Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: 2 October 2026

R. Sharif
CyberSecAI Ltd

31 March 2026

ATTP for Industrial Control Systems: Cryptographic Agent Authentication in SCADA and IoT Environments

draft-sharif-attp-industrial-control-systems-00

Abstract

This document defines an application profile of the Agent Trust Transport Protocol (ATTP) [draft-sharif-attp-agent-trust-transport] for use in Industrial Control Systems (ICS), Supervisory Control and Data Acquisition (SCADA) environments, and Internet of Things (IoT) deployments. It specifies how ATTP mandatory message signing, agent identity passports, and trust-gated access control apply to industrial protocols including Modbus/TCP, OPC UA, MQTT, and CoAP.

The profile addresses the absence of per-message authentication in legacy industrial protocols, which has been exploited in numerous critical infrastructure attacks. It defines a gateway architecture that enables ATTP protection for legacy devices without firmware modification, maps ATTP trust levels to IEC 62443 Security Levels, and specifies real-time revocation mechanisms suitable for safety-critical environments.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 2 October 2026.

Copyright Notice

Table of Contents

1. Introduction

Industrial Control Systems (ICS) and Supervisory Control and Data Acquisition (SCADA) systems increasingly incorporate autonomous AI agents for predictive maintenance, load balancing, anomaly detection, and process optimisation. These agents communicate with Programmable Logic Controllers (PLCs), Remote Terminal Units (RTUs), Human-Machine Interfaces (HMIs), and other field devices using protocols that were designed decades before autonomous software agents existed.

The core industrial protocols -- Modbus (1979), DNP3 (1990), and OPC UA (2008) -- lack mandatory per-message authentication. Modbus/TCP has no authentication mechanism whatsoever. DNP3 Secure Authentication (SA) is an optional extension that is rarely deployed. OPC UA supports signing but implementations frequently disable it for performance reasons. MQTT, widely used in IoT, has optional TLS but no per-message signing.

This absence of message-level authentication has been exploited in every major ICS attack:

- Stuxnet (2010): Injected fake commands to centrifuge PLCs via unsigned Profinet/OPC messages.
- BlackEnergy/Ukraine (2015): Sent unsigned SCADA commands to disconnect power substations, causing widespread blackouts.
- TRITON/TRISIS (2017): Impersonated a Triconex safety controller by sending unsigned commands to the Safety Instrumented System.
- Oldsmar Water Treatment (2021): Attacker accessed HMI remotely and sent unsigned commands to increase sodium hydroxide to dangerous levels.
- Colonial Pipeline (2021): Compromised VPN with no device identity verification led to ransomware deployment on OT networks.
- IOCONTROL (2025): Iranian APT targeting US and Israeli critical infrastructure IoT and OT systems, exploiting the absence of device authentication.

The Agent Trust Transport Protocol (ATTP) [draft-sharif-attp-agent-trust-transport] provides mandatory ECDSA P-256 message signing, cryptographic agent identity via passports,

and trust-gated access control (L0-L4). This document defines how ATTP applies to industrial environments, addressing the unique requirements of safety-critical systems, legacy device compatibility, real-time constraints, and regulatory compliance with IEC 62443, NIST SP 800-82, and the EU NIS2 Directive.

1.1. Scope

This profile covers:

- ATTP signing for Modbus/TCP, OPC UA, MQTT, and CoAP protocols
- Gateway architecture for legacy device protection
- Trust level mapping to IEC 62443 Security Levels (SL 1-4)
- Safety Integrity Level (SIL) integration per IEC 61508
- Real-time revocation for safety-critical environments
- Audit trail requirements for regulatory compliance

This profile does not cover serial Modbus (RS-485), fieldbus protocols (PROFIBUS, Foundation Fieldbus), or wireless industrial protocols (WirelessHART, ISA100.11a), though the gateway architecture may be extended to these in future work.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

ATTP: Agent Trust Transport Protocol, as defined in [draft-sharif-attp-agent-trust-transport].

Agent Passport: A signed JSON structure containing the agent identity, trust level, capabilities, owner, and public key.

ATTP Gateway: A network appliance or software process that intercepts industrial protocol traffic and applies ATTP signing and verification on behalf of legacy devices.

Field Device: A PLC, RTU, IED, sensor, actuator, or other physical device in an industrial control system.

Safety-Critical Command: Any command that, if executed incorrectly, could cause physical harm, environmental damage, or equipment destruction.

SIL: Safety Integrity Level, as defined in IEC 61508, ranging from SIL 1 (lowest) to SIL 4 (highest).

3. Threat Model

This section defines the threat model specific to AI agents operating in industrial control environments.

3.1. Threat Actors

T1 - Nation-State APT: Advanced persistent threat groups targeting critical infrastructure. Capabilities include zero-day exploits, supply chain compromise, and long-duration persistence. Examples: Sandworm (Russia), APT33 (Iran), Volt Typhoon (China).

T2 - Hacktivist: Ideologically motivated actors targeting industrial systems for disruption. Capabilities include exploitation of exposed HMIs and known vulnerabilities. Examples: Z-Pentest, CyberAv3ngers, Dark Engine.

T3 - Insider Threat: Authorised personnel with legitimate access who misuse their access or whose credentials are compromised.

T4 - Compromised AI Agent: An autonomous agent that has been subverted through prompt injection, model poisoning, supply chain attack, or credential theft.

