power and low-bandwidth applications. This necessitates a tiered architectural approach wherein the intensive cryptographic handshake is delegated to the Edge/Fog layer, which has the high-bandwidth backhaul (e.g., Ethernet or Wi-Fi) requisite for managing PQC-sized artifacts.

- Deployment Cost: Continuous verification under ZTA requires policy enforcement engines, identity management services, and telemetry collection mechanisms that are often unsupported by legacy IoT gateways. Remodifying existing deployments to adapt these components introduces non-trivial infrastructure and operational costs [62]. This limits the economic feasibility of full ZTA adoption in large-scale or cost-sensitive IoT deployments.
- Legacy Interoperability and Operational Complexity: Beyond the direct computational costs, the deployment of next-generation security models faces severe friction from legacy infrastructure. Industrial and smart city ecosystems often rely on long-lived devices that use hard-coded, proprietary protocols that lack native support for dynamic policy enforcement agents or modern identity management. Retrofitting these heterogeneous endpoints for ZTA compliance often demands a prohibitive "rip-and-replace" strategy, creating a migration cost barrier that outweighs the security benefits for many organizations. Moreover, the operational complexity of managing fragmented tooling ecosystems across multi-vendor devices creates a significant administrative bottleneck [63]. These practical limitations reinforce the necessity of the proposed Edge-Gateway architectural shift; by abstracting security enforcement to the gateway layer, organizations can envelop legacy devices in a secure ZTA perimeter without requiring infeasible hardware upgrades at the sensor level.
- Latency: The "verify-then-trust" model intrinsic to ZTA introduces additional round-trip communication and processing delays. For latency-critical IoT applications, such as industrial control systems (ICS) with timing constraints below 10 ms, this added delay can violate real-time requirements [64]. That leads to jitter and system instability.

Collectively, these constraints indicate that full ZTA enforcement and PQC-based key exchange are infeasible at the IoT end-node layer under strict latency, bandwidth, and cost limitations. As summarized in Table 4, the sensor layer faces critical bandwidth limitations for PQC and high jitter risks for ZTA, necessitating a tiered approach to viability. Therefore, this review derives a multi-layer security design principle for multi-tenant IoT–cloud systems: lightweight authentication and symmetric cryptography should be applied at the device layer; trust evaluation, policy enforcement, and anomaly detection should be offloaded to hardware-accelerated edge gateways; computationally intensive cryptographic operations and compliance mechanisms should be confined to the cloud layer. This layered deployment model provides strong tenant isolation while accommodating performance constraints inherent to IoT environments and serves as a practical design baseline for future multi-tenant security architectures.

Table 4. Feasibility Analysis of Emerging Technologies per Layer.

Technology	Metric of Concern	Quantified Technical Impact	Sensor Layer Viability	Edge Layer Viability	Cloud Layer Viability	Deployment Cost
PQC (Kyber)	Bandwidth & MTU	~1.6 KB public key; LoRaWAN MTU 51–222 bytes. 7–30 packet fragments per exchange	Critical Fail (Exceeds MTU)	High (Ethernet/WiFi)	High	Energy cost (μJ/bit), firmware memory increase (KB), hardware upgrade expenses

Table 4. *Cont.*

Technology	Metric of Concern	Quantified Technical Impact	Sensor Layer Viability	Edge Layer Viability	Cloud Layer Viability	Deployment Cost
ZTA	RTT & Processing Delay	Additional RTT per access request; >5–10 ms unacceptable in sub −10 ms ICS	Low (Jitter risks)	Medium/High (PEP placement)	High	Compute scaling operating expenses; policy engine licensing

6. Future Research Roadmap

To resolve the conflict between security robustness and IoT performance, future research must move beyond software-only patches and adopt specific architectural shifts. We propose three concrete research directions:

6.1. Federated Learning at the IoT Edge

While centralized AI offers high accuracy, it risks data leakage in multi-tenant clouds. Future architectures should deploy Federated Learning (FL) directly at the IoT Edge layer [65]. Therefore, individual tenants train anomaly detection models locally on their sensor data and share only model updates (e.g., gradients, not raw data) with a global cloud aggregator. Research is urgently needed to optimize lightweight FL algorithms for resource-constrained microcontrollers (e.g., ARM Cortex-M) to enable privacy-preserving intrusion detection without compromising tenant isolation [66].

