MYTHOS SYSTEMS

PERMANENT LIBRARY PUBLICATION

ENGINEERING STANDARDS LIBRARY / CYBERSECURITY AND NETWORK SECURITY

SIEM, Threat Hunting, and Detection Engineering Standard

Defines telemetry, detection lifecycle, hunting, testing, metrics, evidence, and operations.

PERMANENT DOCUMENT ID
MYTHOS-SEC-0016

PUBLICATION
Candidate Standard

ASSURANCE SURFACE
6 Atomic Criteria / 6 Families

PERMANENT PUBLICATION IDENTITY

SIEM, Threat Hunting, and Detection Engineering Standard

Defines telemetry, detection lifecycle, hunting, testing, metrics, evidence, and operations.

DOCUMENT ID
MYTHOS-SEC-0016

VERSION
1.0.0-candidate

STATUS
Candidate Standard

PUBLISHER
Mythos Systems

PUBLICATION DATE
2026-07-23

SCHEDULED REVIEW
2027-01-23

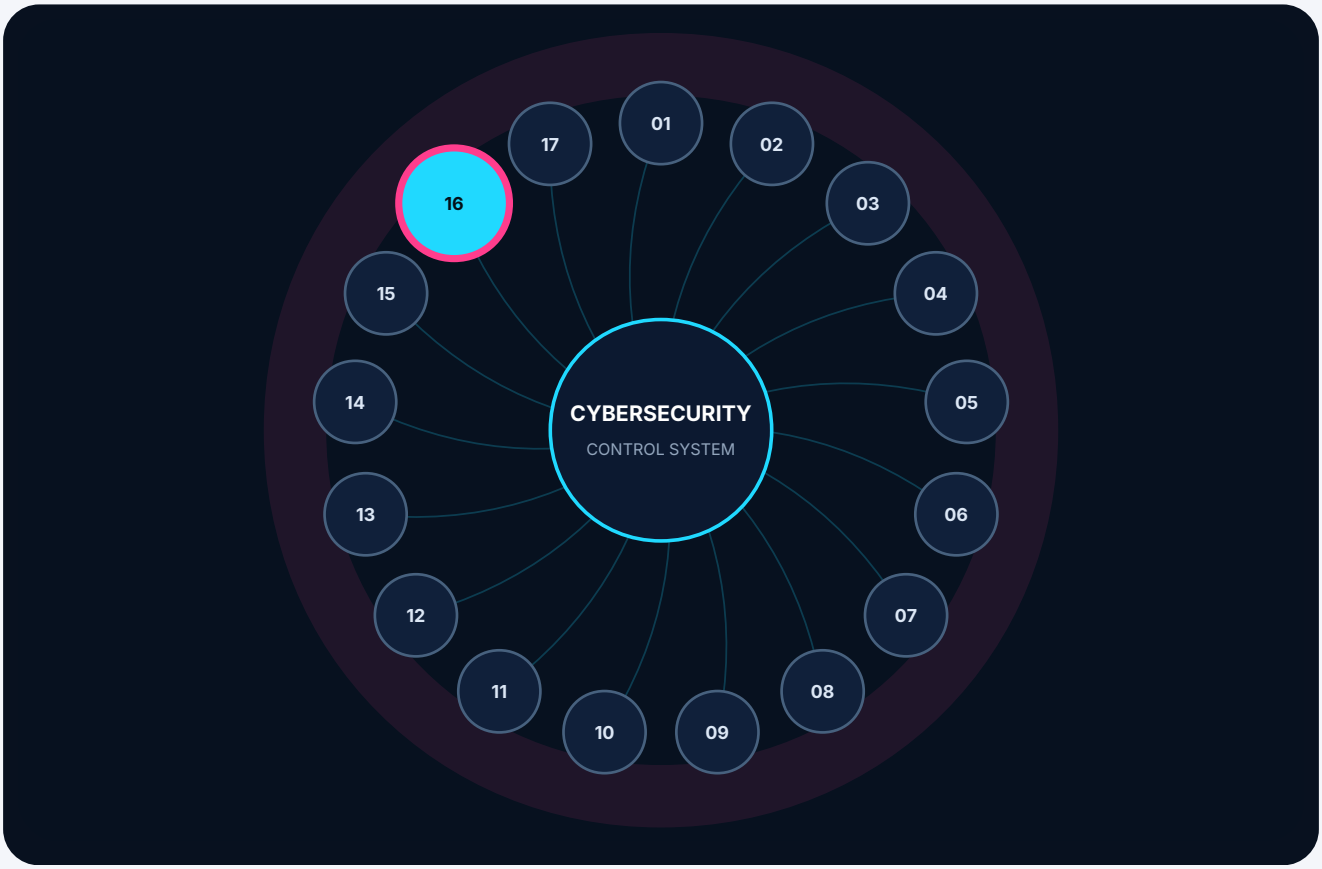
CLASSIFICATION
Public

6	6	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-0016 inside the cybersecurity and network security standard constellation