T5 - Rogue Agent: An unauthorised agent that gains network access to the industrial environment through misconfiguration, VPN compromise, or lateral movement.

3.2. Attack Vectors

AV1 - Command Injection: Sending unsigned commands to PLCs, RTUs, or actuators. Without per-message authentication, any network entity can issue control commands. ATTP prevents this by requiring a valid signature on every command.

AV2 - Device Impersonation: An attacker impersonates a legitimate controller, HMI, or agent by sending messages that appear to originate from a trusted source. ATTP agent passports bind cryptographic identity to each device, preventing impersonation.

AV3 - Command Replay: Capturing and replaying a valid command at a later time. In industrial systems, a replayed "open valve" command can cause physical damage. ATTP nonces and timestamps prevent replay.

AV4 - Man-in-the-Middle: Intercepting and modifying commands in transit. Even with TLS, a compromised proxy or gateway can modify decrypted traffic. ATTP per-message signing provides end-to-end integrity independent of transport encryption.

AV5 - Trust Escalation: A low-privilege agent attempts to execute safety-critical commands. ATTP trust levels (L0-L4) ensure that only agents with sufficient verification can execute high-impact commands.

AV6 - Mass Revocation Failure: A compromised agent continues operating because revocation propagation is too slow. TLS Certificate Revocation Lists (CRLs) can take hours to propagate. ATTP specifies real-time revocation suitable for safety-critical environments (Section 8).

3.3. Assets Under Threat

- Physical processes (valve positions, pump speeds, temperatures)
- Safety instrumented systems (emergency shutdown, fire/gas)
- Firmware and configuration of field devices
- Sensor data integrity (temperature, pressure, flow rates)
- Historian and audit log integrity
- Human safety

4. Architecture

The ATTP-ICS architecture supports two deployment models:

4.1. Native ATTP Devices

New devices with sufficient computational resources (ARM Cortex-M4

or higher, ESP32, Raspberry Pi, industrial gateways) implement ATTP natively. These devices:

- Generate ECDSA P-256 keypairs on first boot
- Carry an agent passport signed by the facility Certificate Authority (CA)
- Sign every outgoing message
- Verify every incoming message
- Reject unsigned or insufficiently trusted messages

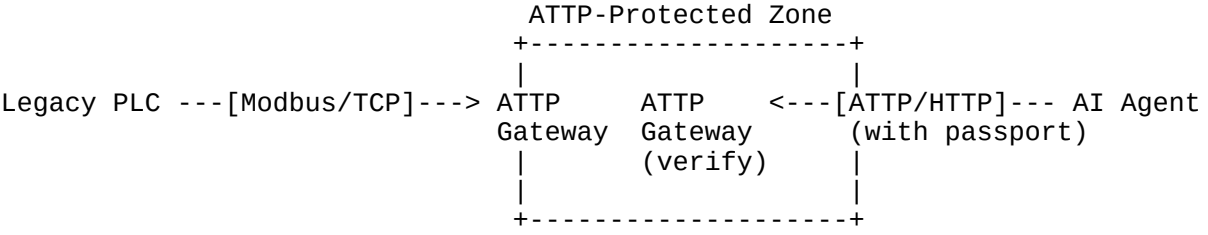
Minimum hardware requirements:

- 32-bit processor, 160 MHz or higher
- 256 KB RAM
- Hardware cryptographic accelerator (recommended, not required)
- WiFi, Ethernet, or cellular connectivity

Note: ECDSA P-256 signing has been demonstrated on ESP32-C3 RISC-V microcontrollers ($3 unit cost, hardware-accelerated) with signing latency under 10ms.

4.2. Gateway-Protected Legacy Devices

Existing PLCs, RTUs, and field devices that cannot be modified are protected by an ATTP Gateway:



The ATTP Gateway:

- Terminates the industrial protocol (Modbus/TCP, OPC UA, etc.)
- Verifies the ATTP signature and passport on incoming commands
- Checks trust level against the command's required minimum
- If verified, forwards the command to the legacy device
- Signs the response from the legacy device before forwarding
- Logs every transaction to the audit trail

The gateway acts on behalf of the legacy device, which does not need firmware changes. This is analogous to a TLS termination proxy, but operating at the application message layer rather than the transport layer.

4.3. Network Architecture

The recommended deployment follows the Purdue Enterprise Reference Architecture (PERA) model:

```
Level 4 (Enterprise):  ATTP-enabled AI agents with L3/L4 passports
Level 3.5 (DMZ):       ATTP Gateways (signing and verification)
Level 3 (Operations):  Historians, MES (ATTP-signed data feeds)
Level 2 (Control):     SCADA, DCS (gateway-protected)
Level 1 (Field):       PLCs, RTUs (gateway-protected)
Level 0 (Process):     Sensors, actuators (physical)
```

ATTP Gateways sit at Level 3.5, the Demilitarised Zone between the enterprise network and the control network. All traffic crossing this boundary MUST be ATTP-signed.

5. Industrial Protocol Profiles

5.1. Modbus/TCP Profile

Modbus/TCP [RFC 793] carries no authentication. Function codes (read coils, write registers, etc.) are accepted from any TCP client. The ATTP Modbus profile wraps Modbus Application Data Units (ADUs) in ATTP envelopes:

```
ATTP-Modbus Message:
{
  "attp_version": "1.0",
  "protocol": "modbus-tcp",
  "agent_passport": { ... },
  "payload": {
    "transaction_id": 1,
    "unit_id": 1,
    "function_code": 6,
    "register_address": 100,
    "register_value": 500
  },
  "signature": "<ECDSA-P256-signature>",
  "nonce": "<unique-nonce>",
  "timestamp": 1711900000
}
```

The ATTP Gateway intercepts Modbus/TCP on port 502, verifies the ATTP envelope, and if valid, translates back to native Modbus/TCP for the legacy PLC.