6.2. Hardware Enclaves in Fog Computing

Current software-based isolation (e.g., Docker) is insufficient for high-security applications due to the shared OS kernel surface. Future multi-tenant architectures should leverage Trusted Execution Environments (TEEs), such as ARM TrustZone or Intel SGX, specifically within Fog Computing nodes [67]. By mapping distinct tenants to isolated hardware enclaves, the system can guarantee memory protection even if the underlying OS is compromised. Future work must develop efficient context-switching mechanisms to dynamically allocate limited isolated memory among competing tenants without introducing prohibitive latency.

6.3. Lightweight Distributed Ledgers at the Gateway

To ensure data integrity verification in decentralized IoT systems, future work should investigate Distributed Ledger Technology (DLT) implemented at the IoT Gateway layer. Instead of power-intensive blockchain agreements, these architectures should use Directed Acyclic Graph (DAG) structures to log access and confirm data integrity [68]. This establishes a decentralized and immutable record of tenant interactions, verified by the edge network, thereby offering scalable verification without dependence on centralized trusted third parties.

7. Conclusions

This review analyzes the security and privacy challenges inherent to multi-tenant IoT–Cloud environments. By examining the intersection of resource-constrained IoT devices and shared cloud infrastructure, we mapped primary threats, specifically cross-tenant data leakage, insecure APIs, and side-channel vulnerabilities, against the current landscape of mitigation techniques. The primary finding of this review highlights a critical and unresolved tension between security robustness and system performance. Many current solutions impose significant computational and latency overheads, making them impracti-

cal for lightweight IoT devices. This gap emphasizes the limitations of existing approaches and signals a clear need for more adaptive, low-overhead security frameworks.

To address this gap, this work presents a forward-looking research roadmap focused on adaptive, lightweight, and hardware-assisted security frameworks. Specifically, future research should emphasize: (1) edge-based adaptive AI, such as federated learning, to enable privacy-preserving intrusion detection without exposing tenant data; (2) hardware-backed isolation mechanisms, including TEE in fog nodes, to enforce strong tenant separation beyond shared operating systems; and (3) lightweight decentralized verification mechanisms, such as DAG-based distributed logs at the gateway layer, to provide unmanipulated responsibility with minimal energy and storage overhead. These architectural directions enable ZTA enforcement, scalable anomaly detection, and decentralized auditability while remaining compatible with strict IoT resource constraints. PQC acts as a cross-cutting requirement to secure long-lived IoT data against future quantum threats without degrading real-time performance. Ultimately, the successful deployment of these technologies relies on bridging the implementation gap. Future research should prioritize the development of hardware-accelerated security modules and edge offloading architectures that enable multi-tenant systems to maintain strict isolation while preserving IoT edge performance.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/iot7010021/s1. Supplementary File S1: Systematic review checklist used in this study (e.g., PRISMA checklist) [69].

Author Contributions: Conceptualization, B.A.; methodology, B.A.; validation, B.A. and M.G.A.; formal analysis, B.A. and M.G.A.; investigation, B.A. and M.G.A.; resources, B.A.; data curation, F.T.S.; writing—original draft preparation, B.A.; writing—review and editing, B.A., F.T.S. and M.G.A.; visualization, B.A.; supervision, F.T.S.; project administration, F.T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: During the writing phase of this manuscript, the authors used Gemini 3 only for language editing. All outputs were reviewed and revised by the authors, who take full responsibility for the final content.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IoT	Internet of Things
PQC	Post-Quantum Cryptography
ZTA	Zero Trust Architectures
CoAP	Constrained Application Protocol
LPWANs	Low-Power Wide Area Networks
SSL	Secure Sockets Layer
RBAC	Role-Based Access Control
VM	virtual machine
QoS	Quality of Service
VPCs	Virtual Private Clouds
ABAC	Attribute-Based Access Control
APIs	Application Programming Interfaces
TEEs	Trusted Execution Environments
DAG	Directed Acyclic Graph

ML	Machine Learning
TLS	Transport Layer Security
MTU	Maximum Transmission Unit
DL	Deep Learning
MQTT	Message Queuing Telemetry Transport
SaaS	Software as a Service
IaaS	Infrastructure as a Service
AMQP	Advanced Message Queuing Protocol
ACLs	Access Control Lists
Amazon EC2	Amazon Elastic Compute Cloud
PaaS	Platform as a Service
SDN	Software-Defined Networking
VPN	Virtual Private Networks
VLAN	Virtual Local Area Network
DoS	Denial-of-Service
SCA	side-channel attacks
FL	Federated Learning
DLT	Distributed Ledger Technology
AI	Artificial Intelligence
BLE	Bluetooth Low Energy
PKI	Public Key Infrastructure
ICS	Industrial Control Systems

