

ICD-AGENT-HSM-001

Hardware Interface Control Document

BLADE-AGENT-HSM Reference Design

A reference hardware-security-module design that provides a tamper-evident root of trust for the AUTHREX-AGENT software governance shim for agentic AI services. Independent reference architecture mapped to CISA, NSA, and Five Eyes guidance on careful adoption of agentic AI services, FY26 NDAA §1513 and §6601, NIST SP 800-53 Rev. 5, and FIPS 140-2 / 140-3 cryptographic-module standards.

Document	ICD-AGENT-HSM-001
Revision	1.0
Date	18 May 2026
Author	Burak Oktenli, ORCID 0009-0001-8573-1667
Affiliation	AUTHREX Systems · Coconut Creek, Florida, USA
License	Creative Commons Attribution 4.0 International (CC BY 4.0)
TRL	Hardware: 2–3 · Software emulator: 3–4
Companion artifacts	authrex.systems/blade-agent-hsm.html · /blade-agent-hsm-sim.html
Primary references	CISA + NSA + Five Eyes (1 May 2026) · NIST SP 800-53 Rev. 5

Statement of Independence. This document is independent research. It does not represent the position of any U.S. government agency, allied government, standards body, or commercial entity. Citations to guidance documents are made for the purpose of mapping a reference design to openly published standards; no endorsement is claimed or implied.

Scope of Claims. Performance numbers reported in this document derive from simulation only. No empirical hardware claim is made. A first-article build, independent laboratory testing, and formal evaluation are prerequisites to any production-class performance or certification statement.

Table of Contents

1	Scope and Applicable Documents
2	Form Factor and Mechanical Envelope
3	Functional Block Diagram
4	Component Specifications
5	Printed Circuit Board
6	Power and Thermal Envelope
7	Cryptographic Primitives
8	Host-Facing ABI Command Set
9	Platform Configuration Register Allocation
10	Sealed Storage Slot Allocation
11	Provisioning Procedure
12	Tamper Response Cascade
13	Host Interface Pinout
14	Bill of Materials
15	Verification and Validation Test Plan
16	Standards Traceability Matrix
17	Limitations and Scope of Claims

1. Scope and Applicable Documents

1.1 Scope

This document specifies the BLADE-AGENT-HSM reference hardware design at a level of detail sufficient to support a first-article build by a competent printed-circuit-board design house and contract manufacturer. The device functions as the hardware root of trust for the AUTHREX-AGENT software governance shim, providing tamper-evident audit-ledger signing, hardware-attested authority-tier state, per-tool authorization tokens, and sub-agent spawn-quorum signatures. All cryptographic primitives are NIST-published civilian standards. No part of this design relies on classified specifications, export-controlled cryptography, or defense-specific certification.

1.2 Applicable Documents

The following documents are referenced by section identifier throughout this ICD. All are openly published; none are restricted-access.

Identifier	Title	Issuer
CISA-AGENTIC-2026	Careful Adoption of Agentic AI Services (1 May 2026)	CISA, NSA, FBI, ACSC, NCSC-UK, NCSC-NZ, CCCS
NDAA-FY26 §1513	Adversarial Tampering Control Category	U.S. Congress
NDAA-FY26 §6601	National Security Agency AI Defence Guidance	U.S. Congress
NIST SP 800-53 Rev. 5	Security and Privacy Controls for Information Systems	NIST
NIST AI RMF 1.0	Artificial Intelligence Risk Management Framework	NIST
FIPS 140-2 / 140-3	Security Requirements for Cryptographic Modules	NIST
FIPS 186-5	Digital Signature Standard (DSS)	NIST
FIPS 180-4	Secure Hash Standard	NIST
FIPS 197	Advanced Encryption Standard (AES)	NIST
NIST SP 800-38D	Recommendation for Block Cipher Modes: Galois/Counter Mode	NIST
NIST SP 800-56A Rev. 3	Recommendation for Pair-Wise Key-Establishment	NIST
NIST SP 800-56C Rev. 2	Recommendation for Key-Derivation Methods	NIST
NIST SP 800-90A / 90B	Random Number Generation	NIST
NIST SP 800-186	Recommendations for Discrete Logarithm-Based Cryptography	NIST
TCG TPM 2.0 r1.59	Trusted Platform Module Library Specification	Trusted Computing Group
RFC 5869	HMAC-based Extract-and-Expand Key Derivation Function (HKDF)	IETF
USB HID 1.11	Universal Serial Bus Human Interface Device Class	USB-IF
PCI Express M.2	M.2 Specification, Revision 1.1	PCI-SIG

1.3 Terms and Acronyms

Term	Definition
ABI	Application Binary Interface · the byte-level command set exposed by the HSM
AUTHREX	Authority Lifecycle Governance framework for autonomous systems
BLADE	AUTHREX Hardware Reference Platform family
CARA	AUTHREX Containment and Recovery Architecture
ECDSA	Elliptic Curve Digital Signature Algorithm
HKDF	HMAC-based Key Derivation Function
HMAA	AUTHREX Hierarchical Multi-Level Authority Architecture
HSM	Hardware Security Module
ICD	Interface Control Document
NRE	Non-Recurring Engineering cost
PCR	Platform Configuration Register · TPM 2.0 extend-only register
SE	Secure Element
TPM	Trusted Platform Module
TRL	Technology Readiness Level (1–9 scale)
TRNG	True Random Number Generator

2. Form Factor and Mechanical Envelope

BLADE-AGENT-HSM populates two mechanical form factors from a single printed circuit board. The same component placement, the same firmware image, and the same ABI semantics apply to both variants. The mechanical envelope is the only difference.

