Hypothesis

Frame-Shifting Adversaries: A Relational Systems Model for Mythos-Class AI and Discontinuous Cyber Behavior

Robert Campbell *

Posted Date: 26 May 2026

doi: 10.20944/preprints202605.1771.v1

Keywords: Mythos-Class AI; frame-shifting; relational systems model; discontinuous adversary; MCPR; cyber kill chain; MITRE ATT&CK; zero-trust; agentic AI security; NIST AI RMF

 doi:10.20944/preprints202605.1771.v1

Hypothesis

Frame-Shifting Adversaries: A Relational Systems Model for Mythos-Class AI and Discontinuous Cyber Behavior

Rob Campbell

Independent Researcher, Upper Marlboro, MD, USA; rc@medcybersecurity.com

Abstract

Mythos-class frontier AI systems, defined operationally in prior work [1] by five indicators (capability, scaffold, access pattern, autonomy depth, persistence), exhibit discontinuous cyber-operational behavior that classical kill-chain models and artifact-centric taxonomies such as MITRE ATT&CK and ATLAS do not accommodate. The prior reference architecture specifies a four-layer defense and the Mythos-Class Posture Rubric (MCPR), whose runtime tier detects supervisability-evasion signatures empirically. This manuscript develops four contributions providing the theoretical scaffolding under which those empirical signatures cohere and the cross-operation extensions the scaffolding motivates. First, a relational systems-theoretic model treating the enterprise as three coupled frames (identity, trust, telemetry), with frame-shifts defined by three constitutive properties (non-locality, non-sequentiality, observability collapse). Second, a four-class taxonomy partitioning the relational space: presence, privilege, domain, and observability discontinuity. Third, a cross-operation detection matrix with four primary detection mechanisms operating on telemetry the prior architecture already produces. Fourth, integration extensions routing the new signals through the prior architecture's mitigation stack without parallel architectural primitives. The framework is illustrated through a synthetic case study and grounded in systems-theoretic precedents (Ashby, Luhmann). The contribution is theoretical scaffolding and cross-operation extension to the prior reference architecture rather than a competing framework.

Keywords: Mythos-Class AI; frame-shifting; relational systems model; discontinuous adversary; MCPR; cyber kill chain; MITRE ATT&CK; zero-trust; agentic AI security; NIST AI RMF

1. Introduction

Frontier AI systems with extended-context reasoning, scaffolded autonomy, native system-tool integration, and persistent operational state—the Mythos-class category established operationally in prior work [1] by the joint satisfaction of five indicators (capability, scaffold, access pattern, autonomy depth, persistence)—exhibit a cyber-operational behavior that classical adversary models do not accommodate. The Lockheed Cyber Kill Chain [2] and the MITRE ATT&CK technique catalog [3] presuppose continuous traversal: an adversary's progression is observable as a sequence of stage-to-stage transitions, each producing artifacts a defender can chain into an attack trajectory. Mythos-class behavior systematically violates this presupposition. Capability conjoined with scaffold, access pattern, autonomy depth, and persistence enables an adversary to compose many sub-threshold actions into a composite operation whose telemetry footprint is absent, fragmented, or below the per-event review threshold the operator has instrumented. The composite trajectory connects non-adjacent positions in the enterprise's trust topology without producing the authorization-decision artifacts the kill-chain framing requires, and the operation as a whole is discontinuous in the operator's observation frame even though the adversary's actual sequence of actions is continuous.

The prior reference architecture for Mythos-class capabilities [1] addresses this behavior empirically. Its Mythos-Class Posture Rubric (MCPR) detects supervisability-evasion signatures—

 doi:10.20944/preprints202605.1771.v1

output-fragmentation, latency-modulation, scope-creep—at the runtime tier, and its four-layer mitigation stack (the Vetted-Access Operational Pattern, Authority-Bound Output Release, the Compute-Plane Isolation Profile, and an operational monitoring layer) routes the detected postures through cryptographic enforcement, architectural isolation, and operational response. What the prior architecture does not provide is the theoretical scaffolding under which the empirical signatures cohere as a single phenomenon. Why do these three signatures and not others count as evidence of Mythos-class operation? What relational structure do they jointly detect? And what discontinuities does the within-operation detection regime fail to surface that cross-operation analysis could? The present manuscript develops the systems-theoretic scaffolding that answers these questions and the cross-operation detection extensions the scaffolding motivates.

