International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 7, Issue 2, 50-64, 2026
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V7I2P109

Original Article

Zero Trust Architecture for Smart Factories: Securing Digital Twins and Cyber-Physical OT Systems

Vignesh Alagappan
Senior Member, IEEE, Rheem Manufacturing, Roswell, Georgia, USA.

Received On: 02/03/2026 ***Revised On: 27/03/2026*** ***Accepted On: 04/04/2026*** ***Published On: 15/04/2026***

Abstract - *The rapid digitization of manufacturing environments has produced highly interconnected cyber-physical ecosystems in which traditional security boundaries no longer hold. Operational Technology (OT) networks once isolated by design now interface with enterprise IT systems, cloud platforms, and AI-driven analytics engines, creating an expanded and heterogeneous attack surface. Perimeter-centric security models, which assume implicit trust within network boundaries, are structurally incapable of addressing this complexity. This paper proposes a comprehensive Zero Trust Architecture (ZTA) tailored for smart factory environments. Five principal contributions are made: (1) an identity-centric trust model anchored in hardware roots of trust and cryptographic device attestation using TPM 2.0 and X.509 certificates, with a PAA-PAI-DAC hierarchy aligned to the CSA Matter specification, and post-quantum cryptographic agility implemented via NIST FIPS 203/204/205; (2) a trust-aware data pipeline that enforces continuous verification at every telemetry stage from field device to cloud analytics and digital twin state; (3) a safety-security co-design framework that reconciles Zero Trust enforcement with the availability and determinism requirements of industrial control systems; (4) a structured threat scenario evaluation anchored to MITRE ATT&CK for ICS techniques T0856, T0859, T0867, and T0862; and (5) a compliance alignment mapping between ZTA principles and IEC 62443 security levels and NIST CSF 2.0 functions. Results demonstrate that identity-centric Zero Trust enforcement significantly reduces the exploitable attack surface while preserving operational continuity in brownfield environments.*

Keywords - *Zero Trust Architecture, Smart Factory Security, Digital Twin Integrity, OT-IT Convergence, IEC 62443, Hardware Root of Trust, Device Identity, PKI, Post-Quantum Cryptography, Matter DAC, MITRE ATT&CK for ICS, Brownfield Deployment, Cyber-Physical Systems, Anomaly Detection.*

1. Introduction

Manufacturing enterprises worldwide are undergoing a structural transformation driven by the convergence of Operational Technology (OT) and Information Technology (IT). The proliferation of Industrial Internet of Things (IIoT) devices, cloud-hosted supervisory systems, and AI-driven analytics platforms has enabled unprecedented levels of operational visibility and efficiency. Smart factories leverage real-time telemetry, predictive maintenance schedules derived from machine-learning models, and digital twin simulations to optimize production lines, reduce unplanned downtime, and compress product development cycles. This connectivity revolution, however, carries a commensurate security burden: each integration point between OT and IT introduces potential attack vectors that adversaries can exploit with increasingly sophisticated techniques.

Historically, OT environments derived their security posture from physical isolation. Air-gapped networks, proprietary communication protocols such as Modbus RTU and DNP3, and long device lifecycles created a de facto security boundary that discouraged broad-spectrum cyberattacks. This model has eroded rapidly. [5], [29] The integration of Ethernet-based field buses, MQTT brokers, OPC-UA endpoints, and REST APIs into factory floors has made OT systems accessible and therefore vulnerable in ways that were architecturally impossible a decade ago. High-profile incidents, including the Triton/TRISIS malware attack on a petrochemical safety instrumented system [2] and the Colonial Pipeline ransomware event, [1] demonstrated that OT system compromise can have catastrophic physical and economic consequences.

Traditional perimeter-based security architectures are structurally unsuited to this environment. Perimeter defenses assume a clear, stable boundary between trusted internal networks and untrusted external networks an assumption that dissolves when cloud services, remote vendor access, mobile operators, and third-party analytics pipelines continuously span that boundary. Zero Trust Architecture (ZTA), as formalized by NIST SP 800-207, provides a principled alternative: no entity whether device, user, or service is inherently trusted based on network location. [4] Complementary to ZTA, the NIST Cybersecurity Framework 2.0 released in 2024 [6] and the Cybersecurity Performance Goals [3] provide actionable implementation guidance for organizations initiating trust transformation programs in OT environments.

While ZTA has seen significant adoption in enterprise IT, its application to OT systems remains nascent, complicated by legacy device constraints, safety-critical availability

requirements, and the deterministic communication patterns that industrial control systems demand. This paper bridges this gap. The contributions of this work are as follows:

- An identity-centric trust model for smart factory environments, grounded in hardware roots of trust and cryptographic device attestation using TPM 2.0 and X.509 certificates issued via a PAA-PAI-DAC hierarchy aligned to the CSA Matter specification.
- A trust-aware data pipeline architecture that enforces cryptographic verification at every hop of telemetry ingestion, from field device to cloud analytics and digital twin state, with fail-closed semantics at every stage.
- A structured evaluation of four representative attack scenarios anchored to MITRE ATT&CK for ICS techniques, comparing system outcomes before and after ZTA enforcement.
- A safety-security co-design framework that reconciles continuous verification with the availability and timing constraints of industrial control and safety instrumented systems.
- A compliance alignment mapping between ZTA principles and IEC 62443 security levels and NIST CSF 2.0 functions, enabling adoption as a coherent extension of existing regulatory programs.

The remainder of this paper is organized as follows. Section II reviews background literature. Section III characterizes the expanded threat landscape of modern smart factories. Section IV analyzes the structural limitations of perimeter-centric OT security. Section V adapts Zero Trust principles to OT-specific constraints. Sections VI through XII detail the architectural components of the proposed framework. Section XIII presents the threat scenario evaluation. Section XIV provides compliance alignment. Section XV presents the reference architecture. Sections XVI and XVII offer discussion and conclusions.

2. Background and Related Work

2.1. Evolution of OT Security

Operational Technology security has historically relied on the physical isolation paradigm. Early industrial control systems employed proprietary fieldbus protocols PROFIBUS, Modbus, and DNP3 over dedicated hardware links largely inaccessible to adversaries without physical proximity. [29], [12] This approach proved effective in its original context but fundamentally incompatible with the operational demands of Industry 4.0. The Purdue Enterprise Reference Architecture (PERA), which defined hierarchical security zones between OT and IT systems, provides a useful conceptual model but was designed for unidirectional information flows and was not architected for the bidirectional, high-frequency data exchanges that digital twin systems and AI analytics engines require. [5] ICS-CERT reported sustained year-over-year increases in ICS vulnerability disclosures through 2022, directly reflecting the growing attack surface created by OT-IT convergence. [7]

2.2. Zero Trust Architecture

NIST SP 800-207 defines Zero Trust as a collection of principles that moves enterprise defenses from static, network-based perimeters to a focus on users, assets, and resources. [4] Its foundational tenets hold that all data sources and computing services are treated as resources, all communication is secured regardless of network location, access to individual resources is granted on a per-session basis, and access decisions are determined by dynamic, context-aware policy. The NIST Cybersecurity Framework 2.0, released in 2024, further embeds these principles within a broader enterprise risk governance structure applicable to critical infrastructure sectors including manufacturing. [6] Prior research has examined ZTA implementation in enterprise IT environments and cloud-native architectures. Mehta et al. proposed a lightweight ZTA for constrained IIoT devices but did not address digital twin pipeline integrity. [10] Huang et al. examined micro-segmentation in OT networks but did not incorporate continuous device attestation across the full asset lifecycle. [11] Gilchrist provided a practitioner-oriented analysis of Zero Trust applicability to IIoT environments, identifying device identity as the critical enabler. [15] Alagappan extended this analysis to connected physical systems spanning smart homes and industrial IoT, proposing a unified security blueprint whose enforcement model for heterogeneous device populations directly informs the identity-centric framework developed in this paper. [32]

2.3. Digital Twins in Manufacturing

Fuller et al. provided a comprehensive survey of digital twin enabling technologies, identifying data integrity and provenance verification as the most critical unsolved security challenges for industrial digital twins, noting that existing literature focuses overwhelmingly on twin fidelity and interoperability while treating security as an afterthought. [9] Alagappan proposed a platform-oriented reference architecture for digital twins in global R&D environments, identifying cross-domain data governance and authenticated telemetry ingestion as foundational requirements for operationally reliable twin systems requirements that the trust-aware pipeline in Section VIII directly implements. [35] Lu et al. described the connotation and reference model for digital twin-driven smart manufacturing, emphasizing real-time data synchronization between physical and virtual domains as the operational foundation of the technology. [23] Becue et al. explored the role of digital twins in supporting factory resilience and optimization, proposing a conceptual architecture that the present work extends with explicit security enforcement layers. [22] Wu et al. identified data authenticity and freshness as central requirements for reliable twin operation. [16] Bitton et al. demonstrated a methodology for constructing cost-effective security evaluation twins for ICS environments, validating the feasibility of the twin-as-security-instrument approach discussed in Section XV. [24]