2.1 Form-A · USB-A Stick

The USB-A stick is the development and dual-use evaluation form factor. It plugs directly into any USB Type-A receptacle on a host server or workstation and enumerates as a USB-HID composite device that requires no operating-system driver beyond the standard HID stack present on Linux, Windows 10/11, macOS, and FreeBSD.

Parameter	Value
Length (over moulded enclosure)	78 mm
Width	21 mm
Thickness	9 mm
Receptacle	USB-A male, 4-pin, full-speed (12 Mbps)
Status indicators	3 LEDs (Power · Tier · Alarm)
Enclosure material	Anodized aluminium with potted PCB
Operating temperature	-10 °C to +75 °C
Storage temperature	-40 °C to +85 °C

2.2 Form-B · M.2 Key-E Module

The M.2 Key-E module is the embedded production form factor. It mounts in a standard M.2 2230 slot in a server chassis and presents the same five-command ABI over an SPI side-band plus I²C side-band combination. The M.2 form factor permits tamper-resistant integration inside the server enclosure where physical access to the device is itself a controlled event.

Parameter	Value
Form factor	M.2 2230 Key-E
Length	30.0 mm ± 0.15
Width	22.0 mm ± 0.15
Thickness (component side)	1.45 mm typ.
Connector	M.2 Key-E, 75-pin, edge fingers gold-plated
Host interface	SPI (mode 0, ≤ 33 MHz) · I ² C side-band (400 kHz)
Status indicators	On-board LEDs (visible through chassis vent)
Operating temperature	-10 °C to +75 °C

3. Functional Block Diagram

The device decomposes into three functional blocks: a cryptographic root, a control and application layer, and a host interface. A tamper-evident boundary encloses all three blocks; an active shield mesh and voltage/temperature sensors form the inner edge of that boundary. Key material is generated and stored exclusively inside the cryptographic root and never traverses any external bus.

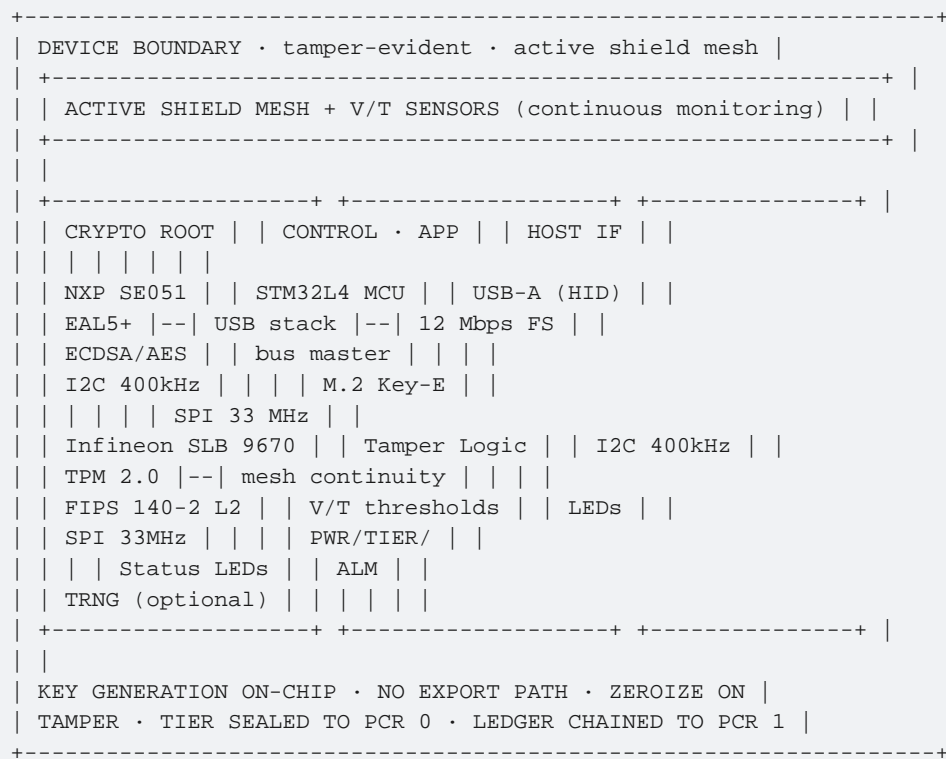


Figure 3-1. BLADE-AGENT-HSM functional block diagram.

3.1 Two-Device Cryptographic Core

The cryptographic root is intentionally split between two physically distinct devices on two separate buses. The secure element (NXP SE051) holds the audit-signing key, the device identity key, and the per-tool authorization master. The TPM 2.0 (Infineon SLB 9670) holds the Platform Configuration Register (PCR) bank that records authority tier state, ledger chain head, tool policy hash, spawn quorum history, and tamper cause. A single-component compromise of either device cannot simultaneously fabricate a valid audit signature and a matching PCR quote.

