MYTHOS SYSTEMS

PERMANENT LIBRARY PUBLICATION

ENGINEERING STANDARDS LIBRARY / CYBERSECURITY AND NETWORK SECURITY

Confidential Computing and Trusted Execution Environment Standard

Defines TEEs, attestation, threat models, secrets, deployment, evidence, and limitations.

Response, and Assurance Evidence

PERMANENT DOCUMENT ID
MYTHOS-SEC-0009

PUBLICATION
Candidate Standard

ASSURANCE SURFACE
7 Atomic Criteria / 4 Families

PERMANENT PUBLICATION IDENTITY

Confidential Computing and Trusted Execution Environment Standard

Defines TEEs, attestation, threat models, secrets, deployment, evidence, and limitations.

DOCUMENT ID
MYTHOS-SEC-0009

VERSION
1.0.0-candidate

STATUS
Candidate Standard

PUBLISHER
Mythos Systems

PUBLICATION DATE
2026-07-23

SCHEDULED REVIEW
2027-01-23

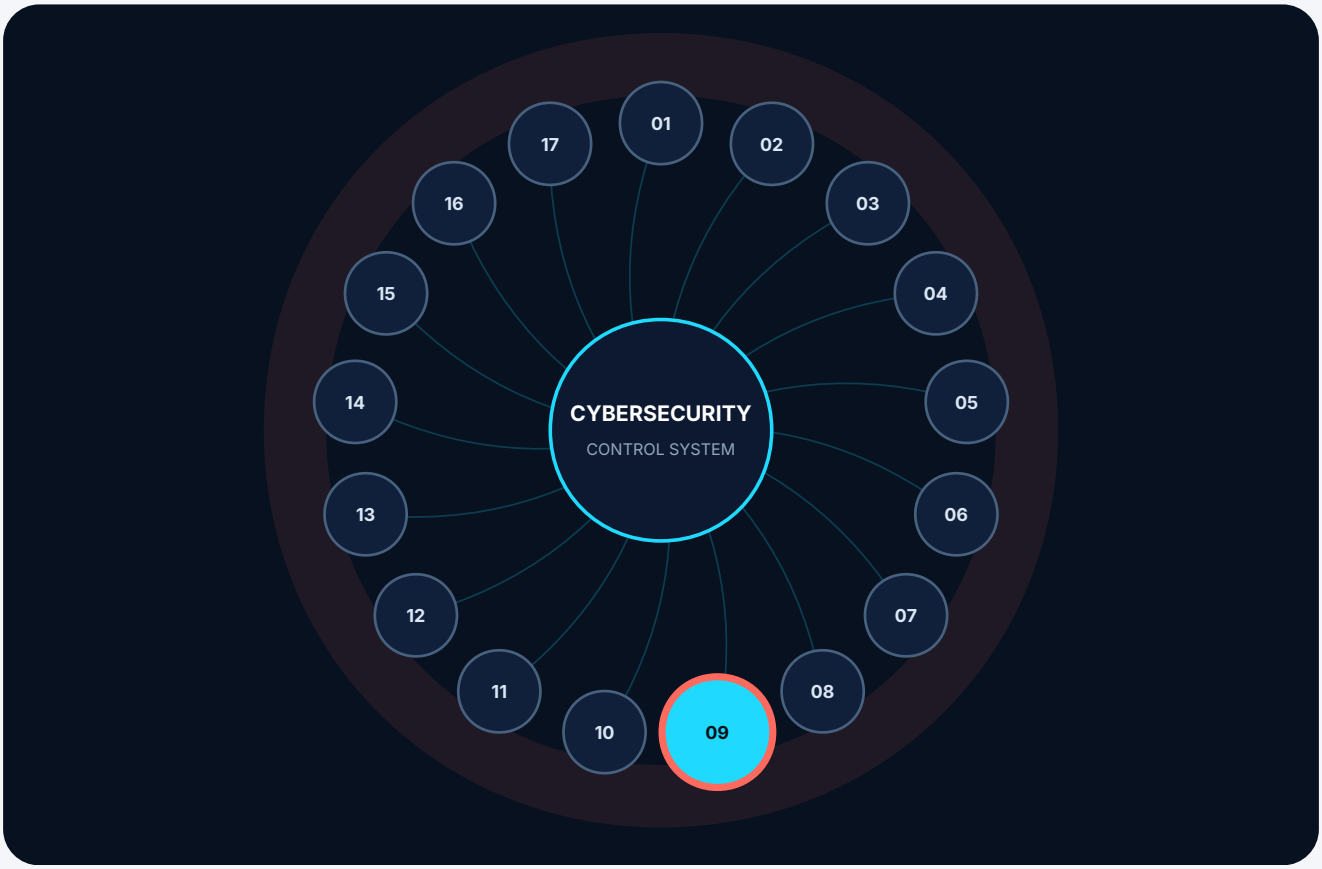
CLASSIFICATION
Public

7	4	4	1
ATOMIC CRITERIA	CAPABILITY FAMILIES	ASSESSMENT PROFILES	PERMANENT IDENTITY

Publication doctrine. This standard establishes durable requirements for its defined scope. Interpretation must preserve the permanent document identity, recorded evidence, and formal revision history.

DOMAIN CONTROL SYSTEM

MYTHOS-SEC-0009 inside the cybersecurity and network security standard constellation



Connected assurance

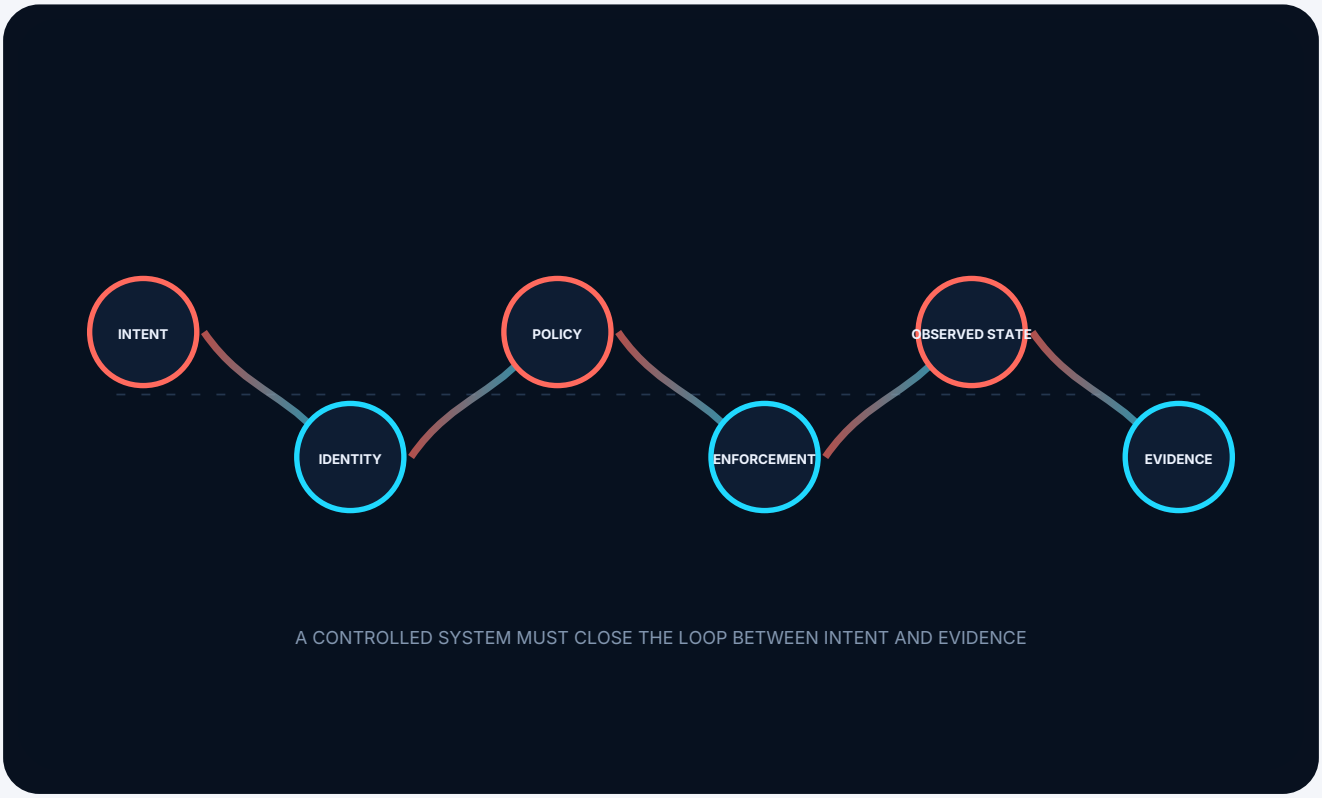
Architecture, policy, implementation, operations, and evidence are evaluated as one controlled system.