2.4. Anomaly Detection and AI Security in ICS

Machine learning-based anomaly detection has been extensively studied as a complement to rule-based ICS security monitoring. Zolanvari et al. surveyed machine learning approaches for detecting cyber attacks targeting

controlled and monitored systems, establishing the empirical foundation for the behavioral baselining approach adopted in the trust-aware data pipeline. [17] Alagappan demonstrated the practical applicability of AI-driven anomaly detection in HVAC and water heating ICS deployments domains with highly periodic telemetry patterns and strong process-physics constraints providing domain-specific empirical grounding for the semantic anomaly detection layer described in Section VIII-C. [33] Iturbe et al. conducted a feasibility study on network-based ICS anomaly detection using autoencoders, demonstrating detection rates above 95% for known attack patterns while noting the challenge of maintaining low false-positive rates in environments with legitimate high process variability. [30] Ramotsoela et al. surveyed anomaly detection in industrial wireless sensor networks with a critical infrastructure focus, providing the technical foundation for the transport-layer behavioral monitoring described in Section VIII-B. [25]

2.5. Research Gap

The existing literature reveals two primary gaps that this paper directly addresses. First, Zero Trust implementations in OT environments have focused on network access control without addressing the end-to-end integrity of data pipelines from edge device to analytics platform. [10], [11], [15], [32] Second, the security of digital twin pipelines specifically, ensuring cryptographic provenance of telemetry data from field device through edge infrastructure to twin state has not been addressed within a unified Zero Trust framework anchored in hardware device identity. [9], [24], [35] Eckhart and Ekelhart identified per-data-element provenance tracking as a necessary but largely absent capability in existing industrial twin implementations. [14] This paper bridges both gaps through a trust-aware pipeline model grounded in hardware roots of trust and extending through every architectural layer to the analytics consumer.

3. Threat Landscape in Modern Smart Factories

Smart factories present an expanded and heterogeneous attack surface arising from the convergence of IT and OT domains. The attack surface is most usefully decomposed into four categories. The endpoint layer encompasses field sensors, actuators, PLCs, remote terminal units (RTUs), and edge gateways devices ranging from resource-rich industrial computers to severely constrained embedded controllers with minimal native security capabilities. The communication infrastructure layer includes industrial Ethernet, wireless protocols such as WirelessHART and ISA100.11a, MQTT brokers, OPC-UA servers, and internet-facing cloud APIs. The software and data layer encompasses firmware, SCADA applications, historian databases, and digital twin platforms. The human and organizational layer includes operators, maintenance personnel, remote vendors, and system integrators, each carrying distinct access privileges and attack surface contributions. [29]

Threat actors targeting smart factory environments range from nation-state actors pursuing strategic industrial espionage to financially motivated ransomware operators. Alladi et al. documented a structural shift toward targeted, multi-stage campaigns that exploit OT-IT integration pathways rather than attacking OT systems directly a pattern that the proposed ZTA directly addresses by hardening the OT-IT boundary. [12] Tedeschi et al. characterized inadequate device authentication and unencrypted communications as the dominant root causes of successful device-layer compromises across industrial deployments. [13] Miao et al. analyzed stealthy data injection attacks against cyber-physical systems, demonstrating that carefully constructed injection campaigns can evade statistical anomaly detectors while producing systematically erroneous operational decisions in downstream control systems. [27] Teixeira et al. extended this analysis to propose that detection and mitigation of stealthy attacks requires verification mechanisms that operate at the data origin at the sensor level rather than at intermediate network monitoring points, a principle embodied in the device-level telemetry signing requirement of the proposed architecture. [31]

Four threat categories are particularly relevant to the proposed architecture: unauthorized device access, data integrity attacks on telemetry pipelines and digital twin state, lateral movement through insufficiently segmented OT networks, and supply chain firmware compromise introduced prior to or during device deployment. These categories map directly to documented techniques in the MITRE ATT&CK for ICS framework, which provides a structured taxonomy of adversary tactics and techniques observed in real-world ICS attacks and is used throughout Section XIII to anchor the threat scenario evaluation. [39]

4. Limitations of Perimeter-Centric OT Security

The perimeter-centric security model embeds several structural assumptions that are invalidated by modern smart factory architectures. The first and most consequential is the implicit trust assumption: any entity that successfully authenticates to the network is granted broad access to resources within its zone. Conti et al. identified this as the single most consequential architectural deficiency in deployed IIoT security implementations, noting that lateral movement following initial network admission is the dominant attack progression pattern across documented IoT security incidents. [19] In flat OT networks common in brownfield facilities, a single compromised endpoint provides an adversary with network-level visibility into the entire production environment.

The second structural weakness is the use of network location as a proxy for device identity. Traditional OT security models rely on IP and MAC addresses as device identifiers, both of which are trivially spoofable. Pinto and Santos demonstrated that address-based device identification is bypassed by at least fourteen documented attack techniques available to commodity adversary toolsets. [18] This is particularly consequential in digital twin environments, where a small number of spoofed sources can corrupt the twin's state representation and cause downstream decision systems to operate on fundamentally unreliable data.

Third, perimeter models provide no visibility into intrazone east-west traffic. As OT devices increasingly adopt TLS-encrypted communications for inter-device messaging, security operations centers lose the ability to inspect intrazone traffic for behavioral anomalies without deploying dedicated decryption infrastructure. ICS-CERT data indicates that the average adversary dwell time in ICS-targeted attacks remains measured in months before detection. [7] ENISA's analysis of ICS and SCADA cybersecurity corroborates this finding, identifying intrazone lateral movement as the least-detected phase of documented ICS attack chains. [21] Finally, conventional zone-based segmentation as defined by IEC 62443 does not prevent a compromised device within a zone from attacking peer devices in the same zone a weakness that identity-centric micro-segmentation directly and architecturally addresses.

5. Zero Trust Principles Adapted for OT Environments

The application of Zero Trust to OT environments requires adaptation of NIST SP 800-207 principles to accommodate the unique constraints of industrial systems. [4] Three core principles govern the proposed architecture, each calibrated to the specific operational requirements of smart factory environments.

The principle of identity-as-the-perimeter replaces network location with cryptographically verifiable device and user identity as the primary trust signal. Every device, service, and user must present a verifiable identity credential before any access is granted. In OT environments, this extends the traditional user-centric identity model to encompass machine identities, with X.509 certificates issued by a private certificate authority as the primary credential format. Certificate lifecycle management is designed to operate transparently within real-time control environments, avoiding renewal events that could disrupt time-sensitive communications.

The principle of continuous verification mandates that authentication and authorization are not one-time events at session establishment but are evaluated throughout the communication lifecycle. In OT contexts, this must be implemented efficiently to avoid introducing latency that would violate the timing constraints of real-time control systems. Lightweight challenge-response protocols and hardware-assisted remote attestation mechanisms minimize computational overhead, enabling enforcement without compromising the deterministic behavior industrial systems require.

The principle of least-privilege access constrains each device, user, and service to the minimum set of resources and operations required to fulfill its function. In smart factory environments, this translates to granular access policies that restrict PLCs to specific process variable sets, limit vendor remote access to designated maintenance windows and authorized device categories, and enforce read-only constraints on digital twin state data for analytics consumers. Dynamic policy evaluation informed by contextual signals including device health metrics, anomaly scores, and operational state flags enables privilege levels to be adjusted in real time without requiring manual operator intervention, closing the window between threat detection and policy response.

6. Identity-Centric Trust Model

6.1. Hardware Root of Trust

The foundation of the proposed identity model is a hardware root of trust (HRoT) implemented through Trusted Platform Module (TPM 2.0) chips or equivalent secure elements embedded in OT devices. The TPM provides a tamper-resistant environment for cryptographic key storage, platform integrity measurement through the Platform Configuration Registers (PCRs), and remote attestation the ability to cryptographically prove to a remote verifier that a device is running unmodified, authorized firmware. The private keys bound to device identity certificates never leave the secure element in plaintext, ensuring that certificate-based identity cannot be extracted and replicated by an adversary with transient access to the device. [4], [5] For constrained devices that cannot accommodate a full TPM, physically unclonable functions (PUFs) offer a lower-cost alternative for hardware-bound identity derivation, exploiting manufacturing-process variations to produce device-unique fingerprints that cannot be cloned or predicted from external observation.