References

1. Kuchuk, H.; Malokhvii, E. Integration of Iot With Cloud, Fog, and Edge Computing: A Review. *Adv. Inf. Syst.* **2024**, *8*, 65–78. [CrossRef]
2. Zoting, S.; Aditi, S. Multi-Tenant Data Centers Market Size, Share and Trends 2025 to 2034. 2025. Available online: https://www.precedenceresearch.com/multi-tenant-data-centers-market (accessed on 15 November 2025).
3. Hashim, W.; Hussein, N.A.-H.K. Securing Cloud Computing Environments: An Analysis of Multi-Tenancy Vulnerabilities and Countermeasures. *SHIFRA* **2024**, *2024*, 8–16. [CrossRef]
4. Surianarayanan, C.; Chelliah, P.R. Integration of the Internet of Things and Cloud: Security Challenges and Solutions—A Review. *Int. J. Cloud Appl. Comput. (IJCAC)* **2023**, *13*, 1–30. [CrossRef]
5. Almutairi, M.; Sheldon, F.T. IoT–Cloud Integration Security: A Survey of Challenges, Solutions, and Directions. *Electronics* **2025**, *14*, 1394. [CrossRef]
6. Botta, A.; De Donato, W.; Persico, V.; Pescapé, A. Integration of Cloud computing and Internet of Things: A survey. *Future Gener. Comput. Syst.* **2016**, *56*, 684–700. [CrossRef]
7. Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **2013**, *29*, 1645–1660. [CrossRef]
8. Mrabet, H.; Belguith, S.; Alhomoud, A.; Jemai, A. A Survey of IoT Security Based on a Layered Architecture of Sensing and Data Analysis. *Sensors* **2020**, *20*, 3625. [CrossRef]
9. Bello, O.; Zeadally, S.; Badra, M. Network layer inter-operation of Device-to-Device communication technologies in Internet of Things (IoT). *Ad Hoc Netw.* **2017**, *57*, 52–62. [CrossRef]
10. Hussain, B.; Elmedany, W.; Sharif, M.S. The Internet of Things Security Issues and Countermeasures in Network Layer: A Systematic Literature Review. In Proceedings of the 2022 International Conference on Data Analytics for Business and Industry (ICDABI), Sakhir, Bahrain, 25–26 October 2022.
11. Dallaf, A.A.A. Edge Computing in IoT Networks: Enhancing Efficiency, Reducing Latency, and Improving Scalability. *Int. J. Adv. Netw. Monit. Control.* **2025**, *10*, 103–115. [CrossRef]
12. Sarwar, K.; Yongchareon, S.; Yu, J.; ur Rehman, S. Efficient privacy-preserving data replication in fog-enabled IoT. *Future Gener. Comput. Syst.* **2022**, *128*, 538–551. [CrossRef]
13. Mavridis, I.; Karatza, H. Orchestrated sandboxed containers, unikernels, and virtual machines for isolation-enhanced multitenant workloads and serverless computing in cloud. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e6365. [CrossRef]
14. Lynn, T.; Rosati, P.; Lejeune, A.; Emeakaroha, V. A Preliminary Review of Enterprise Serverless Cloud Computing (Function-as-a-Service) Platforms. In Proceedings of the 2017 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), Hong Kong, China, 11–14 December 2017.