Trust requirements for Modbus function codes:

- Read operations (FC 1-4): Minimum trust L1
- Write single register (FC 5-6): Minimum trust L2
- Write multiple registers (FC 15-16): Minimum trust L3
- Diagnostics and device identification (FC 8, 17): Minimum L3
- Encapsulated interface transport (FC 43): Minimum L4

5.2. OPC UA Profile

OPC UA supports message-level signing via its native security model, but implementations frequently disable it. The ATTP OPC UA profile adds ATTP signing as an additional layer:

- ATTP signatures are carried in the OPC UA message header as extension objects
- The ATTP agent passport is presented during the OPC UA CreateSession handshake
- Trust levels are mapped to OPC UA user roles
- ATTP revocation is checked before session activation

For implementations that already use OPC UA signing, ATTP provides the agent identity layer (passports, trust levels) that OPC UA lacks.

5.3. MQTT Profile

MQTT [RFC 9431] is the dominant IoT messaging protocol. MQTT v5 supports user properties in message headers, which carry ATTP metadata:

```
MQTT Publish with ATTP:
Topic:   factory/line-3/temperature
Payload: {"value": 72.4, "unit": "celsius"}
User Properties:
  X-ATTP-Version:   1.0
  X-ATTP-Signature: <ECDSA-P256-signature-of-payload>
```

```
X-ATTP-Agent-ID:  sensor-agent-line3-temp01
X-ATTP-Trust:     L2
X-ATTP-Nonce:     <unique-nonce>
X-ATTP-Timestamp: 1711900000
```

The MQTT broker or a subscribing ATTP Gateway verifies signatures before processing or forwarding messages.

Trust requirements for MQTT operations:

- Subscribe to sensor data topics: Minimum trust L1
- Publish sensor readings: Minimum trust L2
- Publish setpoint changes: Minimum trust L3
- Publish to safety-critical topics: Minimum trust L4
- Broker administration ($SYS topics): Minimum trust L4

5.4. CoAP Profile

The Constrained Application Protocol (CoAP) [RFC 7252] is used in resource-constrained IoT devices. ATTP metadata is carried in CoAP options:

- Option 65001: ATTP-Signature (opaque, variable length)
- Option 65003: ATTP-Agent-ID (string)
- Option 65005: ATTP-Trust-Level (uint, 0-4)
- Option 65007: ATTP-Nonce (opaque, 16 bytes)
- Option 65009: ATTP-Timestamp (uint)

For devices too constrained for ECDSA P-256 (8-bit microcontrollers with less than 32 KB RAM), the ATTP Gateway model (Section 4.2) MUST be used.

6. Trust Level Mapping

ATTP trust levels (L0-L4) map to established industrial security frameworks:

ATTP	Description	IEC 62443 SL	Typical ICS Use
L0	Unverified	Below SL 1	Blocked at gateway
L1	Self-Signed	SL 1	Read-only sensors
L2	Verified Device	SL 2	Standard telemetry
L3	Org-Certified	SL 3	Write operations
L4	Hardware-Bound	SL 4	Safety-critical

The mapping to IEC 62443 Security Levels ensures that ATTP deployments align with existing industrial cybersecurity certification requirements.

NIST SP 800-82 mapping:

- L1: Aligns with NIST SP 800-82 "Monitoring" access
- L2: Aligns with NIST SP 800-82 "Operator" access
- L3: Aligns with NIST SP 800-82 "Engineer" access
- L4: Aligns with NIST SP 800-82 "Administrator" access

7. Gateway Specification

7.1. Functional Requirements

An ATTP-ICS Gateway MUST:

- Terminate one or more industrial protocols (Modbus/TCP, OPC UA, MQTT, CoAP)

- Verify ATTP signatures on all incoming messages
- Verify agent passports against a local or remote trust store
- Enforce minimum trust levels per command type (Section 5)
- Reject unsigned messages with no fallback to insecure mode
- Sign all outgoing responses on behalf of legacy devices
- Maintain an audit log of all transactions (Section 10)
- Support real-time revocation checks (Section 8)
- Operate with maximum added latency of 5ms per message

7.2. Performance Requirements

Industrial control systems have strict timing requirements. The ATTP Gateway MUST meet:

- Signing latency: less than 2ms (with hardware acceleration)
- Verification latency: less than 2ms (with hardware acceleration)
- Total added latency: less than 5ms per message
- Throughput: minimum 10,000 messages per second
- Availability: 99.999% (five nines)

These requirements are achievable with ECDSA P-256 hardware acceleration available on modern ARM processors (Cortex-A53 and above) and FPGA-based industrial platforms.