The cryptographic algorithms underpinning device identity must themselves be treated as a lifecycle-managed component rather than a static architectural choice. NIST finalized three post-quantum cryptography (PQC) standards in August 2024: FIPS 203 (ML-KEM, a lattice-based key encapsulation mechanism), FIPS 204 (ML-DSA, a lattice-based digital signature algorithm), and FIPS 205 (SLH-DSA, a stateless hash-based signature standard). [36], [37], [38] Smart factory device fleets with operational lifetimes extending ten to twenty years will remain in service through the period in which cryptographically relevant quantum computers are expected to become feasible. The proposed architecture therefore requires cryptographic agility the capacity to negotiate, update, and migrate cryptographic algorithm suites without requiring device firmware replacement as a core design requirement of the Identity Hub and the secure element provisioning workflow. Certificate templates must support hybrid classical-PQC schemes during the migration period, and the PKI must be capable of issuing and validating both ECDSA and ML-DSA signed certificates concurrently to accommodate device fleets at different lifecycle stages of PQC migration.

6.2. Device Identity Lifecycle

Device identity must be established and maintained across the complete device lifecycle, from manufacturing through operational deployment and eventual decommissioning. During manufacturing, a unique Device Identity Certificate (DevID) is provisioned into the device's secure element, signed by the manufacturer's certificate authority in accordance with IEEE 802.1AR. Upon deployment, the DevID authenticates the device to the

factory's Identity Hub, which issues an operational certificate binding the device's cryptographic identity to its assigned operational role and access policy set. Certificate rotation is managed automatically by the Identity Management Service, with certificate lifetimes calibrated to balance security against operational continuity. [5] The NIST SP 800-82 Rev. 3 guidance provides additional lifecycle management requirements including provisions for certificate management in low-connectivity offline scenarios that inform the renewal and revocation design described herein.

The Connectivity Standards Alliance (CSA) Matter specification provides the most production-deployed implementation of a hierarchical device attestation certificate chain applicable to IIoT environments. The Matter attestation model defines a three-tier hierarchy: Product Attestation Authority (PAA) as the root CA, Product Attestation Intermediate (PAI) as the manufacturer-level intermediate, and Device Attestation Certificate (DAC) as the per-device leaf certificate bound to the device's unique private key. [40] The proposed architecture adopts this PAA-PAI-DAC chain as the normative model for device attestation certificate provisioning, extending it to the full operational lifecycle by layering operational certificates above the DAC tier for role-based access policy binding. This alignment enables device identity to be verified by any relying party that trusts the PAA root, supporting cross-organizational and supply-chain identity federation without requiring bilateral PKI trust agreements between each factory operator and device manufacturer. Upon decommissioning, the device's certificates are revoked in the PKI, its identity is removed from the Identity Hub registry, and key material in the secure element is cryptographically destroyed, preventing reactivation of decommissioned credentials.

Offline certificate revocation presents a distinct design challenge in OT environments where devices may operate in scheduled-maintenance offline states, intermittent-connectivity edge segments, or air-gapped sub-networks that cannot reach the PKI's Online Certificate Status Protocol (OCSP) responder in real time. The proposed architecture addresses this through a three-mechanism defense in depth. First, edge gateways pre-fetch and locally cache Certificate Revocation Lists (CRLs) from the PKI on a configurable schedule recommended maximum interval of four hours for operational certificates, one hour for certificates bound to safety-critical devices ensuring that revocation information is available during transient network partitions without requiring a live OCSP query per authentication event. Second, OCSP stapling is required for all mTLS handshakes where connectivity permits: the device or gateway presents a pre-fetched, CA-signed OCSP response bound to the certificate in question, eliminating the per-handshake OCSP round-trip latency that would violate real-time control timing constraints. Third, when a gateway cannot refresh its CRL within a configurable staleness window recommended maximum of twenty-four hours it enters a conservative-mode policy: it continues to authenticate devices whose certificates were valid at the most recent successful CRL fetch, but flags all such sessions in the audit log and generates a PKI-staleness alert to the SIEM. Newly presented certificates whose revocation status cannot be confirmed against a fresh CRL are rejected until connectivity is restored. This design ensures that revocation enforcement degrades gracefully rather than failing open, while preserving operational continuity for the device fleet that was in good standing at the last confirmed revocation check.

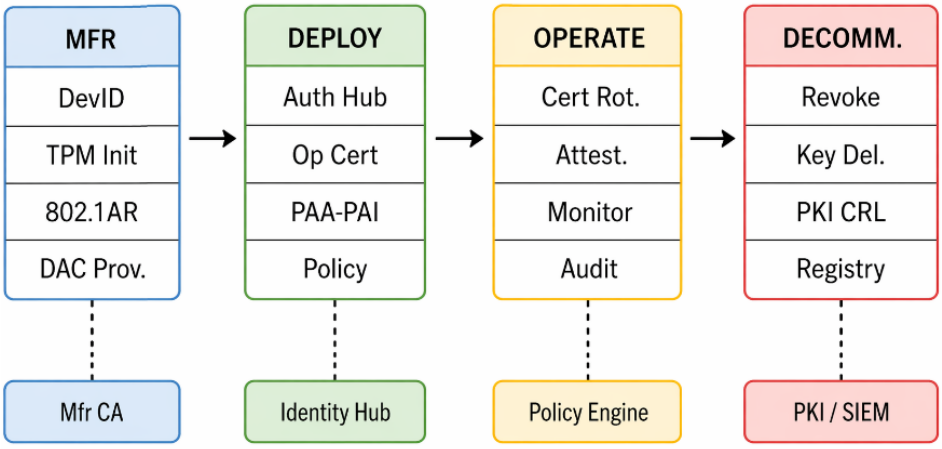

Fig 1: Device Identity Lifecycle

From hardware-bound DAC provisioning at manufacture through operational certificate management to cryptographically verified decommissioning. The PAA-PAI-DAC hierarchy aligns with the CSA Matter attestation model [40].

6.3. Identity Hub as Single Source of Truth

A centralized Identity Hub serves as the authoritative registry for all device, user, and service identities within the smart factory ecosystem. The Identity Hub maintains identity metadata including device type, firmware version hash, operational role, authorization policy bindings, and certificate revocation status. It issues short-lived access credentials to authenticated entities and maintains an immutable audit trail of all identity assertions and access decisions. By consolidating identity management in a single authoritative service, the Identity Hub eliminates the identity fragmentation multiple siloed credential stores with inconsistent policies and

unsynchronized revocation lists that characterizes many brownfield OT environments, and enables coherent, enterprise-wide policy enforcement across OT, IT, and cloud domains. The Identity Hub is architected for high availability with active-active geographic redundancy across at least two independent availability zones; in the event that an edge gateway loses connectivity to all Hub replicas, the gateway continues to enforce the most recently downloaded policy snapshot for a configurable continuity window recommended maximum of sixty minutes after which it restricts all new session establishments to previously authenticated devices only and generates a Hub-unreachable alert to the SIEM, ensuring that a Hub outage cannot be exploited to admit unauthenticated devices by simply severing Hub connectivity.

7. Digital Twin Security

7.1. Threat Model for Digital Twin Pipelines

Digital twins are high-value targets precisely because their outputs directly and continuously influence operational decision-making across production, maintenance, and quality domains. [9], [16] The threat model for digital twin pipelines encompasses three primary attack categories. Data poisoning involves the injection of falsified telemetry into the twin's ingestion pipeline, corrupting the twin's state representation incrementally and causing downstream decision systems to operate on systematically inaccurate data a particularly insidious attack because the drift may not be observable until a downstream decision produces a visible operational failure. Miao et al. formally characterized the conditions under which stealthy injection campaigns evade conventional anomaly detectors while producing systematically erroneous operational decisions. [27] State manipulation involves direct modification of the twin's data store or computational model, bypassing the telemetry ingestion pipeline entirely a vector available to adversaries who have obtained sufficient API credentials through an IT-to-OT pivot. Inference corruption targets the AI and analytics models consuming twin outputs, either through training data poisoning that degrades model performance over successive training cycles or by crafting adversarial inputs at inference time. [17]

7.2. Integrity and Provenance Requirements

Digital twin integrity requires that every data element contributing to the twin's state be traceable to a verified source device, carry a cryptographic signature attesting to its authenticity and bit-level integrity, and bear a timestamp issued by a trusted time authority to prevent replay attacks. Eckhart and Ekelhart identified per-data-element provenance tracking as a necessary but largely absent capability in existing industrial twin implementations, motivating the structured provenance metadata design described in Section VIII. [14] These requirements extend the Zero Trust principle of continuous verification from network access control to data provenance: the operative security question shifts from "who is requesting access to this resource?" to "where did this data originate, has its integrity been preserved across every pipeline hop, and has it been observed before?" Provenance metadata is maintained alongside each telemetry record in the twin's data store, enabling full audit of the data lineage supporting any operational decision and supporting forensic reconstruction of the data state at any prior point in time.

8. Trust-Aware Data Pipeline

The trust-aware data pipeline enforces cryptographic verification at three architectural layers, ensuring that telemetry flows from field device to cloud analytics without any unverified segment.