Specialization rule

Companion standards may add precision, but they cannot silently weaken this publication's requirements.

CLOSED-LOOP ARCHITECTURE

Intent becomes trustworthy only when observed state returns as evidence



State separation

Desired, approved, deployed, observed, historical, and simulated states must remain distinguishable.

Recoverable change

Material change requires validation, authorization, monitoring, containment, and a credible recovery path.

EVALUATION TOPOLOGY

The capability families and atomic criteria of MYTHOS-SEC-0009

Capability families

Families expose the major engineering surfaces and preserve the source weighting model.

Atomic criteria

Each criterion carries a question, maturity spectrum, evidence expectations, assessment method, and operational signals.

NAVIGATION

Contents

ENGINEERING DOCTRINE

0. Executive Mandate

The protection of AI data has failed at the final frontier: Data in Use.

Organizations encrypt data at rest (AES-256) and in transit (TLS 1.3). But the moment an LLM processes a prompt, that prompt - and the model's proprietary weights - are decrypted into plaintext in system RAM. The cloud provider, the hypervisor, and malicious insiders can read or modify this data. For sovereign AI, this is an unacceptable compromise.

This standard exists because:

1. **Data in Use is the Unprotected Frontier:** Standard cryptography cannot protect data while it is being computed. Confidential Computing - using Trusted Execution Environments (TEEs) like Intel SGX, AMD SEV-SNP, and AWS Nitro Enclaves - is the only architectural solution to isolate memory and CPU registers from the host OS.
2. **AI Agents are High-Value Targets:** AI agents process proprietary code, infrastructure credentials, and sensitive user context. If an agent's memory is compromised, its entire capability grant and secret lease is exfiltrated. Agent memory MUST be enclave-protected.
3. **Model Weights are Crown Jewels:** Fine-tuned models represent millions of dollars in compute and proprietary data. Deploying them to untrusted cloud infrastructure is an IP exfiltration risk. Model inference MUST occur within a TEE.
4. **Attestation is Mandatory:** You cannot trust a cloud provider's claim that a TEE is secure. The system MUST cryptographically verify the hardware, firmware, and application state via Remote Attestation before transmitting any data.
5. **FHE is the Ultimate Sovereignty:** Fully Homomorphic Encryption (FHE) allows computation on encrypted data without ever decrypting it. While computationally expensive, FHE is the gold standard for sovereign AI, allowing models to process data without ever exposing the plaintext.

This document establishes the **Mythos Confidential Computing Standard (MCCS)**, a comprehensive, evidence-based methodology for evaluating hardware-backed isolation, remote attestation, and encrypted computation for AI workloads.

ENGINEERING DOCTRINE

Part I. Scope and Applicability

1.1 In Scope

This standard covers the evaluation of:

- Trusted Execution Environments (TEEs) and secure enclaves (Intel SGX/TDX, AMD SEV-SNP, AWS Nitro)
- Confidential computing for AI inference and model deployment
- Remote attestation protocols and cryptographic verification
- Memory encryption and hypervisor isolation
- Fully Homomorphic Encryption (FHE) and Zero-Knowledge Proofs (ZKPs)
- AI agent memory isolation and secret protection
- Secure multi-party computation (MPC)

1.2 Out of Scope

- General cryptographic agility (covered in MYTHOS-SEC-0001)
- Software-level sandboxing (covered in MYTHOS-AI-0001)

1.3 Audience

- Security engineers and architects designing enclave-based AI pipelines
- Principal engineers selecting confidential computing platforms
- Cloud infrastructure teams managing TEE fleets
- AI governance, risk, and compliance teams
- Standards bodies seeking a reference TEE methodology

ENGINEERING DOCTRINE

Part II. Definitions and Taxonomy

2.1 Confidential Computing Dimensions

Dimension	Definition
TEE	Trusted Execution Environment. A secure area of a main processor that guarantees code and data loaded inside it is protected with respect to confidentiality and integrity.
Remote Attestation	The cryptographic process by which a client verifies the hardware, firmware, and software state of a TEE before sending data.
Data in Use	Data actively being processed in RAM/CPU. The only state of data not protected by traditional cryptography.
FHE	Fully Homomorphic Encryption. Allows computation on ciphertext, generating an encrypted result which, when decrypted, matches the result of the same operations performed on the plaintext.
Side-Channel Attack	An attack based on information gained from the physical implementation of a system (e.g., timing, power consumption, cache states) rather than theoretical weaknesses.