15. Jain, P.; Munjal, Y.; Gera, J.; Gupta, P. Performance Analysis of Various Server Hosting Techniques. *Procedia Comput. Sci.* **2020**, *173*, 70–77. [CrossRef]
16. Van Eyk, E.; Grohmann, J.; Eismann, S.; Bauer, A.; Versluis, L.; Toader, L.; Schmitt, N.; Herbst, N.; Abad, C.L.; Iosup, A. The SPEC-RG Reference Architecture for FaaS: From Microservices and Containers to Serverless Platforms. *IEEE Internet Comput.* **2019**, *23*, 7–18. [CrossRef]
17. Armoogum, S.; Khonje, P. Healthcare Data Storage Options Using Cloud. In *The Fusion of Internet of Things, Artificial Intelligence, and Cloud Computing in Health Care*; Springer: Cham, Switzerland, 2021.
18. Hu, Z.; Zhang, H.; Sun, S.; Gao, C.; Li, Y.; Li, X. FDRA: Fully Distributed Routing Architecture for Private Virtual Network in Public Cloud. In Proceedings of the 11th International Symposium on Parallel Architectures, Algorithms and Programming, Shenzhen, China, 28–30 December 2021.
19. Um, T.W.; Lee, H.; Ryu, W.; Choi, J.K. Dynamic Resource Allocation and Scheduling for Cloud-Based Virtual Content Delivery Networks. *ETRI J.* **2014**, *36*, 197–205. [CrossRef]
20. Zahariadis, T.; Papadakis, A.; Alvarez, F.; Gonzalez, J.; Lopez, F.; Facca, F.; Al-Hazmi, Y. FIWARE Lab: Managing Resources and Services in a Cloud Federation Supporting Future Internet Applications. In Proceedings of the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, London, UK, 8–11 December 2014.
21. Zhang, X.; Du, H.T.; Chen, J.Q.; Lin, Y.; Zeng, L.J. Ensure Data Security in Cloud Storage. In Proceedings of the 2011 International Conference on Network Computing and Information Security, Guilin, China, 14–15 May 2011.
22. Waghchaude, K. A Review on Cloud Computing Security Issues, Applicable Solutions and Implementation. *Int. J. Sci. Res. Eng. Manag.* **2024**, *8*, 1–3. [CrossRef]
23. Sood, S.K. A combined approach to ensure data security in cloud computing. *J. Netw. Comput. Appl.* **2012**, *35*, 1831–1838. [CrossRef]
24. Pal, S.; Khatua, S.; Chaki, N.; Sanyal, S. A New Trusted and Collaborative Agent Based Approach for Ensuring Cloud Security. *arXiv* **2011**, arXiv:1108.4100. [CrossRef]
25. Gibson, J.; Rondeau, R.; Eveleigh, D.; Tan, Q. Benefits and challenges of three cloud computing service models. In Proceedings of the 2012 4th International Conference on Computational Aspects of Social Networks (CASoN), Sao Carlos, Brazil, 21–23 November 2012.
26. De Napoli, C.; Forestiero, A.; Fortino, G.; Giordano, A.; Guerrieri, A.; Lagana, D.; Lupi, G.; Mastroianni, C.; Spataro, L. IoT-HC: A Novel IoT Architecture for the Hybrid Cloud. In Proceedings of the 2019 28th International Conference on Computer Communication and Networks (ICCCN), Valencia, Spain, 29 July–1 August 2019.
27. Yassein, M.B.; Shatnawi, M.Q.; Al-zoubi, D. Application layer protocols for the Internet of Things: A survey. In Proceedings of the 2016 International Conference on Engineering & MIS (ICEMIS), Agadir, Morocco, 22–24 September 2016.
28. Altayaran, S.; Elmedany, W. Security threats of application programming interface (API's) in internet of things (IoT) communications. In Proceedings of the 4th Smart Cities Symposium (SCS 2021), Online, 21–23 November 2021.
29. Ferdows, J.; Mehedi, S.T.; Hossain, A.D.; Shamim, A.A.M.; Rasiq, G.R.I. A Comprehensive Study of IoT Application Layer Security Management. In Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 6–8 November 2020.
30. Kaur, K.; Kaur, A.; Gulzar, Y.; Gandhi, V. Unveiling the core of IoT: Comprehensive review on data security challenges and mitigation strategies. *Front. Comput. Sci.* **2024**, *6*, 1420680. [CrossRef]
31. Khan, Y.; Su'ud, M.B.M.; Alam, M.M.; Ahmad, S.F.; Salim, N.A.; Khan, N. Architectural Threats to Security and Privacy: A Challenge for Internet of Things (IoT) Applications. *Electronics* **2023**, *12*, 88. [CrossRef]
32. Banerjee, S.; Parisa, S.K. Secure Multi-Tenancy in Cloud Computing: Challenges and Solutions. *Trans. Recent Dev. Ind. IoT* **2025**, *17*.
33. Sharma, R.K. Multi-Tenant Architectures in Modern Cloud Computing: A Technical Deep Dive. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2025**, *11*, 307–317. [CrossRef]
34. Njeguš, A. Intelligent Software Systems for Multi-Tenant Cloud Environments: Challenges and Solutions. In Proceedings of the Sinteza 2025—International Scientific Conference on Information Technology, Computer Science, and Data Science, Novi Sad, Serbia, 9 May 2025.
35. Del Piccolo, V.; Amamou, A.; Haddadou, K.; Pujolle, G. A Survey of Network Isolation Solutions for Multi-Tenant Data Centers. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 2787–2821. [CrossRef]
36. Factor, M.; Hadas, D.; Harnama, A.; Har'El, N.; Kolodner, E.K.; Kurmus, A.; Shulman-Peleg, A.; Sorniotti, A. Secure Logical Isolation for Multi-tenancy in cloud storage. In Proceedings of the IEEE Conference on Mass Storage Systems and Technologies, Long Beach, CA, USA, 6–10 May 2013.
37. Bhattacharyya, S.; Kanka, V.; Mohammed, A.S. Optimizing Resource Isolation Techniques in Multi-Tenant PaaS Architectures Using Kubernetes and Virtualization. *J. Artif. Intell. Res.* **2021**, *1*, 197–239.