7.3. Fail-Safe Behaviour

In the event of gateway failure:

- The gateway MUST fail closed (block all traffic)
- A redundant gateway MUST take over within 100ms
- Safety-critical systems MUST have a hardware bypass that requires physical key activation (not software-controlled)
- All bypass activations MUST be logged to a separate audit trail

8. Revocation and Emergency Override

8.1. Real-Time Revocation

TLS Certificate Revocation Lists (CRLs) can take hours to propagate, which is unacceptable for safety-critical systems. ATTP-ICS specifies real-time revocation:

- Revocation messages are signed by the facility CA
- Revocation propagates to all gateways within 1 second
- Gateways maintain a local revocation cache
- Revoked agents are immediately blocked at all gateways
- Revocation is irrevocable (a new passport must be issued)

Revocation message format:

```
{
  "type": "revocation",
  "agent_id": "<agent-id-to-revoke>",
  "reason": "compromised|decommissioned|policy_violation",
  "effective": "<ISO-8601-timestamp>",
  "issuer": "<facility-CA-id>",
  "signature": "<CA-ECDSA-signature>"
}
```

8.2. Emergency Override

For safety-critical emergencies where an authorised agent’s passport has been incorrectly revoked or has expired:

- Emergency override requires physical presence (hardware token or biometric) at the gateway

- Override creates a temporary L4 passport valid for a maximum of 60 minutes
- All actions during override are logged with enhanced detail
- Override events trigger immediate alerts to all registered security personnel
- Override cannot be activated remotely

9. Safety Integrity Level Integration

For systems requiring functional safety certification per IEC 61508 / IEC 61511:

- SIL 1: ATTP trust L2 minimum for all safety function inputs
- SIL 2: ATTP trust L3 minimum, dual-signature verification
- SIL 3: ATTP trust L4 minimum, hardware-bound keys, dual gateway verification
- SIL 4: ATTP trust L4, hardware-bound keys, triple modular redundancy on gateway verification, formal verification of signing implementation

The ATTP signing implementation used in SIL 3 and SIL 4 environments SHOULD be formally verified and certified to the appropriate Common Criteria Evaluation Assurance Level (EAL).

10. Audit Trail Requirements

All ATTP-ICS deployments MUST maintain audit trails compliant with the ATTP Audit Trail format [draft-sharif-agent-audit-trail]:

- Every signed command MUST be logged with: agent ID, trust level, timestamp, nonce, command details, signature, and verification result
- Audit records MUST be hash-chained for tamper evidence
- Audit logs MUST be retained for minimum 3 years (IEC 62443) or as required by sector-specific regulation
- Audit logs MUST support export to SIEM systems via syslog (RFC 5424) or structured JSON (JSONL)

Regulatory mapping:

- EU NIS2 Directive: Article 21 (risk management), Article 23 (incident reporting)
- IEC 62443-3-3: SR 6.1 (audit log accessibility), SR 6.2 (continuous monitoring)
- NIST SP 800-82: Section 6.2.6 (audit and accountability)
- EU AI Act: Article 12 (record-keeping), Article 14 (human oversight)

11. Deployment Considerations

11.1. Migration Strategy

Deploying ATTP in existing industrial environments requires a phased approach:

Phase 1 - Monitor: Deploy ATTP Gateways in passive mode. Log all unsigned traffic without blocking. Identify all agents and communication patterns.

Phase 2 - Warn: Enable ATTP verification. Log violations but allow unsigned traffic through with warnings. Issue passports to known agents.

Phase 3 - Enforce: Block unsigned traffic. All agents must present valid passports. Legacy devices protected by gateways.

Phase 4 - Harden: Increase minimum trust levels. Enable safety-critical trust requirements. Deploy hardware-bound keys for L4.

Each phase should run for a minimum of 30 days before advancing to the next phase.

11.2. Key Management

- Facility CA: Generates and signs agent passports. MUST be stored in a Hardware Security Module (HSM).
- Gateway keys: Generated on the gateway. SHOULD use hardware key storage (TPM 2.0 or secure enclave).
- Agent keys: Generated on the agent device. Hardware key storage RECOMMENDED for L3+, REQUIRED for L4.
- Key rotation: Minimum annually for L1-L2, quarterly for L3, monthly for L4.

12. Worked Examples

This section provides complete worked examples demonstrating ATTP-ICS in realistic industrial scenarios.

12.1. Example 1: AI Agent Adjusts Boiler Temperature

Scenario: An AI predictive maintenance agent detects that boiler pressure is rising above optimal parameters and needs to reduce the burner setpoint from 850C to 780C via a Modbus write command.

Step 1: Agent constructs the command payload:

```
{
  "protocol": "modbus-tcp",
  "unit_id": 3,
  "function_code": 6,
  "register_address": 40001,
  "register_value": 780,
  "context": {
    "reason": "pressure_above_threshold",
    "current_pressure_psi": 142.7,
    "target_pressure_psi": 125.0,
    "model_confidence": 0.94
  }
}
```

Step 2: Agent signs the payload with its ECDSA P-256 private key:

```
Hash:     SHA-256(payload) = a7c3f8e2...
Signature: MEYCIQC7xL8kN2pRvT...  (72 bytes, DER-encoded)
```