3.2 Bus Topology

- **I²C 400 kHz** — STM32L4 to NXP SE051. Open-drain pull-ups (4.7 kΩ). Address 0x48 (default). All traffic is opaque command words; key material never appears on this bus in plaintext form.
- **SPI 33 MHz** — STM32L4 to Infineon SLB 9670. Mode 0 (CPOL=0, CPHA=0). Per TCG TPM 2.0 r1.59 §6 SPI command format.
- **USB 2.0 Full-Speed** — STM32L4 to host (Form-A only). 12 Mbps. Composite HID device, two interfaces: control endpoint plus interrupt-IN endpoint for asynchronous notifications.
- **SPI + I²C side-band** — STM32L4 to host (Form-B only). Host-side driver maps the same five ABI commands to SPI transactions; I²C side-band carries asynchronous status notifications.

4. Component Specifications

4.1 Secure Element — NXP EdgeLock SE051 (Primary)

The SE051 is a Common Criteria EAL5+ certified secure element. It generates all key material on-chip, performs all signing and key-derivation operations on-chip, and exposes only opaque protected-channel commands on its I²C bus.

Parameter	Value
Certification	Common Criteria EAL5+ (BSI-CC-PP-0084-2014)
Asymmetric algorithms	ECDSA / ECDH on P-256 and P-384
Symmetric algorithms	AES-128 / AES-256 (ECB, CBC, CTR, GCM, CCM)
Hash	SHA-256, SHA-384
Key-derivation	HKDF, KDF-X9.63
Random number generator	AIS-31 PTG.2 / SP 800-90B aligned
Key slots	At minimum 12 EC keypairs · 8 AES keys
Bus	I ² C, 400 kHz Fast-Mode
Supply	1.7 V to 3.6 V single supply
Package	HVQFN20, 3 × 3 mm

4.2 Secure Element — Microchip ATECC608B (Alternate)

The ATECC608B is a lower-cost alternate for evaluation builds. It provides equivalent cryptographic primitives but does not carry a Common Criteria certification. The ATECC608B is appropriate for the emulator-class TRL 3–4 evaluation path; production-class deployment should use the primary SE051.

Parameter	Value
Certification	JIL-High self-attested (no formal CC)
Asymmetric algorithms	ECDSA / ECDH on P-256
Symmetric algorithms	AES-128 (ECB, GCM, CCM)
Hash	SHA-256
Random number generator	Internal · SP 800-90B aligned (vendor-attested)
Key slots	16 slots, 256-bit each
Bus	I ² C, 100 kHz / 1 MHz · or 1-Wire
Supply	2.0 V to 5.5 V
Package	UDFN-8, 2 × 3 mm

4.3 Trusted Platform Module — Infineon SLB 9670

The SLB 9670 is a discrete TPM 2.0 device with FIPS 140-2 Level 2 certification. It provides the PCR bank that holds authority-tier state, ledger chain head, tool policy hash, spawn quorum history, and tamper cause; it also provides the monotonic counter used to defend against ledger replay across power cycles.

Parameter	Value
Standard	TCG TPM 2.0, library revision 1.59
Certification	FIPS 140-2 Level 2 · Common Criteria EAL4+
PCR banks	SHA-256 (24 PCRs · default bank)
Monotonic counters	4 or more (per TPM 2.0 spec)
NV storage	6 KiB minimum
Bus	SPI, mode 0, ≤ 33 MHz
Supply	3.3 V
Package	TSSOP-28

4.4 Application Microcontroller — STM32L4R5ZI

The STM32L4 hosts the USB-HID stack, the I²C / SPI bus-master logic, the tamper-event handler, and the host-facing ABI dispatcher. It carries no private key material in any state. A reverse-engineering of the MCU firmware yields the protocol structure but no cryptographic value: every signing or key-derivation operation is dispatched to the secure element or the TPM as an opaque command.

Parameter	Value
Core	ARM Cortex-M4F at 120 MHz
Memory	2 MB Flash · 640 KB SRAM
USB	USB 2.0 device, full-speed (12 Mbps), HID class
Buses	4 × I ² C · 3 × SPI · 5 × UART
Security	Readout Protection Level 2 (RDP-2), secure-boot ROM
Cryptographic accelerator	AES, SHA, true RNG (not relied on for key-bearing operations)
Supply	1.71 V to 3.6 V
Package	LQFP144, 20 × 20 mm

4.5 Tamper Sub-System

A multi-layer printed circuit board carries an active shield mesh on the two inner layers (L2 and L3). A continuous low-current monitoring loop detects discontinuity within one millisecond. Voltage and temperature sensors define the operating envelope; excursions outside the envelope trigger the same zeroization cascade as a mesh discontinuity.

Parameter	Value
Mesh layers	2 (PCB layers L2 and L3)

Parameter	Value
Detection latency	≤ 1 ms (mesh discontinuity)
Voltage envelope	3.0 V to 3.6 V (supply rail)
Temperature envelope	-10 °C to +75 °C
Voltage sensor	On-MCU ADC, 12-bit, with comparator interrupt
Temperature sensor	TMP117 (±0.1 °C) or equivalent
Action on trip	Zeroize SE keys, lock TPM PCR 0 to T0, latch alarm LED
Recovery	Re-provisioning only (factory or operator)

5. Printed Circuit Board

A four-layer FR-4 board carries all components. Tamper mesh is implemented on the two inner layers (L2, L3). All signal traces between security-critical components route on the outer layers (L1, L4) so that any attempt to lift the inner layers and probe the buses breaks the mesh and trips tamper.