38. Basu, M.T.; Sastry, J.K.R. Enhancing Data Security under Multi-Tenancy within Open Stack. *Int. J. Adv. Trends Comput. Sci. Eng.* **2020**, *9*, 533–544. [CrossRef]
39. Kumar, R. Multi-Tenant SaaS Architectures: Design Principles and Security Considerations. *J. Softw. Eng. Simul.* **2020**, *6*, 28–41. [CrossRef]
40. Panguraj, A.R.R. Systematic Approach to Security Testing in Multi-Tenant Cloud Systems. *Int. J. Multidiscip. Res. Growth Eval.* **2025**, *6*, 2139–2412. [CrossRef]
41. Al-Balasmeh, H. Zero Trust Architecture for IoT Device Ecosystems. *Int. J. Basic Appl. Sci.* **2025**, *14*, 818–825. [CrossRef]
42. Saxena, D.; Gupta, I.; Gupta, R.; Singh, A.K.; Wen, X. An AI-Driven VM Threat Prediction Model for Multi-Risks Analysis-Based Cloud Cybersecurity. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 6815–6827. [CrossRef]
43. Yadav, S.; Abidin, S. Enhancing Security in Multi-Tenant Cloud Environments: Threat Detection, Prevention, and Data Breach Mitigation. *J. Inf. Syst. Eng. Manag.* **2025**, *10*, 61–72. [CrossRef]
44. Pandit, A.; Pandit, R. Side-Channel Attacks in Multi-Tenant Cloud Environments: Prevention & Mitigation. *Int. J. Innov. Sci. Eng. Manag.* **2025**, *4*, 93–105.
45. Malikireddy, S.K.R. Securing Multi-Tenant Cloud Environments with Graph-Based Models. *Int. J. Sci. Res. Eng. Manag.* **2024**, *6*, 1–21. [CrossRef]
46. Hariharan, R. Zero Trust Security in Multi-Tenant Cloud Environments. *J. Inf. Syst. Eng. Manag.* **2025**, *10*, 623–644. [CrossRef]
47. Neto, E.C.P.; Dadkhah, S.; Ghorbani, A.A. Collaborative DDoS Detection in Distributed Multi-Tenant IoT using Federated Learning. In Proceedings of the 2022 19th Annual International Conference on Privacy, Security & Trust (PST), Fredericton, NB, Canada, 22–24 August 2022.
48. Sebestyen, H.; Popescu, D.E.; Zmaranda, R.D. A Literature Review on Security in the Internet of Things: Identifying and Analysing Critical Categories. *Computers* **2025**, *14*, 61. [CrossRef]
49. Dauda, A.; Flauzac, O.; Nolot, F. A Survey on IoT Application Architectures. *Sensors* **2024**, *24*, 5320. [CrossRef]
50. Kyriakidou, C.N.; Papathanasiou, A.M.; Pittaras, I.; Fotiou, N.; Thomas, Y.; Polyzos, G.C. Attribute-Based Access Control Utilizing Verifiable Credentials for Multi-Tenant IoT Systems. In Proceedings of the 2024 IEEE 4th International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB), Taipei, Taiwan, 19–21 April 2024.
51. Paracha, M.T.; Dubois, D.J.; Vallina-Rodriguez, N.; Choffnes, D. IoTLS: Understanding TLS usage in consumer IoT devices. In Proceedings of the 21st ACM Internet Measurement Conference, Virtual, 2–4 November 2021.
52. Lazzaro, S.; De Angelis, V.; Mandalari, A.M.; Buccafurri, F. A black-box assessment of authentication and reliability in consumer IoT devices. *Pervasive Mob. Comput.* **2025**, *110*, 102045. [CrossRef]
53. Mahdi, L.H.; Abdullah, A.A. Fortifying Future IoT Security: A Comprehensive Review on Lightweight Post-Quantum Cryptography. *Eng. Technol. Appl. Sci. Res.* **2025**, *15*, 21812–21821. [CrossRef]