Step 3: Agent constructs the full ATTP-ICS message:

```
{
  "attp_version": "1.0",
  "attp_profile": "ics-modbus",
  "agent_passport": {
    "agent_id": "pred-maint-agent-boiler-3",
    "owner": "ACME Manufacturing Ltd",
    "trust_level": 3,
    "capabilities": ["read_sensors", "write_setpoints"],
    "issued_by": "acme-facility-ca",
    "issued_at": "2026-03-15T10:00:00Z",
    "expires_at": "2026-06-15T10:00:00Z",
    "public_key": "MFkwEwYHKoZIzj0CAQY..."
  },
```

```
  "payload": { ... },
  "signature": "MEYCIQC7xL8kN2pRvT...",
  "algorithm": "ES256",
  "nonce": "a8f3c7e2d1b94f60",
  "timestamp": 1711900000
}

Step 4: ATTP Gateway receives the message and performs:

4a. Parse ATTP envelope                          [0.1ms]
4b. Check attp_version is "1.0"                  [0.01ms]
4c. Check attp_profile is "ics-modbus"           [0.01ms]
4d. Verify agent_passport signature (CA key)     [0.8ms]
4e. Check passport not expired                   [0.01ms]
4f. Check passport not revoked (local cache)     [0.05ms]
4g. Check trust_level >= 3 (write operation)     [0.01ms]
4h. Check nonce not seen before                  [0.05ms]
4i. Check timestamp within 30-second window      [0.01ms]
4j. Verify payload signature (agent public key)  [0.8ms]
4k. Check agent has "write_setpoints" capability [0.01ms]
4l. Validate register_value is within safe range [0.01ms]
                                      Total:     [1.87ms]

Step 5: Gateway translates to native Modbus/TCP:

Transaction ID: 0x0001
Protocol ID:    0x0000 (Modbus)
Length:          0x0006
Unit ID:         0x03
Function Code:   0x06 (Write Single Register)
Register Addr:   0x9C41 (40001)
Register Value:  0x030C (780)

Step 6: Gateway sends native Modbus to the PLC, receives
response, signs the response, and returns to the agent:

{
  "attp_version": "1.0",
  "status": "verified_and_executed",
  "gateway_id": "gw-boiler-zone-3",
  "original_nonce": "a8f3c7e2d1b94f60",
  "modbus_response": {
    "function_code": 6,
    "register_address": 40001,
    "register_value": 780
  },
  "gateway_signature": "MEUCIQDnR7vK...",
  "timestamp": 1711900002
}

Step 7: Gateway logs the complete transaction to the audit trail:

{
  "audit_version": "1.0",
  "sequence": 847293,
  "previous_hash": "e4a2c8f1...",
  "event": "command_executed",
  "agent_id": "pred-maint-agent-boiler-3",
  "trust_level": 3,
  "command": "modbus_write_register",
  "target": "plc-boiler-3:40001",
  "value": 780,
  "previous_value": 850,
  "verification_result": "pass",
  "verification_time_ms": 1.87,
  "gateway_id": "gw-boiler-zone-3",
```

```
  "timestamp": "2026-03-31T14:26:40Z",
  "hash": "b7d1e3f5..."
}
```

12.2. Example 2: Rogue Agent Blocked at Gateway

Scenario: A compromised agent attempts to open a safety relief valve by sending a Modbus write command. The agent has trust level L1 (self-signed) but the valve command requires L4.

Agent sends:

```
{
  "attp_version": "1.0",
  "attp_profile": "ics-modbus",
  "agent_passport": {
    "agent_id": "unknown-agent-x",
    "trust_level": 1,
    "issued_by": "self",
    ...
  },
  "payload": {
    "function_code": 5,
    "coil_address": 1024,
    "coil_value": 65280
  },
  "signature": "MEQCIAx9...",
  ...
}
```

Gateway response:

```
{
  "attp_version": "1.0",
  "status": "rejected",
  "reason": "insufficient_trust_level",
  "required_trust": 4,
  "presented_trust": 1,
  "command": "write_coil",
  "target": "safety-valve-1024",
  "safety_classification": "SIL-3",
  "gateway_id": "gw-safety-zone-1",
  "timestamp": 1711900100,
  "gateway_signature": "MEUCIQD..."
}
```

Gateway audit log entry:

```
{
  "event": "command_rejected",
  "severity": "critical",
  "agent_id": "unknown-agent-x",
  "trust_level": 1,
  "required_trust": 4,
  "command": "write_coil:1024",
  "reason": "insufficient_trust_level",
  "safety_classification": "SIL-3",
  "alert_triggered": true,
  "alert_recipients": ["security-ops", "plant-manager"],
  "timestamp": "2026-03-31T14:28:20Z"
}
```

The gateway immediately triggers a security alert because a low-trust agent attempted a safety-critical command.

12.3. Example 3: MQTT Sensor Network with ATTP

Scenario: A factory floor has 200 temperature sensors publishing readings via MQTT. An AI agent subscribes to aggregate data and detect anomalies. All MQTT messages carry ATTP signatures.

Sensor publishes (MQTT v5):

```
Topic:    factory/zone-2/temp/sensor-147
QoS:      1
Payload:  {"value": 74.2, "unit": "celsius", "ts": 1711900200}
User Properties:
  X-ATTP-Version:     1.0
  X-ATTP-Profile:     ics-mqtt
  X-ATTP-Agent-ID:    temp-sensor-147
  X-ATTP-Trust:       2
  X-ATTP-Signature:   MEYCIQDk... (base64)
  X-ATTP-Nonce:       c4e8a2f1b7d39605
  X-ATTP-Timestamp:   1711900200
  X-ATTP-Passport:    eyJhZ2VudF9pZCI... (base64-encoded passport)
```

MQTT broker with ATTP plugin verifies:

1. Decode X-ATTP-Passport from base64
2. Verify passport signature against facility CA public key
3. Check passport.trust_level >= 2 (publish to sensor topics)
4. Verify X-ATTP-Signature against payload using passport public key
5. Check X-ATTP-Nonce uniqueness (dedup cache, 60-second window)
6. Check X-ATTP-Timestamp within 30 seconds of broker time
7. If all pass: deliver to subscribers
8. If any fail: drop message, log violation, alert

AI agent subscribes with its own L3 passport:

```
Topic:    factory/zone-2/temp/#
User Properties:
  X-ATTP-Agent-ID:    anomaly-detector-zone2
  X-ATTP-Trust:       3
  X-ATTP-Passport:    eyJhZ2VudF9pZCI... (L3 passport)
```

The broker verifies the subscriber’s trust level before allowing the subscription. An L1 agent attempting to subscribe to setpoint topics (requiring L3) would be rejected.