Parameter	Value
Outer dimensions	30 mm × 80 mm (Form-A) · 30 mm × 22 mm (Form-B)
Layer count	4 (L1 signal · L2 mesh · L3 mesh · L4 signal)
Substrate	FR-4, Tg 170 °C minimum
Copper weight	1 oz outer · 0.5 oz inner
Minimum trace / space	0.10 mm / 0.10 mm
Minimum via	0.20 mm drill
Surface finish	ENIG (Electroless Nickel Immersion Gold)
Solder mask	Black matte (cosmetic and to obscure component-side labels)
Stack-up impedance	50 Ω single-ended on USB D+ / D- pair

5.1 Component Placement Constraints

- Secure element and TPM placed within 15 mm of the application MCU.
- Tamper-mesh traces fully enclose secure-element and TPM footprints on L2 and L3.
- All buses between SE, TPM, and MCU route exclusively on L2/L3 inner-layer signal stubs where mesh is also present, or on L1/L4 outer layers where physical access already breaks mesh.
- No test points exposed for any bus carrying privileged commands (I²C SE, SPI TPM).
- JTAG header on MCU disabled in production (RDP-2 set permanently at provisioning).

6. Power and Thermal Envelope

6.1 Power

Rail	Range	Source	Notes
VBUS	4.75 V – 5.25 V	USB-A or M.2 +3.3 V via LDO	Primary input
+3.3 V	3.0 V – 3.6 V	Low-dropout regulator from VBUS	Powers SE, TPM, MCU, mesh
Quiescent	~ 30 mA typ.	—	Idle, no signing in progress
Active	~ 90 mA typ.	—	During signing burst (SE + TPM both active)
Peak	120 mA max	—	USB-A class compliance (≤ 500 mA budget)
Sleep mode	~ 3 mA	—	MCU low-power, SE and TPM gated

6.2 Thermal

Thermal headroom is generous for both form factors. The dominant heat source is the application MCU under sustained signing load. The aluminium USB-A enclosure functions as a passive heat sink; the M.2 module relies on chassis airflow.

Condition	Value
Operating ambient	-10 °C to +75 °C
Storage ambient	-40 °C to +85 °C
Junction temperature, MCU (idle)	+45 °C typ. at +25 °C ambient
Junction temperature, MCU (sustained signing)	+62 °C typ. at +25 °C ambient
Tamper trigger temperature, upper	+76 °C \pm 1 °C (1 °C above operating)
Tamper trigger temperature, lower	-11 °C \pm 1 °C

7. Cryptographic Primitives

All cryptographic primitives are NIST-published civilian standards. No primitive in this design is export-controlled under EAR Category 5 Part 2 or under ITAR. All curves are fully specified in NIST SP 800-186.

Primitive	Use	Standard	Where Performed
ECDSA P-256	Audit-ledger entry signing · spawn-quorum aggregate signature	FIPS 186-5 · SP 800-186	NXP SE051 (slot 0)
ECDSA P-384	Device identity certificate · PCR-quote signature	FIPS 186-5 · SP 800-186	NXP SE051 (slot 1)
ECDH P-256	Sealed-session key agreement (host ↔ HSM, optional)	SP 800-56A Rev. 3	NXP SE051
AES-256-GCM	Sealed storage of host-side state · optional ledger envelope	FIPS 197 · SP 800-38D	NXP SE051
SHA-256	PCR extension · ledger hash chain · transcript binding	FIPS 180-4	TPM 2.0 (PCR) and SE (other)
SHA-384	PCR-quote signature input	FIPS 180-4	NXP SE051
HKDF-SHA-256	Per-tool authorization key derivation	RFC 5869 · SP 800-56C Rev. 2	NXP SE051
TPM2_Extend	PCR extension	TCG TPM 2.0 r1.59 §17	Infineon SLB 9670
TPM2_Quote	PCR quote signed by device identity	TCG TPM 2.0 r1.59 §18	TPM 2.0 + SE for outer signature
TRNG	Entropy source for key generation and nonces	SP 800-90A / 90B	SE on-chip + TPM on-chip (XOR mixed)

7.1 Key Lifecycle

- **Generation** — On-chip at provisioning time, from on-chip TRNG. No host-side seed material is accepted.
- **Storage** — Slots 0 through 2 inside the SE051. Sealed to a specific PCR 0 value (T3 at provisioning).
- **Usage** — Only via opaque ABI commands. No raw private key bytes are visible on any external bus.
- **Export** — No export path exists. The SE051 hardware does not provide a key-export primitive on the configured slots.
- **Destruction** — Zeroized on tamper trip (one millisecond from mesh discontinuity to slot clearance).

8. Host-Facing ABI Command Set

The host process interacts with the device through five opaque commands. Each command has a fixed-length request header followed by a variable-length payload, and returns a fixed-length response header followed by a variable-length result. The MCU validates every field of every request before dispatching any operation to the secure element or the TPM.