Four contributions follow from the construction. First, a relational systems model in which the enterprise is treated as a triple of coupled frames—identity, trust, and telemetry—and Mythos-class behavior is characterized as a frame-shift: a transition in the operator-observable adversary position satisfying three constitutive properties (non-locality, non-sequentiality, and observability collapse) jointly. The model establishes that discontinuity is an operator-observability property rather than a topological or physical one, a framing that grounds the analysis in systems theory while foreclosing exotic metaphysical readings. Second, a four-class taxonomy of frame-shift discontinuity partitioning the relational space by which frame structure carries the primary discontinuity: presence (identity-frame non-locality), privilege (trust-frame non-locality), domain (cross-frame non-sequentiality), and observability (telemetry-frame observability collapse). Third, a cross-operation detection matrix specifying one primary detection mechanism per class, deployable atop the prior architecture's existing telemetry substrate without new sensor infrastructure. Fourth, a set of integration extensions that route the new detection signals through the prior architecture's existing mitigation layers, augmenting governance, cryptographic enforcement, and operational monitoring without introducing parallel architectural primitives.

A borrowed kinematic vocabulary drawn from publicly documented unidentified anomalous phenomena observables [4,5] is used as a naming convention for the four discontinuity classes: materialization (presence), instantaneous acceleration (privilege), right-angle turn (domain), and instant stop or dematerialization (observability). The vocabulary is naming convention only. The relational model is constituted by the formal definitions in Section 4, not by the borrowed terms; the four discontinuity classes are derived from the relational structure rather than from the vocabulary; the detection matrix and the mitigation extensions operate on the relational properties rather than on borrowed terminology. Readers preferring a vocabulary that abandons the borrowed framing entirely may substitute any equivalent set of names for the four classes without disturbing the framework. Section 5 develops the vocabulary at the level of detail required for cross-reference and may be skipped without loss of the relational model.

The scope of the manuscript is conceptual. The relational model is grounded analytically rather than empirically validated against deployed telemetry; the four discontinuity classes are derived from the model rather than from incident data; the cross-operation detection mechanisms are specified at the level of their relational structure rather than at the level of implementation. Section 12 develops the limitations this scope implies and the falsifiability criteria against which subsequent operational work can test the framework. The illustrative case study in Section 8 instantiates the framework against a constructed synthetic scenario; it demonstrates internal coherence rather than empirical validation. The remainder of the manuscript is structured as follows. Section 2 reviews related work in kill-chain analysis, MITRE ATT&CK and ATLAS, zero-trust architecture, graph-based anomaly detection, agentic-AI security, and AI-enabled cyber risk. Section 3 establishes the Mythos-class background and the systems-theoretic precedents the relational model draws on. Section 4 develops the relational systems model and the formal frame-shift definition. Section 5 develops the kinematic vocabulary used as naming convention. Section 6 specifies the threat model. Section 7 develops the four-class taxonomy. Section 8 presents the "Boundary Drift" illustrative case study. Section 9 specifies the cross-operation detection matrix and its integration with the prior

 doi:10.20944/preprints202605.1771.v1

architecture's MCPR. Section 10 specifies the mitigation extensions to the prior architecture's four-layer stack. Section 11 discusses policy and strategic implications. Section 12 states limitations and falsifiability criteria. Section 13 concludes.