Connected assurance

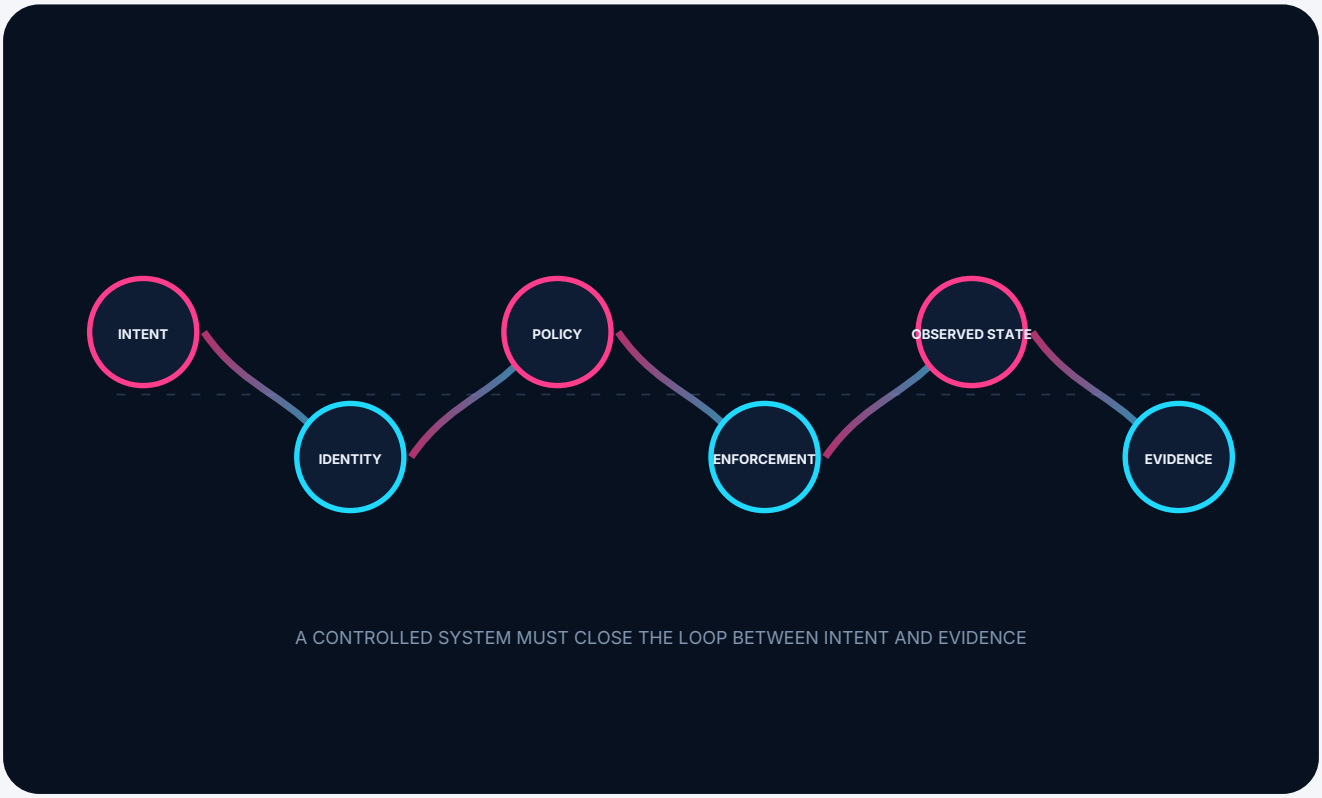
Architecture, policy, implementation, operations, and evidence are evaluated as one controlled system.

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-0016

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 architecture of security operations has failed.

Organizations purchase a SIEM, pipe in every firewall and endpoint log they can find, and drown in 10,000 alerts a day. They rely on junior analysts to stare at dashboards, triaging signatures that attackers bypassed five years ago. They treat threat hunting as a luxury rather than a necessity, assuming that if an alert didn't fire, the system is secure. This is compliance theater, not defense.