2.2 The Confidential Computing Doctrine

> Data at rest is safe. Data in transit is safe. Data in use is exposed. A model that cannot protect its memory is a liability. A cloud that cannot prove its hardware is untrusted.

ENGINEERING DOCTRINE

Part III. Evaluation Methodology

3.1 Scoring Model

Each capability is scored from 0 to 5.

Score	Meaning
0	Fails basic task; plaintext processing in untrusted RAM
1	Basic container isolation but no hardware TEE
2	TEE exists but lacks remote attestation
3	Solid production performance; TEE, attestation, memory isolation
4	Strong performance; FHE integration, agent memory protection
5	Best-in-class; sets the standard for confidential computing

Weighting

Category	Weight	Rationale
Hardware Isolation & Memory Protection	25%	Untrusted RAM is the primary exfiltration vector
Remote Attestation	20%	Unverified enclaves are untrusted
AI Model & Agent Protection	20%	Model weights and agent context are high-value IP
FHE & ZKP Integration	15%	FHE is the ultimate sovereignty tool
Side-Channel Resistance	10%	Enclaves are susceptible to physical attacks
Operational Lifecycle	10%	Enclaves must be governed continuously

Total: 100%

A system MUST score at least 3.0 in Hardware Isolation and Remote Attestation to be considered for production use.

EVALUATION AND ASSURANCE SYSTEM

Capability claims are bounded by evidence, context, and reproducible assessment.

7 atomic criteria across 4 capability families define the evaluation surface of this standard.

4.1 / CAPABILITY FAMILY

Hardware Isolation and Memory Protection

SOURCE WEIGHT 25%

4.1.1
Enclave Isolation

4.1.2
Hypervisor Protection

MYTHOS-SEC-0009-EVAL-4.1.1

Enclave Isolation

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Hardware Isolation and Memory Protection	4.1.1	25%	Evidence-led

Does the system execute AI models and agent logic within a hardware-isolated TEE?

0
Execution in standard, untrusted RAM

1
Basic software containerization (Docker)

2
Virtual Machine (VM) isolation, but no hardware enclave

3
Hardware TEE (SGX, SEV-SNP, Nitro) for inference

4
TEE protection for agent memory and secret broker

5
Leading isolation: formally verified TEE boundaries, hardware-backed memory encryption, and zero host OS access

ENGINEERING INTERPRETATION

This criterion evaluates whether enclave isolation is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- approved architecture or policy record
- current configuration or platform export
- change and exception records
- monitoring, test, or assessment evidence

ASSESSMENT PROCEDURE

1. Examine authoritative design, policy, configuration, and lifecycle records.
2. Interview the accountable owner and operators responsible for the control.
3. Test a representative positive, negative, failure, and recovery scenario.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- The control exists only in diagrams or policy text.
- Observed state differs from approved intent without a governed exception.
- Evidence cannot establish scope, time, owner, or operating effectiveness.

OPERATIONAL MEASURES

- coverage of in-scope assets or flows
- age and closure rate of unresolved findings
- percentage of changes passing pre-deployment validation
- time to detect and contain control failure

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

MYTHOS-SEC-0009-EVAL-4.1.2

Hypervisor Protection

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Hardware Isolation and Memory Protection	4.1.2	25%	Evidence-led

Is the TEE protected from the hypervisor/host OS?

0
Host OS can read enclave memory

1
Basic isolation but hypervisor can still map memory

2
Memory encryption keys generated within CPU

3
Full memory encryption (e.g., AMD SEV-SNP)

4
Hypervisor cannot intercept interrupts or I/O

5
Leading protection: formally verified hypervisor isolation and cryptographic memory separation

ENGINEERING INTERPRETATION

This criterion evaluates whether hypervisor protection is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- approved architecture or policy record
- current configuration or platform export
- change and exception records
- monitoring, test, or assessment evidence

ASSESSMENT PROCEDURE

1. Examine authoritative design, policy, configuration, and lifecycle records.
2. Interview the accountable owner and operators responsible for the control.
3. Test a representative positive, negative, failure, and recovery scenario.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- The control exists only in diagrams or policy text.
- Observed state differs from approved intent without a governed exception.
- Evidence cannot establish scope, time, owner, or operating effectiveness.