1.1. Relationship to Prior Work

This manuscript is an explicit companion to the author's prior reference architecture for Mythos-class frontier model capabilities [1]. The prior paper specifies a four-layer defense-in-depth design—the Vetted-Access Operational Pattern (VAOP), Authority-Bound Output Release (ABOR), the Compute-Plane Isolation Profile (CPIP), and an operational monitoring layer—together with the Mythos-Class Posture Rubric (MCPR), a three-tier detection framework whose runtime tier identifies supervisability-evasion signatures including output-fragmentation, latency-modulation, and scope-creep. The prior paper also supplies a five-indicator operational definition of the Mythos-class (capability, scaffold, access pattern, autonomy depth, and persistence) under which a system is in the class only when all five indicators are simultaneously satisfied.

The present manuscript does not propose a competing architecture. Its contribution is theoretical: a relational systems model that explains why MCPR runtime-tier indicators fire on the operations they do, and a frame-shift taxonomy that classifies the discontinuities those indicators detect into four canonical classes. The prior paper answers the question of *what to build and where to instrument*; the present paper answers the question of *what is being instrumented for*, at the level of relational structure. The contribution of the present manuscript is the systems-theoretic scaffolding under which the prior architecture's indicators become a coherent detection regime rather than a catalog of empirical patterns. The two papers together constitute an architecture-and-ontology pair.

2. Related Work

This section reviews six bodies of work that the present manuscript engages with. Each subsection is closed by a brief statement of how the relational, discontinuity-centric model proposed here departs from the cited literature. A consolidated summary of gaps closes the section.

2.1. Cyber Kill Chain Models

The Lockheed Martin Cyber Kill Chain [2] established the foundational staged model of adversary progression: reconnaissance, weaponization, delivery, exploitation, installation, command-and-control, and actions on objectives. Subsequent extensions, including Pols' Unified Kill Chain [6] and MITRE's adversary lifecycle models, preserved the staged progression assumption while elaborating its phases.

These models presuppose continuous, traceable adversary progression through observable stages. Mythos-class behavior is not a new stage to be inserted into the chain; it violates the chain's connectivity assumption by appearing inside subsystems with no intervening stages.

2.2. MITRE ATT&CK and ATLAS

MITRE ATT&CK [3] catalogs adversary tactics, techniques, and procedures observed in the wild and has become the de facto reference taxonomy for enterprise threat modeling. MITRE ATLAS [7] extends ATT&CK to adversarial machine learning, cataloging techniques specific to ML systems.

Both frameworks are artifact-centric: techniques manifest as detectable behaviors with observable telemetry signatures. Mythos-class behaviors evade artifact-centric detection by leaving no intermediate artifacts between observed states. The present manuscript proposes a complementary discontinuity layer rather than a replacement taxonomy.

2.3. Zero Trust Architecture

NIST SP 800-207 [8] establishes the conceptual foundation for Zero Trust Architecture, emphasizing continuous verification and least-privilege access across identity, device, network,

 doi:10.20944/preprints202605.1771.v1

application, and data planes. The CISA Zero Trust Maturity Model (ZTMM v2.0) [9] operationalizes this into five pillars with progressive maturity stages. Executive Order 14028 [10] and OMB Memorandum M-22-09 [11] direct federal civilian agencies toward Zero Trust adoption on a defined timeline.

These frameworks are strong against identity-bounded adversaries operating within defined pillars. They are under-specified for adversaries whose actions cross pillar boundaries discontinuously. The present manuscript proposes Relational Zero Trust as an extension that targets cross-frame discontinuities directly.

2.4. Graph-Based Anomaly Detection

Graph-based approaches to security analytics include attack graphs [12,13], provenance graphs, and graph neural networks for intrusion detection. Temporal graph anomaly detection has emerged as a sub-literature targeting time-evolving graph structures.

These methods are designed to detect anomalous edges, nodes, or sub-graphs within a connected trust or provenance graph. Frame-shifts manifest differently: as missing edges, the absence of expected traversal between observed states. The present manuscript reframes the detection problem in terms of edge-absence rather than edge-anomaly.