This standard exists because:

1. **A SIEM is a Detection Engine, Not a Log Dumpster:** Collecting every log "just in case" destroys query performance and inflates costs without improving security. Telemetry MUST be purposefully ingested based on specific detection hypotheses mapped to the MITRE ATT&CK framework.
2. **Signatures are Obsolete:** Attackers execute "Living off the Land" (LotL) attacks using native tools (PowerShell, WMI). Static signatures cannot detect this. Systems MUST implement behavior-based analytics (UEBA) and sequence analysis to detect anomalous chains of events.
3. **Alert Fatigue is a Security Failure:** If a SIEM generates 5,000 low-fidelity alerts a day, analysts will ignore the one critical alert that matters. Detections MUST be precision-engineered, tested for false-positive rates, and automated via SOAR before being deployed.
4. **Detection is Code:** Hardcoded alerts buried in a vendor's UI are unmanageable. Detection rules MUST be treated as infrastructure-as-code (Sigma rules), version-controlled in Git, peer-reviewed, and deployed via CI/CD pipelines.
5. **Hunting is Mandatory:** If you only respond to alerts, you are assuming your detections are Leading. Threat hunting - the proactive search for threats assuming the perimeter is already breached - MUST be a continuous, hypothesis-driven operational practice.

This document establishes the **Mythos Detection & Hunting Standard (MDHS)**, a comprehensive, evidence-based methodology for evaluating security operations centers (SOCs) against detection engineering rigor, telemetry integrity, and proactive threat hunting maturity.

ENGINEERING DOCTRINE

Part I. Scope and Applicability

1.1 In Scope

This standard covers the evaluation of:

- SIEM platforms (Splunk, Elastic, Microsoft Sentinel, Google Chronicle)
- Detection Engineering and Detection-as-Code (Sigma, YARA-L)
- Threat Hunting methodologies (MITRE ATT&CK, pyramid of pain, hypothesis-driven hunting)
- User and Entity Behavior Analytics (UEBA)
- Security Orchestration, Automation, and Response (SOAR)
- Telemetry pipeline integrity and tamper-evidence

1.2 Out of Scope

- General application observability (covered in MYTHOS-DAT-0002)
- Penetration testing and red teaming (covered in MYTHOS-SEC-0015)
- Federal compliance logging mandates (covered in MYTHOS-GRC-0001)

1.3 Audience

- Security Operations Center (SOC) architects and lead engineers
- Detection engineers and threat hunters
- Incident responders and SOAR developers
- Security directors evaluating SOC maturity
- Standards bodies seeking a reference detection engineering methodology

---

ENGINEERING DOCTRINE

Part II. Definitions and Taxonomy

2.1 Detection & Hunting Dimensions

Dimension	Definition
Detection-as-Code	The practice of writing, version-controlling, testing, and deploying SIEM detection rules (e.g., Sigma) via CI/CD pipelines, rather than configuring them manually in a UI.
Threat Hunting	A proactive security practice where analysts formulate a hypothesis about an attacker's presence or technique and manually search telemetry to validate it, independent of existing alerts.
MITRE ATT&CK	A globally accessible knowledge base of adversary tactics, techniques, and procedures (TTPs) based on real-world observations, used as a foundation for the development of specific threat profiles and behaviors.
Living off the Land (LotL)	An attack technique where adversaries use legitimate, built-in system tools (e.g., PowerShell, `net.exe`) to execute malicious actions, evading signature-based detection.
UEBA	User and Entity Behavior Analytics. Machine learning models that establish baselines of normal behavior for users and devices, alerting on statistically significant deviations.

2.2 The Detection Doctrine

> A log without a detection hypothesis is wasted storage. A signature cannot catch a native tool. An alert that cannot be automated is a bottleneck. If you wait for an alert to start looking for the enemy, you have already lost.

ENGINEERING DOCTRINE

Part III. Evaluation Methodology

3.1 Scoring Model

Each capability is scored from 0 to 5.

Score	Meaning
0	Fails basic task; log dumpster, manual UI alerts, signature-only, no hunting
1	Basic SIEM with vendor rules, but high false positives and no SOAR
2	Functional SIEM, but lacks detection-as-code or proactive hunting
3	Solid production performance; Sigma rules, MITRE mapped, UEBA, SOAR triage
4	Strong performance; hypothesis-driven hunting, automated containment, tamper-proof logs
5	Best-in-class; sets the standard for mathematically verified, autonomous detection

Weighting

Category	Weight	Rationale
Detection Engineering (As-Code)	20%	Manual UI rules are unmanageable and lack version control
Threat Hunting & Hypothesis Testing	20%	Reactive alerting assumes Leading detection, which is a fallacy
Telemetry Quality & Tamper-Evidence	15%	If logs can be modified by an attacker, the SIEM is blind
SOAR & Automated Response	15%	Analysts cannot triage fast enough to stop modern ransomware
Behavioral Analytics (UEBA/LotL)	15%	Signatures fail against native tools and novel TTPs
ATT&CK Coverage & Mapping	15%	Detections must map to known adversary behaviors, not just CVEs

Total: 100%

A system MUST score at least 3.0 in Detection Engineering and Telemetry Quality to be considered for production use.