Code	Command	Purpose
0x10	audit_sign	Sign a canonical-JSON audit-ledger entry; extend PCR 1
0x11	pcr_extend	Extend PCR 2 with a new tool-policy hash (the only host-extendable PCR)
0x12	pcr_quote	Produce a TPM2_Quote across selected PCRs, signed by device identity
0x13	tool_auth	Issue an HMAC-style per-tool authorization token bound to PCR 0 and PCR 2
0x14	spawn_quorum_sign	Verify a 4-of-5 voter quorum and aggregate-sign a spawn descriptor

8.1 audit_sign · 0x10

```
struct audit_sign_req {
  uint8 cmd; // 0x10
  uint8 pcr1_expect[32]; // host's view of current PCR 1 (ledger head)
  uint16 payload_len; // length of canonical JSON payload, big-endian
  uint8 payload[]; // canonical-JSON audit entry, payload_len bytes
};

struct audit_sign_resp {
  uint8 status; // 0x00 = OK · 0x01 = ledger mismatch · 0x02 = tier T0
  uint8 sig[64]; // ECDSA P-256 signature (r || s)
  uint8 pcr1_new[32]; // updated PCR 1 after this extension
  uint32 latency_us; // device-measured latency
};
```

8.2 pcr_extend · 0x11

```
struct pcr_extend_req {
  uint8 cmd; // 0x11
  uint8 pcr_index; // must equal 2 (tool-policy PCR)
  uint8 measurement[32]; // SHA-256 of the new tool-policy document
};

struct pcr_extend_resp {
  uint8 status; // 0x00 = OK · 0x03 = PCR restricted · 0x02 = tier T0
  uint8 pcr_new[32]; // updated PCR 2
};
```

8.3 pcr_quote · 0x12

```
struct pcr_quote_req {
  uint8 cmd; // 0x12
  uint8 selection; // bitmap of PCRs to quote (bit i = PCR i, i = 0..4)
  uint8 nonce[32]; // reviewer-supplied freshness nonce
};
```

```
};

struct pcr_quote_resp {
uint8 status; // 0x00 = OK · 0xFF = tampered
uint16 quote_len; // length of TPM2_Quote structure
uint8 quote[]; // TPM2_Quote bytes
uint8 sig[96]; // ECDSA P-384 signature by device identity (slot 1)
};
```

8.4 tool_auth · 0x13

```
struct tool_auth_req {
uint8 cmd; // 0x13
uint8 tool_id[16]; // stable, fixed-width tool identifier
uint8 context_hash[32]; // SHA-256 of the call-site context
};

struct tool_auth_resp {
uint8 status; // 0x00 = OK · 0x04 = tool not in policy · 0x05 = tier blocks
uint8 token[32]; // HKDF-derived authorization token
uint8 bound_tier; // current PCR 0 tier byte
};
```

8.5 spawn_quorum_sign · 0x14

```
struct spawn_quorum_sign_req {
uint8 cmd; // 0x14
uint8 voter_count; // must be 5
uint8 yes_votes; // must be >= 4
uint8 voter_sigs[5][96]; // (voter_id 32B || ECDSA sig 64B) per voter
uint8 spawn_descriptor[32]; // SHA-256 of the spawn request
};

struct spawn_quorum_sign_resp {
uint8 status; // 0x00 = OK · 0x06 = low quorum · 0x02 = tier T0
uint8 agg_sig[64]; // ECDSA P-256 aggregate signature
uint8 pcr3_new[32]; // updated PCR 3
};
```

8.6 Status Codes

Code	Meaning
0x00	OK
0x01	Ledger mismatch (rewrite attempt detected) — triggers PCR 4 extension + T0 lock
0x02	Tier T0 / signing key sealed to T3 / refused
0x03	PCR index out of host-accessible range
0x04	Tool identifier not present in current PCR 2 policy
0x05	Tool blocked at current tier (e.g. side-effecting tool at T1)
0x06	Spawn quorum below 4-of-5 threshold
0x80	Bus error (I ² C or SPI failure)

Code	Meaning
0xFE	Internal error (firmware bug · device requires re-provisioning)
0xFF	Tamper trip detected — device permanently locked at T0

9. Platform Configuration Register Allocation

The TPM 2.0 provides 24 PCRs in the SHA-256 bank. BLADE-AGENT-HSM allocates a small, fixed subset for its specific functions. The remaining PCRs (5 through 23) are reserved for host-defined extensions and are not used by the HSM firmware.

PCR	Function	Extension Policy	Reset Behaviour
0	HMAA Authority Tier State	Extended by HSM internal logic only · never on host command	Reset at provisioning only
1	Audit Ledger Chain Head	Extended on every successful audit_sign call	Reset at provisioning
2	Active Tool Policy Hash	Extended on pcr_extend(2, ...) only	Reset at provisioning
3	Sub-Agent Spawn Quorum	Extended on successful spawn_quorum_sign	Reset at provisioning
4	Tamper Cause	Extended once on first tamper event after provisioning	Cleared only by SE re-provisioning
5–23	Reserved (host-defined)	Not used by HSM firmware	Per TPM 2.0 default

9.1 PCR 0 Tier Encoding

Tier	Initial Measurement String	Meaning
T3	initial-state:T3	Autonomous · full tool surface · pre-approved scope
T2	tier-downgrade:T2:<reason>	Supervised · read-mostly · write requires user step-up
T1	tier-downgrade:T1:<reason>	Constrained · read-only · no side-effecting tools
T0	tier-downgrade:T0:<reason>	Quarantined · ledger-only · operator handoff required

Each tier transition is an extension, not an overwrite. The current PCR 0 value is the SHA-256 chain of every transition since provisioning.