12.4. Example 4: Emergency Revocation

Scenario: The security team detects that pred-maint-agent-boiler-3 has been compromised. They issue an immediate revocation.

Revocation broadcast (signed by facility CA):

```
{
  "type": "revocation",
  "version": "1.0",
  "agent_id": "pred-maint-agent-boiler-3",
  "reason": "compromised",
  "evidence": "anomalous_command_pattern_detected",
  "effective": "2026-03-31T14:30:00Z",
  "issuer": "acme-facility-ca",
  "broadcast_id": "rev-2026-0331-001",
  "signature": "MEUCIQD2xK..."
}
```

Propagation timeline:

```
T+0ms:     Security operator issues revocation command
T+50ms:    Facility CA signs revocation message
```

```
T+100ms:   Revocation broadcast to all gateways (multicast)
T+200ms:   Gateway gw-boiler-zone-3 receives revocation
T+210ms:   Gateway verifies CA signature on revocation
T+220ms:   Gateway adds agent to local revocation cache
T+230ms:   Gateway drops any in-flight commands from this agent
T+500ms:   All gateways in facility have received revocation
T+1000ms:  Revocation confirmed across all zones
```

Any subsequent command from pred-maint-agent-boiler-3 is immediately rejected at every gateway without signature verification (revocation cache check is faster).

13. Cryptographic Specifications

13.1. Signature Algorithm

All ATTP-ICS implementations MUST support:

- ECDSA with NIST P-256 (secp256r1) and SHA-256 (ES256)

Implementations SHOULD also support:

- ECDSA with NIST P-384 (secp384r1) and SHA-384 (ES384) for environments requiring higher security margins

Implementations MAY support:

- EdDSA with Ed25519 for environments where performance is critical and P-256 hardware acceleration is unavailable

Future revisions of this specification will define profiles for post-quantum signature algorithms (ML-DSA, SLH-DSA) as they are standardised by NIST.

13.2. Signature Encoding

Signatures MUST be encoded in DER format (as specified in SEC 1, Section 4.1.3) and then base64-encoded for transport in JSON and MQTT user properties.

For CoAP, signatures are carried as raw DER bytes in the CoAP option value (no base64 encoding).

13.3. Key Generation

Agent keys MUST be generated using a cryptographically secure random number generator (CSPRNG) seeded with at least 256 bits of entropy.

For constrained devices:

- ESP32: Use hardware RNG via mbedtls_entropy_func
- ARM Cortex-M: Use TRNG peripheral if available
- Linux-based: Use /dev/urandom or getrandom(2)

Private keys MUST NOT be transmitted over any network.
Private keys MUST NOT be logged in audit trails.
Private keys MUST NOT be included in agent passports.

13.4. Passport Format

The ATTP-ICS agent passport extends the base ATTP passport with ICS-specific fields:

```
{
  "version": "1.0",
```

```
  "profile": "ics",
  "agent_id": "pred-maint-agent-boiler-3",
  "agent_type": "predictive_maintenance",
  "owner": "ACME Manufacturing Ltd",
  "facility": "acme-coventry-plant-2",
  "trust_level": 3,
  "iec62443_sl": 3,
  "capabilities": [
    "read_sensors",
    "write_setpoints",
    "read_alarms"
  ],
  "allowed_zones": ["zone-2", "zone-3"],
  "allowed_protocols": ["modbus-tcp", "mqtt"],
  "max_write_rate": 10,
  "safety_clearance": "non-safety",
  "issued_by": "acme-facility-ca",
  "issued_at": "2026-03-15T10:00:00Z",
  "expires_at": "2026-06-15T10:00:00Z",
  "public_key": {
    "kty": "EC",
    "crv": "P-256",
    "x": "f83OJ3D2xF1Bg8vub9tLe1gHMzV76e8Tus9uPHvRVEU",
    "y": "x_FEzRu9m36HLN_tue659LNpXW6pCyStikYjKIWI5a0"
  },
  "issuer_signature": "MEYCIQC..."
}
```

ICS-specific passport fields:

- facility: Identifier of the physical facility where the agent operates. Gateways SHOULD reject passports from other facilities.
- iec62443_sl: IEC 62443 Security Level (1-4) corresponding to the agent's trust level.
- allowed_zones: List of Purdue model zones the agent may communicate with.
- allowed_protocols: List of industrial protocols the agent is authorised to use.
- max_write_rate: Maximum write commands per minute. Gateways MUST enforce this rate limit.
- safety_clearance: "non-safety", "sil-1", "sil-2", "sil-3", or "sil-4". Determines which safety-classified commands the agent may issue.