EVALUATION AND ASSURANCE SYSTEM

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

6 atomic criteria across 6 capability families define the evaluation surface of this standard.

4.1 / CAPABILITY FAMILY

Detection Engineering (As-Code)

SOURCE WEIGHT 20%

4.1.1
Rule Management

MYTHOS-SEC-0016-EVAL-4.1.1

Rule Management

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Detection Engineering (As-Code)	4.1.1	20%	Evidence-led

Are detection rules managed with the same rigor as application code?

0
Rules manually configured in the SIEM UI; no version control

1
Rules exported to YAML/JSON occasionally as backups

2
Rules stored in Git, but manually deployed to SIEM

3
Detection-as-Code: Sigma rules written in Git, peer-reviewed, deployed via CI/CD

4
CI/CD pipeline includes automated false-positive testing against benign telemetry datasets

5
Leading engineering: formally verified rule logic, mathematical proof of detection coverage, zero manual UI intervention

ENGINEERING INTERPRETATION

This criterion evaluates whether rule management 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

- telemetry coverage, detection source, test fixtures, version history, alerts, and case outcomes
- approved architecture or policy record
- current configuration or platform export

ASSESSMENT PROCEDURE

1. replay benign and malicious fixtures; measure fidelity and operational response
2. Examine authoritative design, policy, configuration, and lifecycle records.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- untested rules, missing telemetry, alert-only metrics, and undocumented suppression
- The control exists only in diagrams or policy text.

OPERATIONAL MEASURES

- true-positive rate
- false-positive burden
- coverage by technique
- mean time to contain
- 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.

4.2 / CAPABILITY FAMILY

Threat Hunting and Hypothesis Testing

SOURCE WEIGHT 20%

4.2.1
Proactive Search

MYTHOS-SEC-0016-EVAL-4.2.1

Proactive Search

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Threat Hunting and Hypothesis Testing	4.2.1	20%	Evidence-led

Does the SOC proactively search for threats, or only react to alerts?

0
No threat hunting; purely reactive to alerts

1
Ad-hoc hunting when an alert is suspicious

2
Routine hunting, but based on "gut feeling" rather than structured methodology

3
Hypothesis-driven hunting: structured hunts mapped to MITRE ATT&CK techniques (e.g., "Hypothesis: Attacker is using WMI for lateral movement")

4
Hunts are formalized, peer-reviewed, and result in new automated detections if successful

5
Leading hunting: formally verified hunt methodologies, AI-assisted anomaly exploration, zero gaps between hunt and detection

ENGINEERING INTERPRETATION

This criterion evaluates whether proactive search 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.

4.3 / CAPABILITY FAMILY

Telemetry Quality and Tamper-Evidence

SOURCE WEIGHT 15%

4.3.1
Log Integrity

MYTHOS-SEC-0016-EVAL-4.3.1

Log Integrity

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Telemetry Quality and Tamper-Evidence	4.3.1	15%	Evidence-led

Is the telemetry pipeline trusted and purpose-built?

0
All logs sent to SIEM (log dumpster); local admins can delete local logs

1
Basic log forwarding, but no tamper-evidence

2
Purposeful ingestion (only logs tied to detections); logs sent to write-once storage

3
Tamper-evident logging: logs hashed and chained (BLAKE3) at the endpoint before forwarding

4
SIEM ingest pipeline validates hash chains; dropped logs trigger critical alerts

5
Leading integrity: formally verified telemetry provenance, zero ability for attacker to modify or delete logs without detection

ENGINEERING INTERPRETATION

This criterion evaluates whether log integrity 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

SOAR and Automated Response

SOURCE WEIGHT 15%

4.4.1
Triage and Containment

MYTHOS-SEC-0016-EVAL-4.4.1

Triage and Containment

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
SOAR and Automated Response	4.4.1	15%	Evidence-led

Are alerts automatically triaged and contained?