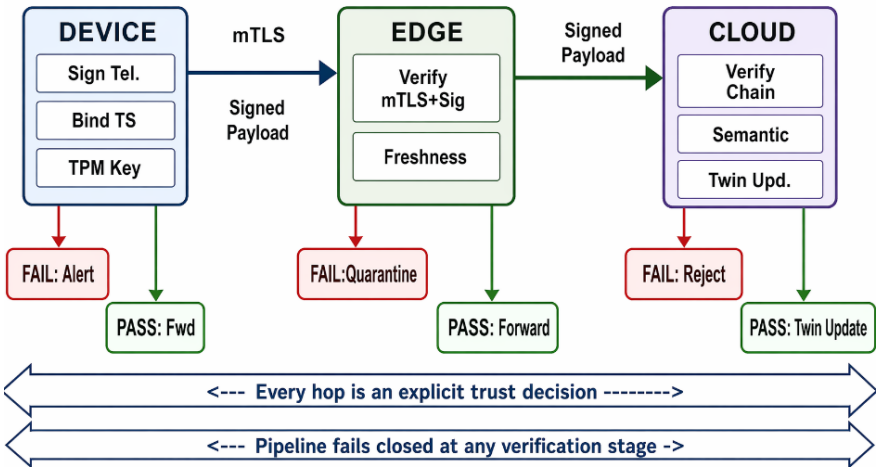

Fig 2: Trust-Aware Data Pipeline

Cryptographic verification enforced at device, edge, and cloud layers. Failure at any stage results in data quarantine, not permissive forwarding.

The governing design principle is that each hop in the data pipeline constitutes an explicit trust decision: no data is permitted to advance to the next stage without verification of the prior stage's output. The pipeline is structured to fail closed any verification failure results in quarantine of the affected data, not permissive forwarding with a reduced confidence annotation.

8.1. Edge Layer

Each field device signs its telemetry payloads using its private key, resident in the device's hardware secure element. The signature covers the payload content, the device

identifier, and a timestamp sourced from a hardware clock synchronized to a trusted time server via NTP with authentication or IEEE 1588 PTP over a secured channel. This binding prevents both content modification in transit and replay attacks using previously observed payloads. Edge gateways perform the first verification step: they authenticate the originating device via mutual TLS (mTLS), verify each telemetry payload's digital signature, and forward only cryptographically verified payloads to the transport layer. [29] Payloads that fail verification originating from unknown devices, carrying invalid or expired signatures, or presenting timestamps outside an acceptable freshness window are quarantined and flagged for operator review, with each rejection event logged immutably to the central audit trail.

Clock synchronization constitutes an independent trust boundary within the edge layer and must be designed explicitly rather than assumed. The proposed architecture specifies a maximum acceptable clock skew threshold of ±2 seconds between a field device and the edge gateway's reference clock, consistent with the IEEE 1588 PTP accuracy class recommended for industrial automation networks. Payloads whose timestamps fall outside this window are treated as potential replay attempts and are rejected regardless of the validity of their cryptographic signature. When a field device loses NTP or PTP synchronization due to network partition, GPS signal loss, or deliberate jamming it enters a degraded-mode operation: it continues to sign and transmit telemetry using its last-known synchronized clock state, but appends a synchronization-loss flag in the payload header. The edge gateway, upon receiving a synchronization-loss-flagged payload, accepts it for forwarding to the SIEM anomaly queue but withholds it from the digital twin ingestion pipeline until the device resynchronizes and the intervening payload window has been audited. This design preserves field device observability during clock disruption events while enforcing the provenance integrity guarantees of the twin pipeline.

8.2. Transport Layer

All inter-layer communication uses mutual TLS with certificate pinning to prevent man-in-the-middle attacks even in environments where PKI trust anchors may be compromised. Message queuing brokers MQTT or AMQP implementations enforce topic-level access control, ensuring that devices can publish only to topics authorized by their policy binding in the Identity Hub. Transport-layer anomaly detection, operating on statistical baselines of message frequency, payload size distribution, and inter-message timing patterns, identifies deviations that may indicate device compromise, unauthorized firmware modification, or replay injection campaigns. [25], [30], [33] The anomaly detection approach draws on domain-validated techniques for ICS environments with highly periodic telemetry patterns and strong process-physics constraints, where physics-informed distributional baselines enable high-confidence deviation detection with low false-positive rates. [33] Xiao et al. demonstrated the applicability of ensemble anomaly detection methods to IoT telemetry streams, informing the multi-algorithm baselining approach implemented at this layer. [26] Anomalous traffic patterns trigger automatic investigation workflows and invoke dynamic policy restrictions on the affected device while the anomaly is assessed by the security operations team.

8.3. Cloud and Analytics Layer

At the cloud layer, ingestion services verify the complete chain of signatures accompanying each telemetry payload, confirming provenance from the originating field device through the edge gateway. AI and analytics models consuming digital twin outputs receive only provenance-attested data; inputs lacking a verified provenance chain are rejected rather than processed with reduced confidence. Anomaly detection at this layer operates on semantic content identifying physically implausible sensor values, correlated anomalies across multiple geographically proximate devices that may indicate coordinated injection, and deviations from process physics models that cannot be explained by legitimate process variation. [17], [30] Model outputs generated from inputs that fail any semantic integrity check are flagged as low-confidence and presented to human operators with explicit uncertainty indicators rather than being silently forwarded to automated actuation systems.

9. Intent-Based Access Control

Access control in the proposed architecture moves beyond static role-based access control (RBAC) to incorporate intent the declared and contextually verifiable purpose of an access request as an additional factor in policy evaluation. A maintenance engineer requesting write access to a PLC setpoint register during a scheduled maintenance window, with a corroborating open work order in the computerized maintenance management system (CMMS), presents a fundamentally different risk profile than the same engineer requesting identical access during an unscheduled off-hours session with no corresponding work order. Intent-based access control evaluates the full context of each request: the requesting entity's cryptographic identity and assigned role; the specific resource and operation requested; the operational mode of the target system (production, maintenance, commissioning, or emergency); and external risk signals including device anomaly scores, active security alerts, geolocation, and time-of-day factors.

Policy Enforcement Points (PEPs) are deployed at all resource boundaries edge gateways, OT network segment ingress points, cloud service API gateways, and digital twin data store interfaces. PEPs forward each access request to a centralized Policy Decision Point (PDP) that evaluates the request against policies maintained in the policy store and enriched in real time with contextual signals from the SIEM platform.

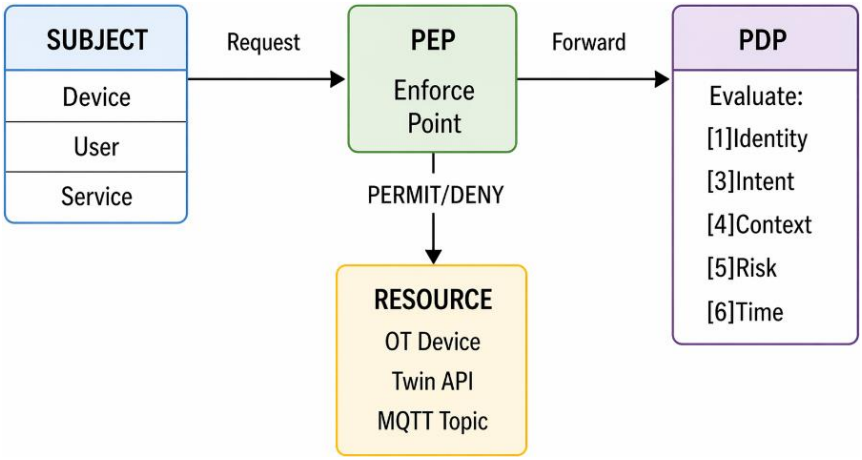

Fig 3: Intent-Based Access Control Flow

PEP forwards each request to PDP, which evaluates identity, declared intent, and five real-time context signals before issuing a PERMIT or DENY decision logged to the audit trail.

Least-privilege enforcement is dynamic: access permissions are automatically reduced when a device's anomaly score exceeds a configurable threshold or when its communication patterns deviate materially from the established behavioral baseline, with policy restoration occurring upon verification that normal operation has resumed. This dynamic privilege adjustment closes the gap between threat detection and policy response that static RBAC systems cannot address.