54. Awasthi, A. Quantum-Resistant Security for IoT Systems Challenges and Implementation Strategies. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2025**, *11*, 671–678. [CrossRef]
55. Oladimeji, G. A Critical Analysis of Foundations, Challenges and Directions for Zero Trust Security in Cloud Environments. *arXiv* **2024**, arXiv:2411.06139. [CrossRef]
56. Harth-Kitzerow, C.; Garrido, G.M. Verifying Outsourced Computation in an Edge Computing Marketplace. *arXiv* **2022**, arXiv:2203.12347. [CrossRef]
57. Donovan, S.; Feamster, N. Alternative Trust Sources: Reducing DNSSEC Signature Verification Operations with TLS. In Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, London, UK, 17–21 August 2015.
58. Zhang, Y.; Wang, X.; Gao, H.; Zhou, Z.; Meng, F.; Zhang, Y.; Su, S. PD3F: A Pluggable and Dynamic DoS-Defense Framework Against Resource Consumption Attacks Targeting Large Language Models. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Suzhou, China, 4–9 November 2025.
59. Andriulo, F.C.; Fiore, M.; Mongiello, M.; Traversa, E.; Zizzo, V. Edge Computing and Cloud Computing for Internet of Things: A Review. *Informatics* **2024**, *11*, 71. [CrossRef]
60. Peng, Y.; Jiang, X.; Wang, S.; Xiang, Y.; Xing, L. An Improved Co-Resident Attack Defense Strategy Based on Multi-Level Tenant Classification in Public Cloud Platforms. *Electronics* **2024**, *13*, 3273. [CrossRef]
61. Liu, T.; Ramachandran, G.; Jurdak, R. Towards Quantum Resilient IoT: A Backward-Compatible Approach to Secure BLE Key Exchange Against Quantum Threats. In Proceedings of the International Conference on Internet-of-Things Design and Implementation, Hong Kong, China, 13–16 May 2024.
62. Mushtaq, S.; Mohsin, M.; Mushtaq, M.M. A Systematic Literature Review on the Implementation and Challenges of Zero Trust Architecture Across Domains. *Sensors* **2025**, *25*, 6118. [CrossRef]
63. Roman, R.; Lopez, J.; Mambo, M. Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges. *Future Gener. Comput. Syst.* **2018**, *78*, 680–698. [CrossRef]
64. Almuseelem, W. Secure Latency-Aware Task Offloading Using Federated Learning and Zero Trust in Edge Computing for IoMT. *IEEE Access* **2025**, *13*, 117808–117830. [CrossRef]

65. Chandu, G.; Karthik, T.; Parag, B. Federated Learning for Distributed IoT Security: A Privacy-Preserving Approach to Intrusion Detection. *IEEE Access* **2025**, *13*, 135863–135875.
66. Belenguer, A.; Navaridas, J.; Pascual, J.A. A review of Federated Learning in Intrusion Detection Systems for IoT. *arXiv* **2022**, arXiv:2204.12443. [CrossRef]
67. Valadares, D.C.G.; Will, N.C.; Spohn, M.A.; de Souza Santos, D.F.; Perkusich, A.; Gorgonio, K.C. Trusted Execution Environments for Cloud/Fog-based Internet of Things Applications. In Proceedings of the 11th International Conference on Cloud Computing and Services Science (CLOSER 2021), Virtual, 28–30 April 2021; pp. 111–121.
68. Cherupally, S.R.; Boga, S.; Podili, P.; Kataoka, K. Lightweight and Scalable DAG based distributed ledger for verifying IoT data integrity. In Proceedings of the International Conference on Information Networking, Jeju Island, Republic of Korea, 13–16 January 2021.
69. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.