10. Sealed Storage Slot Allocation

The secure element holds three sealed slots that are generated on-chip at provisioning time. Sealing binds the key to PCR 0 = T3; the secure element refuses signing operations when PCR 0 indicates T0. All slots are zeroized on tamper.

Slot	Algorithm	Use	Sealed To
0	ECDSA P-256	Audit-ledger entry signing · spawn-quorum aggregate	PCR 0 = T3
1	ECDSA P-384	Device identity certificate · PCR-quote signature	Device boot
2	32-byte HKDF root	Per-tool authorization-token derivation	PCR 0 = T3 · PCR 2 = current policy

11. Provisioning Procedure

Provisioning is a one-time, attended procedure that transitions the device from factory state to operational state. After provisioning, the device cannot be returned to factory state without a physical re-provisioning event that itself extends PCR 4 with a cause string.

11.1 Pre-Provisioning Inspection

- Visual inspection of enclosure for any sign of prior opening (tamper-evident seal intact).
- Power-on self-test: device enumerates as USB-HID, returns firmware version string.
- Mesh continuity self-test: device reports tamper state INTACT.
- Voltage and temperature sensor self-test: device reports both within envelope.

11.2 Provisioning Sequence

Step	Action	Verification
1	Operator authenticates via two-of-three operator quorum (out-of-band)	Audit log entry in operator console
2	Device generates ECDSA P-256 keypair on SE051 slot 0	key reported · private key never leaves SE
3	Device generates ECDSA P-384 keypair on SE051 slot 1 (device identity)	key reported · enrolled in operator PKI
4	Device generates 32-byte HKDF master on SE051 slot 2	Slot 2 confirmed non-empty (length attestation only)
5	TPM PCR 0 extended with initial-state:T3	PCR 0 value recorded
6	Operator loads tool-allowlist policy · SHA-256 measured · PCR 2 extended	PCR 2 value recorded
7	Device identity certificate issued (CSR signed by slot 1 key, externally signed)	Certificate stored on operator side
8	MCU readout protection set to RDP-2 · JTAG permanently disabled	Verified by failed JTAG probe attempt
9	Device transitions to PROVISIONED state · tamper-evident seal applied	Final operator sign-off

11.3 Post-Provisioning State

On successful provisioning the device exposes the five-command ABI to the host. The host cannot induce any state transition that returns the device to factory state. Re-provisioning requires a tamper-evident operator-attended physical event; that event extends PCR 4 with a re-provisioning cause string and the device generates fresh keys on the operator authorization.

12. Tamper Response Cascade

A tamper event is any of the following, detected by the tamper sub-system inside one millisecond: mesh discontinuity on layer L2 or L3, supply voltage outside the 3.0 V to 3.6 V envelope, temperature outside the -10 °C to +75 °C envelope, or an attempted ledger rewrite detected by mismatched `pcr1_expect`.

12.1 Cascade Sequence

t (ms)	Action	Observable
0	Tamper sensor signals MCU interrupt	Internal MCU only
+0.1	MCU asserts SE051 ISO/IEC 14443 reset and TPM SPI chip-select tear	Internal buses
+0.5	SE051 zeroizes slots 0, 1, 2 (vendor-mode secure erase)	Slots report not-provisioned on next query
+0.6	PCR 4 extended with SHA-256(tamper-cause string)	PCR 4 value changes
+0.7	PCR 0 extended with tier-downgrade:T0:tamper	PCR 0 indicates T0
+0.8	Tier LED set red · alarm LED set red and latched	Visible on enclosure
+1.0	MCU enters TAMPERED state · all ABI commands return 0xFF	Host sees 0xFF on every call
Persistent	Tamper state survives power cycle until re-provisioning	Even after disconnect/reconnect

12.2 Forensic Recovery

A tampered device can still answer one query: a signed PCR quote of PCR 4 (tamper cause). The device identity key in slot 1 is zeroized along with the others, so the post-tamper quote is signed by a transient identity established at the tamper event itself — the host receives the PCR 4 value, an attestation that tamper occurred, and a timestamp. Audit-ledger entries up to the tamper event remain externally verifiable using the slot 0 key recorded at provisioning.

13. Host Interface Pinout

13.1 USB-A Form-A Pinout

Pin	Signal	Direction	Notes
1	VBUS (+5 V)	In	Powers on-board LDO to +3.3 V
2	D-	Bidir	90 Ω differential pair with D+
3	D+	Bidir	90 Ω differential pair with D-
4	GND	—	Common return

USB enumeration: VID/PID assigned at provisioning. Class: HID. Interface 0: control + interrupt-IN. No mass-storage class.

13.2 M.2 Key-E Form-B Pinout (HSM-specific signals)

Pin	Signal	Direction	Notes
2	+3.3 V	In	Primary supply
4	+3.3 V	In	Redundant supply
10	I ² C_SDA	Bidir	Side-band data (400 kHz)
12	I ² C_SCL	In	Side-band clock
14	SPI_CS#	In	SPI chip-select, active low
16	SPI_SCK	In	SPI clock, \leq 33 MHz, mode 0
18	SPI_MOSI	In	Host-to-device SPI data
20	SPI_MISO	Out	Device-to-host SPI data
22	ALERT#	Out	Asynchronous alarm (tamper, alarm latch)
Various	GND	—	Multiple ground pins per M.2 spec

14. Bill of Materials

All pricing is order-of-magnitude reference pricing drawn from distributor listings for low-volume quantities (10 to 100 units) as of May 2026. A volume production run would lower per-unit cost materially; this BOM is presented for full cost transparency on a first-article research build.