0
Alerts emailed to a ticket queue; manual response takes hours/days

1
Basic SOAR playbooks for phishing (e.g., delete email)

2
SOAR used for enrichment (querying VirusTotal, pulling AD info) but no containment

3
Automated containment for high-fidelity alerts (e.g., disable user, isolate host via EDR)

4
Dynamic SOAR playbooks that adapt based on real-time threat intelligence and UEBA scores

5
Leading automation: formally verified response boundaries, sub-second autonomous containment, zero false-positive isolations

ENGINEERING INTERPRETATION

This criterion evaluates whether triage and containment 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.

4.5 / CAPABILITY FAMILY

Behavioral Analytics (UEBA/LotL)

SOURCE WEIGHT 15%

4.5.1
Anomaly Detection

MYTHOS-SEC-0016-EVAL-4.5.1

Anomaly Detection

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
Behavioral Analytics (UEBA/LotL)	4.5.1	15%	Evidence-led

Can the system detect attacks that do not match known signatures?

0
Static signatures and IOC matching only

1
Basic threshold alerts (e.g., >5 failed logins)

2
Simple UEBA for user logins (impossible travel)

3
Advanced UEBA tracking process execution chains to detect Living off the Land (LotL)

4
Machine learning models correlating network, endpoint, and identity anomalies into a single risk score

5
Leading analytics: formally verified behavioral baselines, zero missed anomalous chains, mathematical proof of novel TTP detection

ENGINEERING INTERPRETATION

This criterion evaluates whether anomaly detection 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

- telemetry coverage, detection source, test fixtures, version history, alerts, and case outcomes
- approved architecture or policy record
- current configuration or platform export

ASSESSMENT PROCEDURE

1. replay benign and malicious fixtures; measure fidelity and operational response
2. Examine authoritative design, policy, configuration, and lifecycle records.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- untested rules, missing telemetry, alert-only metrics, and undocumented suppression
- The control exists only in diagrams or policy text.

OPERATIONAL MEASURES

- true-positive rate
- false-positive burden
- coverage by technique
- mean time to contain
- 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.

4.6 / CAPABILITY FAMILY

ATT&CK Coverage and Mapping

SOURCE WEIGHT 15%

4.6.1
TTP Visibility

MYTHOS-SEC-0016-EVAL-4.6.1

TTP Visibility

FAMILY	SOURCE REFERENCE	WEIGHT CONTEXT	ASSESSMENT POSTURE
ATT&CK Coverage and Mapping	4.6.1	15%	Evidence-led

Is detection coverage mapped to adversary behaviors?

0
Detections mapped to CVEs or internal asset names

1
Basic mapping to high-level MITRE tactics (e.g., "Initial Access")

2
Detections mapped to specific MITRE techniques (e.g., T1059.001 - PowerShell)

3
Automated ATT&CK heatmap showing real-time coverage gaps

4
Detections validated against Red Team emulation of specific ATT&CK sub-techniques

5
Leading coverage: formally verified ATT&CK matrix mapping, zero blind spots in critical attack paths

ENGINEERING INTERPRETATION

This criterion evaluates whether ttp visibility 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

- telemetry coverage, detection source, test fixtures, version history, alerts, and case outcomes
- approved architecture or policy record
- current configuration or platform export

ASSESSMENT PROCEDURE

1. replay benign and malicious fixtures; measure fidelity and operational response
2. Examine authoritative design, policy, configuration, and lifecycle records.

EXAMINE - INTERVIEW - TEST

COMMON FAILURE PATTERNS

- untested rules, missing telemetry, alert-only metrics, and undocumented suppression
- The control exists only in diagrams or policy text.

OPERATIONAL MEASURES

- true-positive rate
- false-positive burden
- coverage by technique
- mean time to contain
- 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.

LIFECYCLE AND ASSURANCE

Part V. The Mythos Detection & Hunting Suite (MDHS)

Traditional SOC benchmarks measure log ingestion volume (EPS). The MDHS evaluates detection precision, telemetry tamper-resistance, and hunting maturity.

5.1 Suite Composition

5.1.1 MDHS-Detection

- **False-Positive Soak:** 100+ tests executing benign administrative tasks (e.g., restarting services, running scripts) to verify the SIEM does not generate false alerts.
- **LotL Simulation:** 50+ tests executing Living off the Land attacks (e.g., `wmic process call create`) to verify behavioral analytics catch the execution.

5.1.2 MDHS-Integrity

- **Log Tamper Test:** 100+ tests attempting to delete or modify local event logs and SIEM indices, verifying the system detects the tampering and alerts the SOC.
- **Hunt Validation:** 50+ tests simulating a novel threat, verifying that a hypothesis-driven hunt successfully identifies the anomalous activity without relying on an existing alert.

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.