14. Conformance Requirements

This section defines conformance levels for ATTP-ICS implementations.

14.1. Gateway Conformance

An ATTP-ICS Gateway claiming conformance MUST:

Level 1 (Basic):
- Support at least one industrial protocol profile (Section 5)
- Verify ECDSA P-256 signatures
- Verify agent passports
- Enforce trust level minimums per command type
- Reject unsigned messages (no insecure fallback)
- Generate audit log entries for all transactions

- Achieve less than 5ms added latency per message

Level 2 (Standard):
- All Level 1 requirements
- Support at least two industrial protocol profiles
- Support real-time revocation (Section 8)
- Support passport rate limiting (max_write_rate)
- Support zone-based access control (allowed_zones)
- Hash-chained audit logs
- Syslog export (RFC 5424)

Level 3 (Safety):
- All Level 2 requirements
- Dual-gateway verification for SIL 2+ commands
- Hardware key storage (TPM 2.0 or equivalent)
- Hardware cryptographic acceleration
- Redundant gateway failover within 100ms
- Emergency override support (Section 8.2)
- Formal verification of signing implementation (SIL 3+)
- Achieve less than 2ms added latency per message

14.2. Agent Conformance

An ATTP-ICS Agent claiming conformance MUST:

- Generate ECDSA P-256 keypairs using a CSPRNG
- Sign every outgoing message
- Carry a valid, unexpired passport
- Include a unique nonce in every message
- Include a timestamp within 30 seconds of current time
- Not cache or reuse nonces
- Store private keys securely (hardware storage for L3+)
- Respond to revocation by ceasing all operations

15. Sector-Specific Guidance

15.1. Energy and Power Grid

Power grid deployments face additional requirements:

- NERC CIP compliance: All ATTP audit logs map to NERC CIP-007 (System Security Management) and CIP-005 (Electronic Security Perimeter)
- Real-time constraints: Power grid protection relays require sub-millisecond response. ATTP Gateways for protection systems MUST use hardware acceleration and pre-verified passport caches
- Islanding support: During grid islanding events, gateways MUST continue operating with locally cached revocation lists even if connectivity to the facility CA is lost
- Smart meter integration: Smart meters publishing consumption data via MQTT SHOULD use L1 passports (self-signed). Meters receiving disconnect commands MUST verify L4 passports

15.2. Water and Wastewater

Water treatment deployments:

- Chemical dosing commands (chlorine, fluoride, NaOH) MUST require minimum trust L4 with safety_clearance "sil-2" or higher
- SCADA commands to pumping stations MUST require minimum L3
- Sensor readings from water quality monitors SHOULD carry L2 passports to prevent data poisoning attacks
- The Oldsmar water treatment attack (2021) would have been prevented by ATTP: the attacker’s unsigned HMI commands

would have been rejected at the gateway

15.3. Oil and Gas

Oil and gas deployments:

- Pipeline pressure setpoint changes MUST require L4 with safety_clearance "sil-3"
- Emergency Shutdown (ESD) commands MUST require L4 passports with dual-gateway verification
- Gas detection sensor readings MUST carry L2+ passports to prevent false negative attacks
- Remote wellhead monitoring agents SHOULD use L2 passports with satellite-tolerant timestamp windows (60 seconds instead of 30 seconds for latency compensation)

15.4. Manufacturing

Manufacturing deployments:

- Robot arm control commands MUST require L3+ with safety_clearance appropriate to the robot's safety zone
- Quality inspection AI agents SHOULD carry L2 passports
- Predictive maintenance agents writing setpoint adjustments MUST carry L3 passports with "write_setpoints" capability
- Tool changeover commands SHOULD require L3 with "reconfigure_tooling" capability

15.5. Healthcare and IoMT

Internet of Medical Things deployments:

- Drug infusion pump dosage commands MUST require L4 with safety_clearance "sil-3" and dual-gateway verification
- Patient monitoring sensor data MUST carry L2+ passports
- Building management commands (HVAC in operating theatres) MUST require L3
- HIPAA audit trail requirements are met by ATTP-ICS audit logs with 6-year retention

15.6. Transportation

Rail, aviation, and maritime deployments:

- Signalling system commands MUST require L4 with safety_clearance "sil-4"
- Train control messages MUST use dual-gateway verification with geographic diversity
- Maritime vessel monitoring sensors SHOULD carry L2 passports with extended timestamp windows (120 seconds for satellite connectivity)
- Aviation ground support equipment MUST require L3+ for any commands affecting aircraft systems

16. Comparison with Existing Approaches

16.1. TLS

TLS provides transport-layer encryption and server authentication but does not provide:

- Per-message signing (decrypted messages can be modified by proxies, gateways, and logging systems)
- Agent identity (TLS authenticates the connection, not the specific software agent using it)
- Trust levels (TLS is binary: authenticated or not)

- Command-level access control
- Tamper-evident audit trails

ATTP-ICS operates above TLS. Both SHOULD be used together.

16.2. DNP3 Secure Authentication

DNP3 SA provides challenge-response authentication but:

- Is specific to DNP3 (does not cover Modbus, MQTT, OPC UA)
- Uses HMAC-SHA-256 (symmetric keys, no non-repudiation)
- Has no trust levels or graduated access control
- Has no revocation mechanism
- Is rarely deployed in practice

ATTP-ICS provides asymmetric signatures (non-repudiation), trust levels, revocation, and works across all industrial protocols.