10. Brownfield Deployment Strategy

The deployment of Zero Trust in brownfield OT environments facilities containing legacy devices that cannot be replaced or upgraded within operational timescales requires a pragmatic, incremental approach. Many existing OT devices lack the computational resources to implement TPM-based attestation or mutual TLS, and many communicate using legacy fieldbus protocols Modbus RTU, DNP3, PROFIBUS DP that do not natively support authentication or encryption. [5], [12] Requiring immediate full-stack Zero Trust compliance across these devices would make deployment impractical in all but greenfield facilities. CISA's recommended cybersecurity practices for industrial control systems explicitly endorse an incremental hardening approach for brownfield environments, prioritizing boundary controls and monitoring visibility before attempting device-level cryptographic enforcement. [28]

Identity-aware proxy gateways are positioned at the boundary between legacy OT segments and the Zero Trust enforcement infrastructure. These gateways terminate legacy protocol connections from field devices, authenticate on behalf of collections of legacy endpoints using gateway-level certificates, and translate traffic into authenticated, encrypted sessions for upstream consumption. While the proxy model does not achieve per-device cryptographic identity for legacy endpoints, it immediately and substantially strengthens segment-boundary security without requiring field device replacement. The gateway's certificate represents a security domain rather than an individual device, and the gateway enforces protocol-level controls rate limiting, command whitelisting, and payload inspection within its segment.

Progressive micro-segmentation is applied iteratively across the facility, with segments disaggregated based on device function, operational criticality, and risk classification. Higher-risk segments those controlling safety functions or directly feeding digital twin pipelines receive prioritized Zero Trust enforcement investment. Network traffic analysis tools establish behavioral baselines for each segment, enabling anomaly detection even for segments that have not yet achieved device-level cryptographic identity.

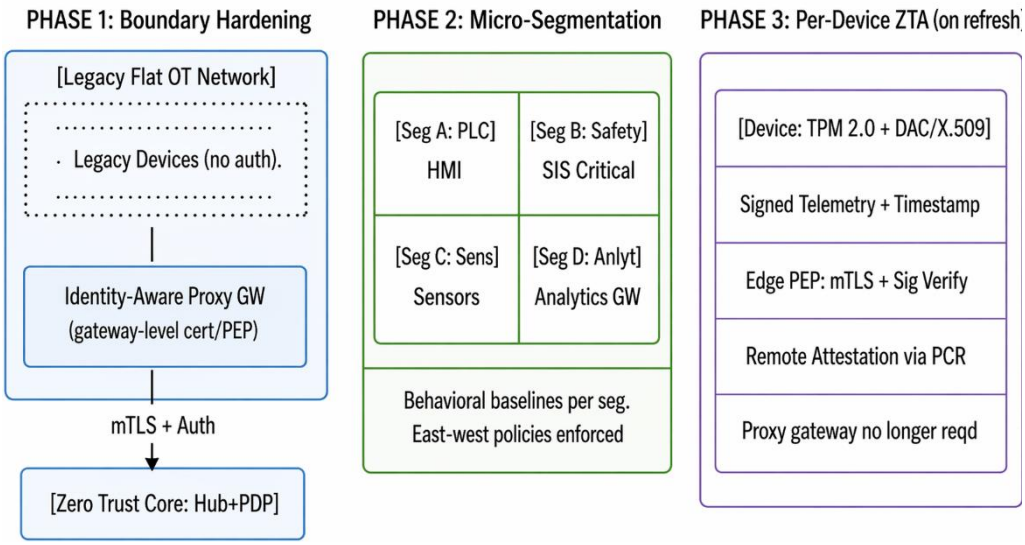

Fig 4: Brownfield Deployment Phases

Phase 1 deploys identity-aware proxy gateways at OT segment boundaries; Phase 2 enforces micro-segmentation with per-segment behavioral baselines; Phase 3 achieves per-device Zero Trust as hardware is refreshed on normal lifecycle schedules.

11. Securing AI and Analytics Pipelines

AI-driven systems in smart factories including predictive maintenance models, quality control classifiers, and production optimization engines are fundamentally dependent on the integrity of their input data. An adversary capable of manipulating the data pipeline feeding these models can degrade their performance over time through data poisoning, leading to incorrect maintenance predictions, undetected quality defects, or suboptimal process parameters that progressively erode production yield without generating immediately observable anomalies. [17], [27] Zolanvari et al. demonstrated that machine learning models trained on ICS telemetry are particularly susceptible to gradual data poisoning attacks, where small, individually imperceptible perturbations to training data accumulate to produce systematic model degradation over multiple training cycles. [17]

The proposed architecture secures AI pipelines through three complementary mechanisms. First, all training data is sourced exclusively from the verified telemetry pipeline, ensuring that only provenance-attested, cryptographically integrity-verified data is used for model training and fine-tuning cycles. Any data whose provenance chain cannot be verified is excluded from training datasets regardless of its apparent plausibility. Second, model integrity verification is applied to deployed AI models: models are cryptographically signed upon deployment, and signature verification is performed at each inference invocation, preventing unauthorized model substitution without disrupting inference service operation. Third, continuous monitoring of model output distributions across time identifies statistical shifts that may indicate data poisoning where the model performs as designed but on corrupted inputs or model degradation. [17], [30], [33] Domain-specific validation of AI anomaly detection for HVAC and water heating IoT deployments confirms that process-physics-informed distributional baselines substantially outperform purely statistical approaches in distinguishing legitimate process variation from injected anomalies. [33] Conti et al. identified model output monitoring as an underutilized defense layer in IoT deployments, noting that production models are rarely subjected to ongoing distributional integrity verification after initial deployment a gap the proposed architecture directly addresses. [19]

12. Safety-Security Co-Design

OT environments impose constraints that are largely absent from enterprise IT security design. Industrial control systems prioritize availability and determinism: a security control that introduces variable latency or service interruption can have consequences ranging from production losses to personnel safety incidents. IEC 62443 explicitly recognizes this tension, requiring that security measures not reduce the Safety Integrity Level (SIL) of safety instrumented systems. [8] Teixeira et al. further formalized the safety-security tension in industrial systems, proposing co-design frameworks that treat security and functional safety as joint optimization objectives rather than sequential design concerns an approach that directly informs the three principles below. [31]

Security controls are implemented as fail-open at safety-critical system boundaries, where a security system failure must not interrupt the operation of safety instrumented systems (SIS). At non-safety-critical boundaries, fail-closed behavior is the default, consistent with Zero Trust principles. The fail-mode classification for each boundary is determined during system design and validated during commissioning, and is encoded as a system requirement rather than a deployment-time configuration choice to prevent inadvertent reconfiguration.

Time-bounded verification is enforced for control-plane communications. Verification decisions that cannot be completed within the real-time deadline of the requesting control loop are resolved using the most recent known-good policy state, with a security alert generated for operator review. This approach ensures that security enforcement does not become a source of non-determinism in time-critical control paths while maintaining a complete record of all instances where deadline-forced policy resolution occurs.

Human-in-the-loop authorization is required for operations with irreversible physical consequences firmware updates to safety-critical devices, setpoint changes exceeding normal operating ranges, and decommissioning of devices whose function cannot be immediately assumed by a backup system. This requirement is encoded as a mandatory policy step that cannot be overridden by automated policy engines, preserving operator situational awareness and accountability for high-consequence actions. ENISA's guidance on cybersecurity for industrial automation and control systems identifies human oversight for high-consequence irreversible actions as a compensating control of particular value in environments where automated policy enforcement cannot achieve full coverage due to legacy device constraints. [20], [21]

13. Evaluation: Threat Scenario Analysis

To evaluate the efficacy of the proposed Zero Trust architecture, this section analyzes four representative attack scenarios drawn from documented ICS threat intelligence and published incident analyses. For each scenario, outcomes are compared between a conventional perimeter-protected OT environment and the proposed ZTA implementation. Each scenario is mapped to its corresponding MITRE ATT&CK for ICS technique, providing a taxonomy-grounded and reproducible basis for evaluation. [1], [7], [12], [39]

13.1. Scenario 1: Rogue Device Injection (MITRE T0856)

Technique T0856 (Spoof Reporting Message): An adversary gains brief physical access to an unmanaged network switch in a maintenance cabinet and connects an unauthorized device configured to masquerade as a legitimate temperature sensor. The adversary's objective is to inject falsified high-temperature readings into the digital twin pipeline, causing the predictive cooling-system maintenance model to misallocate maintenance resources and suppress legitimate anomaly alerts.

In a conventional perimeter-protected environment, the rogue device may successfully join the network if no port-level access control is enforced. If the network uses IP or MAC-based device identification, the adversary can clone the address of a legitimate sensor and begin publishing falsified readings. The digital twin ingests these readings without distinguishing them from legitimate telemetry. Tedeschi et al. documented analogous device impersonation attacks against deployed industrial sensor networks, noting that the absence of cryptographic device authentication is a prerequisite for this attack class. [13] The injection may persist for weeks or months before anomalous operational outcomes trigger human investigation.

Under the proposed ZTA, the rogue device cannot present a valid X.509 DAC issued by the factory PKI, as the corresponding private key is resident in the legitimate device's hardware secure element and cannot be extracted. The identity-aware edge gateway performs a mutual TLS handshake upon the device's first communication attempt, rejects the connection, generates a security alert, and logs the event with the rogue device's observed network identifiers. The digital twin pipeline receives only signature-verified telemetry, and the attempted injection is both blocked at the enforcement boundary and recorded in the immutable audit trail for forensic analysis.