OPERATIONAL MEASURES

- coverage of in-scope assets or flows
- age and closure rate of unresolved findings
- percentage of changes passing pre-deployment validation
- time to detect and contain control failure

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

4.2 / CAPABILITY FAMILY

Remote Attestation

SOURCE WEIGHT 20%

4.2.1
Attestation Verification

4.2.2
Data Binding

MYTHOS-SEC-0009-EVAL-4.2.1

Attestation Verification

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Remote Attestation	4.2.1	20%	Evidence-led

Does the system cryptographically verify the enclave before sending data?

0
No attestation

1
Basic quote generation but no verification

2
Verification of hardware quote but not application state

3
Full attestation of hardware, firmware, and application

4
Continuous attestation during execution

5
Leading attestation: cryptographic binding of data to specific enclave state, verified by adversarial suite

ENGINEERING INTERPRETATION

This criterion evaluates whether attestation verification is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- cryptographic inventory, algorithm registry, key records, certificate data, and compatibility tests
- approved architecture or policy record
- current configuration or platform export

ASSESSMENT PROCEDURE

1. test rotation, hybrid negotiation, rollback, revocation, and algorithm-policy enforcement
2. Examine authoritative design, policy, configuration, and lifecycle records.

EXAMINE – INTERVIEW – TEST

COMMON FAILURE PATTERNS

- hardcoded algorithms, unknown dependencies, missing rollback, and unverifiable keys
- The control exists only in diagrams or policy text.

OPERATIONAL MEASURES

- inventory coverage
- deprecated algorithm count
- rotation success rate
- coverage of in-scope assets or flows

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

MYTHOS-SEC-0009-EVAL-4.2.2

Data Binding

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Remote Attestation	4.2.2	20%	Evidence-led

Is data cryptographically bound to the attested enclave?

0
Data sent in plaintext after initial check

1
Basic TLS to the enclave

2
Data encrypted with enclave's public key

3
Data sealed to the specific hardware TEE

4
Key release gated by attestation state

5
Leading binding: formally verified key release, cryptographic sealing, and zero data exposure outside the TEE

ENGINEERING INTERPRETATION

This criterion evaluates whether data binding is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- cryptographic inventory, algorithm registry, key records, certificate data, and compatibility tests
- approved architecture or policy record
- current configuration or platform export

ASSESSMENT PROCEDURE

1. test rotation, hybrid negotiation, rollback, revocation, and algorithm-policy enforcement
2. Examine authoritative design, policy, configuration, and lifecycle records.

EXAMINE – INTERVIEW – TEST

COMMON FAILURE PATTERNS

- hardcoded algorithms, unknown dependencies, missing rollback, and unverifiable keys
- The control exists only in diagrams or policy text.

OPERATIONAL MEASURES

- inventory coverage
- deprecated algorithm count
- rotation success rate
- coverage of in-scope assets or flows

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

4.3 / CAPABILITY FAMILY

AI Model and Agent Protection

SOURCE WEIGHT 20%

4.3.1
Model Weight Protection

4.3.2
Agent Memory Isolation

MYTHOS-SEC-0009-EVAL-4.3.1

Model Weight Protection

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
AI Model and Agent Protection	4.3.1	20%	Evidence-led

Are proprietary model weights protected inside the TEE?

0
Weights stored in plaintext on disk and RAM

1
Encrypted on disk, decrypted in RAM

2
Decrypted only inside the TEE

3
Weights sealed to the hardware

4
Weights never exist in plaintext outside the TEE

5
Leading protection: formally verified weight sealing, zero exposure, and attestation-gated loading

ENGINEERING INTERPRETATION

This criterion evaluates whether model weight protection is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- approved architecture or policy record
- current configuration or platform export
- change and exception records
- monitoring, test, or assessment evidence

ASSESSMENT PROCEDURE

1. Examine authoritative design, policy, configuration, and lifecycle records.
2. Interview the accountable owner and operators responsible for the control.
3. Test a representative positive, negative, failure, and recovery scenario.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- The control exists only in diagrams or policy text.
- Observed state differs from approved intent without a governed exception.
- Evidence cannot establish scope, time, owner, or operating effectiveness.