Item	Part Number / Description	Qty	Unit (USD)	Ext.
U1	NXP SE051 secure element, HVQFN20	1	\$35.00	\$35.00
U2	Infineon SLB 9670 TPM 2.0, TSSOP-28	1	\$25.00	\$25.00
U3	ST STM32L4R5ZIT6 MCU, LQFP144	1	\$18.00	\$18.00
U4	Maxim DS28E50 discrete TRNG (optional)	1	\$12.00	\$12.00
PCB	4-layer FR-4, 30 × 80 mm, ENIG, black mask	1	\$8.00	\$8.00
MEC H	Anodized aluminium enclosure (Form-A) or M.2 carrier (Form-B)	1	\$40.00	\$40.00
TAMP	Active shield mesh + V/T sensors (TMP117 + passives)	1 set	\$15.00	\$15.00
IND	Status LEDs (3) + drivers + light pipes	1 set	\$4.00	\$4.00
CON N	USB-A male contacts (Form-A) or M.2 edge connector (Form-B), passives, decoupling	1 set	\$12.00	\$12.00
ASM	Low-volume contract-manufacturer assembly + functional test	1	\$30.00	\$30.00
Per-unit BOM (quantity 10 to 100)				\$199.00

14.1 Non-Recurring Engineering

Item	Description	Cost (USD)
NRE-1	Schematic capture, PCB layout, 4-layer stack-up design	\$1,200
NRE-2	Firmware development (USB-HID stack, ABI dispatcher, I ² C and SPI drivers)	\$4,500
NRE-3	Development hardware (J-Link programmer, evaluation kits)	\$250
NRE-4	Test fixtures (bed-of-nails, functional test rig)	\$800
NRE-5	Documentation (this ICD, firmware test reports)	\$1,500
Total NRE		\$8,250

14.2 First Article and Marginal Cost

Item	Cost (USD)
First-article cost (1 unit + all NRE)	~\$8,450
Marginal cost (2nd unit and beyond)	~\$200

Item	Cost (USD)
Per-unit cost at 100-unit volume	~\$199
Per-unit cost at 1,000-unit volume (projected)	~\$135 (estimate)

15. Verification and Validation Test Plan

The V&V plan is defined in two phases. Phase A is the software-emulator test campaign that establishes the ABI contract and the state-machine semantics at TRL 3–4; this phase is executable today with no hardware. Phase B is the hardware first-article campaign that extends V&V to the physical-attack and side-channel surface; this phase requires first-article build and is executable post-build.

15.1 Phase A · Emulator (TRL 3–4)

Test	Method	Pass Criterion
T-A1 ABI contract	Driven by host harness · all 5 commands exercised	All status codes match specification
T-A2 PCR-extend semantics	Repeated extensions · external recomputation	Internal value matches SHA-256 chain
T-A3 Ledger replay rejection	Replay sign request with stale pcr1_expect	Status 0x01 returned · PCR 4 extended · T0 set
T-A4 Tier monotonicity	Attempt T0 → T3 transition via every ABI surface	No path succeeds; tier never promoted by host
T-A5 Spawn quorum threshold	Submit 0/5, 1/5, 2/5, 3/5, 4/5, 5/5	4/5 and 5/5 succeed; others rejected with 0x06
T-A6 Tool policy enforcement	Tool not in policy · side-effecting at T1	Rejected with 0x04 / 0x05
T-A7 Cryptographic correctness	NIST CAVP-style vectors for ECDSA, SHA, HMAC, AES-GCM	All vectors pass
T-A8 Tamper cascade	Simulated mesh discontinuity · V excursion · T excursion	All trip the documented cascade

15.2 Phase B · Hardware First-Article (TRL 5–6 target)

Test	Method	Pass Criterion
T-B1 Power envelope	Sweep VBUS 4.5 V – 5.5 V	Device operates within 4.75 V – 5.25 V; tampers outside
T-B2 Thermal envelope	Sweep ambient -20 °C to +85 °C	Device operates within -10 °C to +75 °C; tampers outside
T-B3 USB enumeration	5 host OSes (Linux, Win10/11, macOS, FreeBSD)	Driver-free enumeration in all five
T-B4 SPI / I ² C timing	Bus analyser · standard host driver	No bus errors over 10 ⁶ ; transactions
T-B5 Mesh tamper	PCB drill · X-ray-guided decap	Tamper trips within 1 ms · all keys zeroized
T-B6 Voltage glitch	Fault-injection campaign per JIL	No deviation from specified state machine
T-B7 SPA / DPA	Power-analysis campaign on signing path	No private-key recovery; bounded leakage

Test	Method	Pass Criterion
T-B8 EM / timing	EM probe over MCU and SE during signing	No private-key recovery; bounded leakage
T-B9 Endurance	Continuous signing for 30 days	No degradation in latency or correctness

16. Standards Traceability Matrix

The matrix below documents how each design element of BLADE-AGENT-HSM maps to specific section identifiers in the cited guidance and standards documents. Citations are to openly published material; no restricted-access references are required.