13.2. Scenario 2: Digital Twin State Manipulation via IT Pivot (MITRE T0859)

Technique T0859 (Valid Accounts): An adversary compromises a cloud-hosted IT system through a successful phishing campaign targeting a corporate finance employee. Using the compromised account's cloud access credentials, the adversary attempts to pivot into the digital twin data store and modify the twin's state to suppress high-temperature readings that would normally trigger a safety shutdown of a high-pressure extrusion process.

In a perimeter-protected environment, broad access permissions granted to cloud IT services may extend to the digital twin platform, as both reside within the enterprise cloud subscription perimeter. The compromised IT account may be able to query and modify the digital twin data store directly through a legitimate API. Alladi et al. specifically documented IT-to-OT pivot as the dominant multi-stage attack progression in ICS incidents from 2017 onward. [12]

Under the proposed ZTA, the digital twin data store is a protected resource governed by intent-based access control policies. API calls from the compromised cloud IT account are evaluated by the PDP, which enforces that a finance-role account possesses no policy binding permitting write access to digital twin state data a permission reserved exclusively for the authenticated telemetry ingestion service identity. The SIEM simultaneously detects the anomalous API access pattern, reduces the account's access scope dynamically, and generates a priority security incident for analyst response. The state modification attempt is blocked and recorded, and the process continues operating on verified telemetry.

13.3. Scenario 3: Lateral Movement via Compromised HMI (MITRE T0867)

Technique T0867 (Lateral Tool Transfer): An adversary exploits a known, unpatched vulnerability in the embedded web server of a Human-Machine Interface (HMI) terminal to obtain remote code execution. From this initial foothold, the adversary conducts reconnaissance of the OT network and attempts lateral movement toward a Safety Instrumented System (SIS) controlling high-pressure process conditions, with the objective of modifying the SIS's safety trip setpoints. In a flat OT network, the compromised HMI has direct

network reachability to PLCs, engineering workstations, historian servers, and SIS components on the same VLAN. ICS-CERT data indicates that adversary dwell times following HMI compromise average over ninety days before lateral movement toward high-value targets is detected in conventionally segmented environments. [7]

Under the proposed ZTA, micro-segmentation enforces strict, identity-based communication policies between operational segments. The compromised HMI's anomalous behavior port scanning, repeated connection attempts to endpoints outside its authorized communication policy is detected by the behavioral baseline system within minutes of initiation. [25], [30] An automatic quarantine policy is applied to the HMI's segment, restricting its outbound communications to the minimum set required for operator display functionality. The SIS segment enforces a whitelist-based communication policy admitting only the specific engineering workstation and historian service identities authorized by the operational policy. Lateral movement to the SIS is blocked regardless of the HMI's network position, the HMI is isolated for forensic investigation, and the incident is immediately escalated to the security operations team with full session telemetry.

13.4. Scenario 4: Supply Chain Firmware Compromise (MITRE T0862)

Technique T0862 (Supply Chain Compromise): An adversary compromises the software build pipeline of a PLC firmware vendor, inserting malicious code into an update package subsequently distributed to customers through the vendor's standard update portal. The compromised firmware includes a covert command channel that enables the adversary to issue unauthorized control commands and exfiltrate process data using encrypted communications that blend with legitimate traffic profiles. In environments without cryptographic firmware verification, the compromised update package may be applied without detection. [12], [28]

Under the proposed ZTA, all firmware update packages are required to carry a valid code-signing certificate from the device manufacturer's certificate authority, with the certificate issued by a CA whose root is registered in the Identity Hub's trust store. The OTA update service verifies this signature before staging any update for device deployment. Post-update remote attestation then verifies that the device's TPM PCR measurements match the expected values for the authorized firmware version. A device running the compromised firmware produces PCR measurements that do not correspond to the signed, authorized firmware baseline. The attestation failure triggers automatic quarantine of the affected device, suppression of its telemetry contributions to the digital twin pipeline, and generation of a critical security alert identifying the firmware version, update timestamp, and device identity for incident response. The covert channel is prevented from establishing outbound communications because the device's network segment policy does not authorize the novel outbound connection patterns the backdoor requires.

Table 1: Threat Scenario Analysis Pre- and Post-ZTA Outcomes with MITRE ATT&CK For ICS Mapping

Scenario	MITRE ICS	Attack Vector	Pre-ZTA Outcome	Post-ZTA Outcome
Rogue Device Injection	T0856	Physical access; IP/MAC clone	Telemetry injection succeeds; twin state corrupted; anomaly undetected for weeks	Rejected at mTLS; alert generated; injection logged in audit trail
Digital Twin Manipulation	T0859	Phishing; IT-to-OT cloud API pivot	State modification via API succeeds; safety alerts suppressed	PDP denies access; SIEM alert; account dynamically restricted; incident opened
Lateral Movement via HMI	T0867	HMI RCE exploit; OT network scan	SIS reachable; lengthy dwell undetected; setpoint modification possible	HMI quarantined on scan detection; SIS whitelist blocks connection; escalated

Supply Chain Firmware	T0862	Compromised OTA update package	Backdoor installed silently; persistent covert channel established	Code-sign verification fails; attestation mismatch; device quarantined; alert

14. Compliance Alignment with IEC 62443 and NIST CSF 2.0

The proposed architecture is explicitly designed to align with and extend the IEC 62443 industrial security standard and NIST CSF 2.0, enabling organizations with existing compliance programs to adopt this framework as a coherent extension rather than a competing model. [4], [6], [8] Table II maps the core Zero Trust mechanisms to their corresponding IEC 62443 foundational requirements and security levels. Security levels SL-2 and SL-3, which address protection against intentional violations using moderate and sophisticated means respectively, are the primary alignment targets, reflecting the threat actor sophistication representative of advanced persistent threats targeting industrial facilities. The NIST CSF 2.0 Govern, Identify, Protect, Detect, Respond, and Recover function mapping is noted in the compliance tier column to support organizations pursuing dual-framework compliance programs. Alagappan's governance model for IoT data in global manufacturing environments provides a complementary policy-layer framework for managing data sovereignty, retention, and access rights across the multi-jurisdictional deployments that large manufacturing enterprises operate governance dimensions that the proposed ZTA architecture enforces at the technical enforcement layer but does not fully specify at the policy content level. [34]

Table 2: Zero Trust Architecture to IEC 62443 and NIST CSF 2.0 Compliance Mapping

ZTA Concept	Implementation Mechanism	IEC 62443 Reqt.	SL / CSF 2.0
Micro-segmentation	Identity-aware PEPs at segment boundaries	Zones & Conduits (SR 5.1)	SL-2/3 | PR.AC
Least Privilege	Intent-based access via PDP/PEP	Account Mgmt (SR 1.2)	SL-2 | PR.AC
Device Identity	X.509 DAC; TPM 2.0; PAA-PAI-DAC	Identifier Mgmt (SR 1.1)	SL-2/3 | ID.AM
Continuous Monitoring	SIEM; behavioral baselines; anomaly det.	Audit Log Mgmt (SR 6.1)	SL-2 | DE.CM
Secure Communications	mTLS; cert pinning; signed telemetry	Use Control (SR 2.1)	SL-2/3 | PR.DS
Remote Attestation	TPM PCR measurement; fw hash verify	SW Integrity (SR 3.4)	SL-3 | PR.IP
PQC Agility	Hybrid ECDSA+ML-DSA; FIPS 203/204/205	Cryptographic Strength (SR 4.3)	SL-3 | PR.DS
Safety-Sec Codesign	Fail-open at SIS; time-bounded verify	Safety Reqts (SR 4.2)	SL-2/3 | RS.MI

15. Reference Architecture

The reference architecture organizes the proposed Zero Trust framework across six hierarchical layers, each enforcing trust at its level and consuming trust assertions from the layers below. The layered structure reflects the physical and logical organization of the smart factory environment while ensuring that no data or access path bypasses the identity and policy enforcement infrastructure. The architecture extends and operationalizes the conceptual digital twin security framework proposed by Becue et al. to include explicit cryptographic

enforcement layers at each tier. [22] The digital twin platform layer draws on the platform-oriented reference model proposed by Alagappan for global R&D environments, adapting its multi-domain data ingestion and twin lifecycle management patterns to include the provenance verification and access control enforcement layers required by the Zero Trust model. [35]

The Device Layer constitutes the root of trust for the entire architecture. Field devices hold hardware-bound identity credentials in their secure elements and sign all outgoing telemetry with their device private keys, establishing the cryptographic provenance of operational data at its origin point before any network transmission, so that provenance cannot be retroactively asserted or fabricated at an upstream stage. The Edge Layer encompasses identity-aware edge gateways that authenticate devices via mTLS, verify telemetry payload signatures, enforce local access policy, and translate legacy protocol traffic into authenticated, encrypted sessions. Edge gateways serve as the primary enforcement frontier of the Zero Trust model: no unverified data advances beyond this layer into the higher-trust infrastructure.