16.3. OPC UA Security

OPC UA has a comprehensive security model but:

- Signing is frequently disabled for performance
- Does not extend to non-OPC UA protocols
- Has no concept of agent identity or trust levels
- Certificate management is complex and often misconfigured
- Has no real-time revocation mechanism

ATTP-ICS complements OPC UA by adding agent identity, trust levels, and a simpler passport-based key management model.

17. Security Considerations

This entire document addresses security. Key residual risks:

- Side-channel attacks on ECDSA implementation: Mitigated by requiring constant-time implementations for SIL 3+.
- Gateway compromise: Mitigated by dual-gateway verification for SIL 2+ and hardware attestation.
- Denial of service via signature verification flood: Mitigated by rate limiting at the gateway and hardware acceleration.
- Supply chain compromise of agent firmware: Mitigated by requiring signed firmware updates and attestation of agent software integrity at passport issuance.
- Time synchronisation attacks: Mitigated by requiring NTP authentication (NTS, RFC 8915) and rejecting timestamps outside the configured window.
- Gateway key extraction via physical access: Mitigated by requiring TPM 2.0 or equivalent hardware key storage for Level 2+ conformance.
- Quantum computing: ECDSA P-256 is not quantum-resistant. Future revisions will define post-quantum signature profiles (ML-DSA / CRYSTALS-Dilithium) per NIST FIPS 204.
- Insider threat with valid L4 passport: Mitigated by behavioural anomaly detection (command rate, target patterns) and mandatory dual-authorisation for safety-critical commands in SIL 3+ environments.

18. IANA Considerations

This document requests the following IANA registrations:

- CoAP Option Numbers: 65001 (ATTP-Signature), 65003 (ATTP-Agent-ID), 65005 (ATTP-Trust-Level), 65007 (ATTP-Nonce), 65009 (ATTP-Timestamp)

(Section 5.4)

- MQTT User Property Names: X-ATTP-Version, X-ATTP-Signature, X-ATTP-Agent-ID, X-ATTP-Trust, X-ATTP-Nonce, X-ATTP-Timestamp, X-ATTP-Profile, X-ATTP-Passport (Section 5.3)

- ATTP-ICS Profile Registry: ics-modbus, ics-opcua, ics-mqtt, ics-coap (Section 5)

19. References

19.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997.

[RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017.

[RFC9431] Banks, A., Borgendale, E., Briggs, R., and K. Gupta, "MQTT Version 5.0", RFC 9431, DOI 10.17487/RFC9431, October 2024.

[draft-sharif-attp-agent-trust-transport] Sharif, R., "Agent Trust Transport Protocol (ATTP)", Internet-Draft draft-sharif-attp-agent-trust-transport, March 2026.

19.2. Informative References

[draft-sharif-agent-audit-trail] Sharif, R., "Agent Audit Trail Format", Internet-Draft draft-sharif-agent-audit-trail, March 2026.

[IEC-62443] International Electrotechnical Commission, "Industrial communication networks - IT security for networks and systems", IEC 62443, 2018.

[IEC-61508] International Electrotechnical Commission, "Functional safety of electrical/electronic/programmable electronic safety-related systems", IEC 61508, 2010.

[NIST-SP-800-82] Stouffer, K., Lightman, S., Pillitteri, V., Abrams, M., and A. Hahn, "Guide to Industrial Control Systems (ICS) Security", NIST SP 800-82 Rev. 3, September 2023.

[NIS2] European Parliament, "Directive (EU) 2022/2555 on measures for a high common level of cybersecurity across the Union (NIS2 Directive)", December 2022.

[OWASP-MCP] OWASP Foundation, "MCP Security Cheat Sheet", https://cheatsheetseries.owasp.org/cheatsheets/MCP_Security_Cheat_Sheet.html, 2026.

[RFC5424] Gerhards, R., "The Syslog Protocol", RFC 5424, DOI 10.17487/RFC5424, March 2009.

[RFC8915] Franke, D., Sibold, D., Dansarie, M., and S. Sundblad, "Network Time Security for the Network Time Protocol", RFC 8915, DOI 10.17487/RFC8915, September 2020.

[NERC-CIP] North American Electric Reliability Corporation, "Critical Infrastructure Protection Standards", NERC CIP-002 through CIP-014, 2024.

[IEC-61511] International Electrotechnical Commission, "Functional safety - Safety instrumented systems for the process industry sector", IEC 61511, 2016.

[STUXNET] Langner, R., "Stuxnet: Dissecting a Cyberwarfare Weapon", IEEE Security and Privacy, vol. 9, no. 3, pp. 49-51, May-June 2011.

[TRITON] Johnson, B., Caban, D., et al., "Attackers Deploy New ICS Attack Framework TRITON", FireEye Threat Intelligence, December 2017.

[OLDSMAR] Pinellas County Sheriff’s Office, "Oldsmar Water Treatment Facility Cyber Intrusion", February 2021.

[COLONIAL] CISA, "DarkSide Ransomware: Best Practices for Preventing Business Disruption from Ransomware Attacks", Alert AA21-131A, May 2021.

[IOCONTROL] Team82 (Claroty), "IOCONTROL: New Malware Targeting Critical Infrastructure IoT/OT Devices", December 2025.

Authors’ Addresses

Raza Sharif
CyberSecAI Ltd
London, United Kingdom

Email: contact@agentsign.dev
URI: https://cybersecai.co.uk