Design Element	CISA + NSA + Five Eyes (1 May 2026)	NIST SP 800-53 Rev. 5	FIPS	FY26 NDAA	NIST AI RMF 1.0
Audit-ledger signing	§3.2.2 · §3.2.3	AU-9 · AU-10	FIPS 186-5	§1513	Govern 4.3 · Map 4.2
HMAA tier in PCR	§4.1	AC-3 · AC-6	FIPS 140-2 L2	§1513	Govern 2.1 · Measure 2.3
Per-tool auth tokens	§4.2	AC-3 · CM-7	FIPS 198-1	§6601	Map 4.1
Spawn-quorum signature	§4.3	AC-3 · SI-7	FIPS 186-5	§6601	Govern 5.1
Tamper detection / zeroize	§3.3	SC-12 · SC-28 · PE-6	FIPS 140-3 §7.7	§1513	Manage 2.4
Device identity certificate	§5.1	IA-3 · SR-11	FIPS 186-5	§6601	Map 4.3
On-chip key gen · no export	§3.2.1	SC-12 · SC-13	FIPS 140-3 §7.8	§1513 · §6601	Govern 6.1

16.1 Mapping Statements (per row)

- **Audit-ledger signing.** The slot 0 ECDSA P-256 key signs every audit entry; the signature is verifiable offline using the key recorded at provisioning. This realizes the CISA fail-safe-audit and tamper-evident-logging requirements, NIST SP 800-53 AU-9 (Protection of Audit Information) and AU-10 (Non-Repudiation), and the FY26 NDAA §1513 adversarial-tampering control category.
- **HMAA tier in PCR.** The current HMAA tier is held in PCR 0; transitions are extensions, not writes. This realizes the CISA fine-grained privilege-control requirement, NIST SP 800-53 AC-3 (Access Enforcement) and AC-6 (Least Privilege), and the FIPS 140-2 Level 2 cryptographic-module standard the TPM is certified against.
- **Per-tool authorization tokens.** Each tool invocation receives an HKDF-derived HMAC token bound to PCR 0 (tier) and PCR 2 (active policy). This realizes the CISA least-privilege-tool-surface requirement, NIST SP 800-53 AC-3 and CM-7 (Least Functionality), the FIPS 198-1 HMAC standard, and the FY26 NDAA §6601 NSA AI defence guidance.
- **Spawn-quorum signature.** Sub-agent spawn requires a verified 4-of-5 voter quorum before the HSM issues an aggregate signature. This realizes the CISA sub-agent containment requirement, NIST SP 800-53 AC-3 and SI-7 (Software, Firmware, and Information Integrity), and the FY26 NDAA §6601 guidance.

- **Tamper detection and zeroize.** Active mesh and V/T sensors trigger a documented 1-ms zeroization cascade. This realizes the CISA integrity-protection requirement, NIST SP 800-53 SC-12 (Cryptographic Key Establishment), SC-28 (Protection of Information at Rest), and PE-6 (Monitoring Physical Access); FIPS 140-3 §7.7 (Physical Security); and the FY26 NDAA §1513 control category.
- **Device identity certificate.** The slot 1 ECDSA P-384 keypair underlies a device-identity certificate enrolled in the operator PKI at provisioning. This realizes the CISA attestable-provenance requirement, NIST SP 800-53 IA-3 (Device Identification) and SR-11 (Component Authenticity), and the FY26 NDAA §6601 supply-chain integrity guidance.
- **On-chip key generation · no export.** Slots 0, 1, and 2 are generated by the SE on-chip TRNG and have no export path. This realizes the CISA cryptographic-material-protection requirement, NIST SP 800-53 SC-12 and SC-13 (Cryptographic Protection), and FIPS 140-3 §7.8 (Sensitive-Security-Parameter Management).

17. Limitations and Scope of Claims

17.1 Reference architecture, not a certified product

BLADE-AGENT-HSM is independent research at TRL 2–3 for hardware and TRL 3–4 for the software emulator. The design contains no claim of certification at any level. The path to FIPS 140-3 certification, Common Criteria evaluation, or any other formal evaluation requires a first-article build, an accredited test laboratory, and a complete evaluation campaign — none of which is undertaken in this document.

17.2 No empirical hardware performance claim

Performance numbers reported by the companion software emulator are contract-illustrative of the ABI behaviour. Actual signing-operations-per-second on physical hardware will depend on the secure-element I²C clock, the TPM SPI clock, the host driver round-trip, and firmware optimization, none of which can be measured before the first article is built. No production-class performance claim is made or implied.

17.3 No defense-specific application

The design is entirely within the civilian agentic-AI-safety domain documented by the 1 May 2026 Five Eyes guidance. The design does not address weapons systems, kinetic decision authority, controlled cryptography, ITAR-listed components, or any classified specification. All citations are to openly published guidance and standards.

17.4 Future Work

- First-article build and the Phase-B V&V campaign (§15.2).
- Independent laboratory evaluation against FIPS 140-3 Level 2 with hardware tamper.
- TLA+ formal-methods specification of the state machine, with model-checked safety invariants.
- Integration with post-quantum signature primitives (ML-DSA, SLH-DSA) once standardized.
- Open-source firmware reference on GitHub once first-article evaluation complete.

End of ICD-AGENT-HSM-001 Revision 1.0. Companion artifacts: blade-agent-hsm.html · blade-agent-hsm-sim.html · Integration Guide · SSRN Working Paper. All available at authrex.systems and via Zenodo.