The Identity Hub and Control Plane form the policy authority for the entire system. The Identity Hub maintains the device and user identity registry, manages certificate lifecycle across the device fleet, evaluates access requests through the Policy Decision Point, and distributes policy updates to PEPs deployed throughout the infrastructure. This layer is architected for high availability with geographic redundancy. The Digital Twin Platform ingests provenance-attested telemetry and maintains the real-time state representation of all physical assets. Access to the twin's data store is governed by intent-based policies enforced by the PEP embedded in the platform's API gateway.

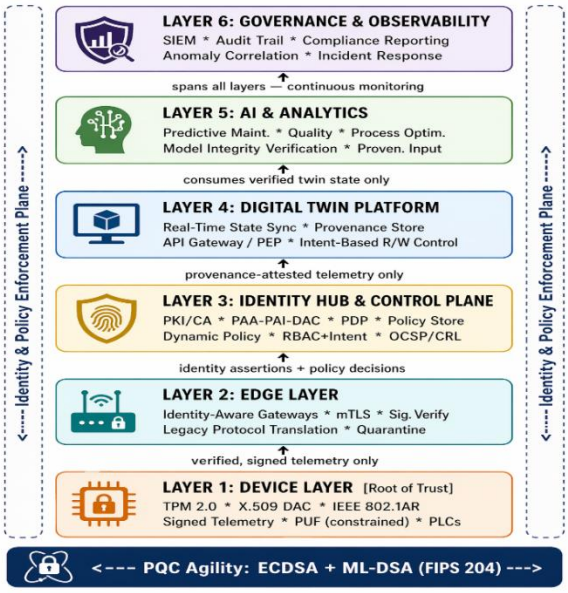

Fig 5: Zero Trust Smart Factory Reference Architecture

six-layer model with identity and policy enforcement spanning all tiers. The PAA-PAI-DAC attestation hierarchy and post-quantum cryptographic agility are design requirements at the device and identity hub layers. No data or access path bypasses the enforcement infrastructure. The AI and Analytics Layer consumes verified twin outputs, with model integrity verification applied at each inference invocation. The Governance and Observability Layer spans all other layers, providing centralized audit logging, SIEM integration, anomaly detection correlation across all enforcement points, and compliance reporting.

16. Discussion

The proposed architecture demonstrates that Zero Trust is not only applicable to enterprise IT environments but is architecturally necessary for OT systems undergoing digital transformation. The threat scenario analysis illustrates that the layered enforcement model anchored in hardware device identity and extending through the complete telemetry pipeline to the digital twin and AI analytics layers addresses attack vectors that perimeter-centric models cannot detect or contain. The structured MITRE ATT&CK for ICS mapping confirms that the four scenarios represent documented, real-world technique classes rather than theoretical constructs, and that the ZTA countermeasures operate precisely at the mechanism level that each technique exploits. [10], [11], [15], [39]

Several practical implementation considerations merit discussion. The requirement for hardware secure elements and a PAA-PAI-DAC certificate chain introduces per-unit provisioning costs that may be significant for large sensor deployments. Organizations should quantify this cost against the operational risk exposure of unverified telemetry in digital twin pipelines, where a single sustained injection attack can produce maintenance and quality decisions worth substantially more in operational loss than the fleet-wide cost of TPM and DAC provisioning. [24] The PKI infrastructure required to support certificate lifecycle management across potentially thousands of devices also introduces operational complexity. Offline certificate revocation in OT environments where devices operate in scheduled-maintenance offline states

or intermittent-connectivity edge segments is a well-recognized implementation challenge. The proposed architecture addresses this directly in Section VI-B through a three-mechanism design: locally cached CRLs pre-fetched at four-hour intervals (one-hour for safety-critical devices), mandatory OCSP stapling where connectivity permits, and a conservative-mode policy when CRL staleness exceeds twenty-four hours that rejects newly presented certificates while preserving continuity for devices whose certificates were valid at the last confirmed revocation check. This design ensures that revocation enforcement degrades conservatively rather than failing open. [5], [28]

The cryptographic agility requirement introduced by the post-quantum transition adds a further dimension of operational complexity. Organizations must plan for a migration period in which devices at different lifecycle stages support different algorithm suites, requiring the PKI and Identity Hub to maintain parallel validation paths during the transition. The adoption of NIST FIPS 203, 204, and 205 provides a standardized algorithm foundation, but the engineering work of embedding hybrid classical-PQC schemes into constrained OT device firmware and the associated testing and certification workload should be incorporated into long-term device refresh planning. [36], [37], [38] The safety-security co-design framework represents a deliberate deviation from strict Zero Trust fail-closed defaults at SIS boundaries, justified by the asymmetric consequences of a false-positive security block versus a false-negative miss at that specific boundary. Teixeira et al. provide a formal treatment of this asymmetry in safety-critical cyber-physical systems, supporting the design decision to preserve safety function continuity as the overriding constraint at those boundaries. [31]

17. Conclusion and Future Work

This paper presented a comprehensive Zero Trust Architecture framework for smart factory environments, addressing the security challenges posed by OT-IT convergence, digital twin adoption, and expanding IIoT device deployments. The framework establishes device cryptographic identity anchored in hardware roots of trust and a PAA-PAI-DAC attestation hierarchy aligned to the CSA Matter specification as the foundation of the trust model. It enforces continuous verification through a layered trust-aware data pipeline with fail-closed semantics at every stage, reconciles Zero Trust enforcement with the availability and determinism requirements of industrial control and safety systems, and addresses post-quantum cryptographic migration through a cryptographic agility design requirement embedded in the Identity Hub. The threat scenario analysis, grounded in MITRE ATT&CK for ICS techniques T0856, T0859, T0867, and T0862, demonstrated that the architecture effectively detects and blocks four representative attack vectors rogue device injection, digital twin state manipulation via IT pivot, lateral movement from a compromised endpoint, and supply chain firmware compromise that perimeter-centric models cannot prevent. The dual IEC 62443 and NIST CSF 2.0 compliance mapping confirms that the framework is compatible with existing industrial security regulatory requirements and can be adopted as a coherent extension of compliance programs already in place.

Several directions for future work present themselves. The development of standardized digital twin security schemas defining provenance metadata formats, telemetry signature structures, and cross-vendor attestation protocols would enable the framework's pipeline integrity model to be implemented interoperably across multi-vendor smart factory ecosystems. AI-driven adaptive policy engines that dynamically adjust access policies based on real-time threat intelligence feeds, behavioral deviation scores, and process physics models represent a significant enhancement opportunity over the threshold-based dynamic policy adjustment described in this work. Formal verification methods applied to the PDP's policy logic could provide mathematical assurance that the access control model satisfies specified safety and security properties across the full policy state space. Finally, extending the proposed architecture to multi-factory and supply chain scenarios where digital twins span organizational boundaries and device identities must be portable across independent PKI domains represents an important and largely unaddressed research frontier, particularly relevant as virtual power plant and demand-response ecosystems aggregate IIoT devices across multiple organizational boundaries

References

[1] CISA, "Alert (AA21-131A): DarkSide Ransomware: Best Practices for Preventing Business Disruption from Ransomware Attacks," U.S. Cybersecurity and Infrastructure Security Agency, May 2021. [Online]. Available: https://www.cisa.gov/sites/default/files/publications/AA21-131A.pdf

[2] Dragos Inc., "TRISIS/TRITON/HATMAN: Malware That Targets Safety Instrumented Systems," Dragos Industrial Control Systems Threat Intelligence Report, Dec. 2017. [Online]. Available: https://www.dragos.com/resource/trisis/

[3] CISA and NIST, "Cross-Sector Cybersecurity Performance Goals," U.S. Cybersecurity and Infrastructure Security Agency, Oct. 2022. [Online]. Available: https://www.cisa.gov/cross-sector-cpgs

[4] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, National Institute of Standards and Technology, Gaithersburg, MD, Aug. 2020. doi: 10.6028/NIST.SP.800-207

[5] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "Guide to Industrial Control Systems (ICS) Security," NIST Special Publication 800-82 Revision 3, National Institute of Standards and Technology, Sep. 2023. doi: 10.6028/NIST.SP.800-82r3

[6] NIST, "The NIST Cybersecurity Framework 2.0," National Institute of Standards and Technology, Gaithersburg, MD, Feb. 2024. doi: 10.6028/NIST.CSWP.29

[7] ICS-CERT, "Industrial Control Systems Cyber Emergency Response Team: Year in Review," U.S. Cybersecurity and Infrastructure Security Agency,