OPERATIONAL MEASURES

- coverage of in-scope assets or flows
- age and closure rate of unresolved findings
- percentage of changes passing pre-deployment validation
- time to detect and contain control failure

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

MYTHOS-SEC-0009-EVAL-4.3.2

Agent Memory Isolation

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
AI Model and Agent Protection	4.3.2	20%	Evidence-led

Is the agent's context window and memory protected?

0
Agent memory in standard RAM

1
Basic process isolation

2
Container isolation

3
TEE isolation for agent context

4
Enclave-protected memory and secret broker

5
Leading isolation: formally verified memory boundaries, zero host access, and cryptographic state sealing

ENGINEERING INTERPRETATION

This criterion evaluates whether agent memory isolation is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- approved architecture or policy record
- current configuration or platform export
- change and exception records
- monitoring, test, or assessment evidence

ASSESSMENT PROCEDURE

1. Examine authoritative design, policy, configuration, and lifecycle records.
2. Interview the accountable owner and operators responsible for the control.
3. Test a representative positive, negative, failure, and recovery scenario.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- The control exists only in diagrams or policy text.
- Observed state differs from approved intent without a governed exception.
- Evidence cannot establish scope, time, owner, or operating effectiveness.

OPERATIONAL MEASURES

- coverage of in-scope assets or flows
- age and closure rate of unresolved findings
- percentage of changes passing pre-deployment validation
- time to detect and contain control failure

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

4.4 / CAPABILITY FAMILY

FHE and ZKP Integration

SOURCE WEIGHT 15%

4.4.1
Homomorphic Computation

MYTHOS-SEC-0009-EVAL-4.4.1

Homomorphic Computation

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
FHE and ZKP Integration	4.4.1	15%	Evidence-led

Can the system perform inference on encrypted data?

0
No FHE support

1
Theoretical FHE support but no implementation

2
Basic FHE for simple models

3
FHE for standard transformer models (high latency)

4
Optimized FHE with acceptable latency

5
Leading FHE: production-latency computation on encrypted data, verified by FHE evaluation suite

ENGINEERING INTERPRETATION

This criterion evaluates whether homomorphic computation is governed as an operating capability rather than a one-time configuration. Assessment must distinguish intended design, approved implementation, observed behavior, historical evidence, and unresolved risk.

REQUIRED EVIDENCE

- approved architecture or policy record
- current configuration or platform export
- change and exception records
- monitoring, test, or assessment evidence

ASSESSMENT PROCEDURE

1. Examine authoritative design, policy, configuration, and lifecycle records.
2. Interview the accountable owner and operators responsible for the control.
3. Test a representative positive, negative, failure, and recovery scenario.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- The control exists only in diagrams or policy text.
- Observed state differs from approved intent without a governed exception.
- Evidence cannot establish scope, time, owner, or operating effectiveness.

OPERATIONAL MEASURES

- coverage of in-scope assets or flows
- age and closure rate of unresolved findings
- percentage of changes passing pre-deployment validation
- time to detect and contain control failure

DECISION BOUNDARY

A maturity score is valid only for the assessed scope and evidence period. Documentation alone does not establish operating effectiveness.

LIFECYCLE AND ASSURANCE

Part V. The Mythos Confidential Computing Suite (MCCS)

Traditional security benchmarks evaluate software configuration. The MCCS evaluates hardware isolation, attestation, and FHE integration.

5.1 Suite Composition

5.1.1 MCCS-Attestation

- **Quote Verification:** 100+ tests verifying the system rejects unattested or compromised enclaves.
- **Data Binding:** 50+ tests verifying data cannot be read outside the attested state.

5.1.2 MCCS-Isolation

- **Memory Exfiltration:** 50+ tests attempting to read agent memory from the host OS.
- **Side-Channel:** 50+ tests simulating cache-timing attacks.

5.2 Execution Protocol

PUBLICATION BOUNDARY

Candidate standard. Permanent identity. Evidence before authority.

Approval requires designated technical, security, legal, accessibility, and governance review with recorded findings and dispositions.