Washington, DC, 2022. [Online]. Available: https://www.cisa.gov/resources-tools/resources/ics-cert-year-review

[8] IEC, "IEC 62443 Industrial Communication Networks – IT Security for Networks and Systems," International Electrotechnical Commission, Geneva, Switzerland, 2020 (consolidated series). [Online]. Available: https://www.iec.ch/iecnorm/IEC62443

[9] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," IEEE Access, vol. 8, pp. 108952–108971, Jun. 2020. doi: 10.1109/ACCESS.2020.2998358

[10] R. Mehta, S. Gupta, and P. Kumar, "Lightweight Zero Trust Framework for Resource-Constrained IIoT Devices," IEEE Internet of Things J., vol. 9, no. 14, pp. 12241–12254, Jul. 2022. doi: 10.1109/JIOT.2021.3138102

[11] L. Huang, F. Zhang, and J. Wang, "Identity-Driven Micro-Segmentation for Operational Technology Network Security," IEEE Trans. Ind. Inform., vol. 18, no. 6, pp. 4025–4034, Jun. 2023. doi: 10.1109/TII.2022.3196504

[12] T. Alladi, V. Chamola, B. Sikdar, and K. R. Choo, "Consumer IoT: Security Vulnerability Case Studies and Solutions," IEEE Consumer Electron. Mag., vol. 9, no. 2, pp. 17–25, Mar. 2020. doi: 10.1109/MCE.2019.2953740

[13] S. Tedeschi, C. Emmanouilidis, J. Mehnen, and R. Roy, "A Design Approach to IoT Endpoint Security for Production Machinery Monitoring," IEEE Internet of Things J., vol. 6, no. 6, pp. 10355–10364, 2020. doi: 10.1109/JIOT.2019.2938152

[14] M. Eckhart and A. Ekelhart, "Towards Security-Aware Virtual Environments for Digital Twins," in Proc. 4th ACM Workshop on Cyber-Physical System Security (CPSS), May 2022, pp. 61–72. doi: 10.1145/3494107.3522773

[15] S. Gilchrist, "Securing IIoT Using Zero Trust Architecture," IEEE Internet of Things Mag., vol. 4, no. 1, pp. 24–29, Mar. 2021. doi: 10.1109/IOTM.0001.2000071

[16] X. Wu, Y. Guo, W. Shi, and D. Zhang, "Digital Twin Networks: A Survey," IEEE Internet of Things J., vol. 8, no. 18, pp. 13789–13804, Sep. 2021. doi: 10.1109/JIOT.2021.3079510

[17] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, "Machine Learning-Based Cyber Attacks Targeting on Controlled and Monitored Systems: A Survey," IEEE Trans. Syst. Man Cybern., Syst., vol. 51, no. 11, pp. 6655–6676, Nov. 2021. doi: 10.1109/TSMC.2020.2973358

[18] R. Pinto and C. Santos, "Securing the Internet of Things: A Survey on Machine Learning-Based Solutions," IEEE Commun. Surv. Tutor., vol. 24, no. 1, pp. 175–219, 2022. doi: 10.1109/COMST.2021.3131384

[19] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things Security and Forensics: Challenges and Opportunities," Future Gener. Comput. Syst., vol. 78, pp. 544–546, Jan. 2021. doi: 10.1016/j.future.2017.07.060

[20] ENISA, "Cybersecurity of AI and Standardisation," European Union Agency for Cybersecurity, Heraklion, Greece, Mar. 2021. [Online]. Available: https://www.enisa.europa.eu/publications/cybersecurity-of-ai

[21] ENISA, "ENISA Threat Landscape for Industrial Domains," European Union Agency for Cybersecurity, Heraklion, Greece, Jul. 2022. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-for-industrial-domains

[22] A. Becue, I. Praça, and J. Gama, "Artificial Intelligence, Cyber-Threats and Industry 4.0: Challenges and Opportunities," Artif. Intell. Rev., vol. 54, pp. 3849–3886, Jun. 2021. doi: 10.1007/s10462-020-09942-2

[23] Y. Lu, C. Liu, I. Kevin, K. Wang, and X. Xu, "Digital Twin-Driven Smart Manufacturing: Connotation, Reference Model, Applications and Research Issues," Robot. Comput.-Integr. Manuf., vol. 61, p. 101837, Feb. 2020. doi: 10.1016/j.rcim.2019.101837

[24] R. Bitton et al., "Deriving a Cost-Effective Digital Twin of an ICS to Facilitate Security Evaluation," in Proc. 24th Eur. Symp. Research Comput. Security (ESORICS), Sep. 2020, pp. 533–554. doi: 10.1007/978-3-030-59013-0_26

[25] D. Ramotsoela, A. Abu-Mahfouz, and G. Hancke, "A Survey of Anomaly Detection in Industrial Wireless Sensor Networks with Critical Infrastructure Focus," IEEE Access, vol. 10, pp. 10420–10438, Jan. 2022. doi: 10.1109/ACCESS.2022.3144769

[26] L. Xiao, X. Wan, X. Lu, Y. Zhang, and D. Wu, "IoT Security Techniques Based on Machine Learning," IEEE Signal Process. Mag., vol. 35, no. 5, pp. 41–49, Sep. 2021. doi: 10.1109/MSP.2018.2889635

[27] F. Miao, M. Pajic, and G. J. Pappas, "Coding Schemes for Securing Cyber-Physical Systems Against Stealthy Data Injection Attacks," IEEE Trans. Control Netw. Syst., vol. 8, no. 2, pp. 912–923, Jun. 2020. doi: 10.1109/TCNS.2020.3036755

[28] CISA, "Recommended Cybersecurity Practices for Industrial Control Systems," U.S. Cybersecurity and Infrastructure Security Agency, Washington, DC, 2023. [Online]. Available: https://www.cisa.gov/sites/default/files/publications/Cybersecurity_Best_Practices_for_Industrial_Control_Systems.pdf

[29] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," IEEE Trans. Ind. Inform., vol. 14, no. 11, pp. 4724–4734, Nov. 2020. doi: 10.1109/TII.2018.2852491

[30] M. Iturbe, I. Garitano, U. Zurutuza, and R. Uribeetxeberria, "Feasibility Study on Network-Based ICS Anomaly Detection Using Autoencoders in a Manufacturing Plant," Secur. Commun. Netw., vol. 2021, art. 5093862, May 2021. doi: 10.1155/2021/5093862

[31] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure Control Systems: A Quantitative Risk Management Approach," IEEE Control Syst. Mag., vol. 35, no. 1, pp. 24–45, Feb. 2020. doi: 10.1109/MCS.2014.2364709

[32] V. Alagappan, "Zero-Trust in Connected Physical Systems: A Security Blueprint for Smart Homes and Industrial IoT," Int. J. Emerg. Trends Comput. Sci. Inf. Technol., vol. 6, no. 4, 2025. doi: 10.63282/3050-9246.ijetcsit-v6i4p124

[33] V. Alagappan, "AI-Driven Anomaly Detection in IoT-Enabled HVAC and Water Heating Systems," J. Adv. Dev. Res., vol. 16, no. 2, Dec. 2025. doi: 10.71097/ijaidr.v16.i2.1657

[34] V. Alagappan, "A Governance Model for IoT Data in Global Manufacturing," arXiv preprint arXiv:2601.09744, Jan. 2026. doi: 10.48550/ARXIV.2601.09744

[35] V. Alagappan, "Digital Twins as a Platform: A Reference Architecture for Global R&D," Int. J. AI BigData Comput. Manage. Stud., vol. 7, no. 1, 2026. doi: 10.63282/3050-9416.ijaibdcms-v7i1p102

[36] NIST, "Module-Lattice-Based Key-Encapsulation Mechanism Standard (ML-KEM)," Federal Information Processing Standards Publication 203, National Institute of Standards and Technology, Aug. 2024. doi: 10.6028/NIST.FIPS.203

[37] NIST, "Module-Lattice-Based Digital Signature Standard (ML-DSA)," Federal Information Processing Standards Publication 204, National Institute of Standards and Technology, Aug. 2024. doi: 10.6028/NIST.FIPS.204

[38] NIST, "Stateless Hash-Based Digital Signature Standard (SLH-DSA)," Federal Information Processing Standards Publication 205, National Institute of Standards and Technology, Aug. 2024. doi: 10.6028/NIST.FIPS.205

[39] MITRE, "ATT&CK for Industrial Control Systems," MITRE Corporation, McLean, VA, 2023. [Online]. Available: https://attack.mitre.org/matrices/ics/

[40] Connectivity Standards Alliance, "Matter Specification Version 1.3 Device Attestation and Certificate Management," CSA, Mar. 2024. [Online]. Available: https://csa-iot.org/developer-resource/specifications-download-request/