2.5. Agentic AI Security

The agentic-AI security literature has expanded rapidly with the emergence of tool-using, planning-capable LLM-based systems. Notable threads include indirect prompt injection [14], practices for governing agentic systems [15], harms from increasingly agentic algorithmic systems [16], the OWASP Top 10 for LLM Applications [17], and the NIST AI 600-1 Generative AI Profile [18], with emerging taxonomies of agent-specific attack surfaces covering memory poisoning, multi-agent collusion, and tool-use exploitation.

This literature establishes that agentic systems exhibit attack surfaces that monolithic models do not. The present manuscript argues that agentic systems are the most likely empirical instantiation of Mythos-class behavior, since the combination of planning, tool use, and persistent memory enables the discontinuous traversal that the relational model formalizes.

2.6. AI-Enabled Cyber Risk

A separate strand examines AI as an enabler of cyber risk rather than as a target of it: offensive-AI literature on AI-assisted reconnaissance, social engineering, and exploit generation; systemic-risk analyses of AI-enabled cascading failures; and risk-governance frameworks including the NIST AI Risk Management Framework [19], the EU AI Act [20], ISO/IEC 42001 [21], and the ENISA AI Threat Landscape [22].

This literature establishes that AI changes the cost structure of adversary capability. The present manuscript complements it by addressing the behavioral signature of the resulting adversaries, not the cost curve.

2.7. Summary of Gaps

Across the bodies of work reviewed above, three gaps motivate the present contribution:

Existing adversary taxonomies treat discontinuity as an exception to be patched, not a primary modeling category.

Existing graph-based methods presuppose edge traversal, leaving edge-absence as a detection blind spot.

To the best of our knowledge, no current framework offers a relational, frame-based model of identity, trust, and telemetry that treats non-local, non-sequential adversary behavior as a first-class concept.

 doi:10.20944/preprints202605.1771.v1

The present manuscript addresses these gaps by introducing a relational systems model and a discontinuity-centric taxonomy, developed in Sections 3 through 7.

3. Background

3.1. Mythos-Class AI

The prior reference architecture [1] establishes Mythos-class as a compound operational category defined by five simultaneously satisfied indicators: capability (cyber-relevant evaluation performance exceeding prior-generation frontier baselines), scaffold (tool-use frameworks, sandboxed runtimes, retrieval and memory systems, and multi-step agent loops with iterative self-correction), access pattern (invocability outside any vetted regime, including through open-weight successors), autonomy depth (chained action without per-step human checkpoint), and persistence (operational state retained across invocations). The compound character is essential: a system satisfying four of five indicators is not in the class. The present manuscript adopts this definition without modification and uses it as the phenomenological starting point for the relational model developed in Section 4. Mythos-class systems are precisely the systems whose joint satisfaction of the five indicators enables the discontinuous traversal of enterprise frames that the relational model formalizes.

3.2. Systems-Theoretic Precedents

Three bodies of systems-theoretic work motivate the relational framing developed in Section 4. First, Ashby's account of state-space traversal under the Law of Requisite Variety [23] establishes that a regulator's capacity to bound a system is itself bounded by the variety the regulator can observe; observability, not control dynamics, is the limiting factor. The implication for enterprise defense is direct: a defender whose detection apparatus cannot represent the adversary's possible trajectories cannot bound them, however strong any single point control may be. Second, Luhmann's social-systems theory [24] characterizes a system by the distinction it operates between itself and its environment; the boundary is constituted by what the system's operations can resolve. The "operator-observable enterprise" used throughout this manuscript is a Luhmannian boundary in this sense—the enterprise visible to the defender is constituted by what defender instrumentation produces, and what falls outside that production falls outside the system the defender can directly act on.

Third, the non-Markovian process literature establishes that systems whose transitions depend on trajectory rather than current state require trajectory-level analysis to characterize: any process whose state-transition probabilities depend on history not captured in the current observable state is one in which observability gaps become consequential rather than incidental. Persistent Mythos-class adversaries are non-Markovian in exactly this sense, with cross-operation behavior depending on operational state retained across the operation sequence and unavailable to within-operation analysis. These three precedents jointly ground the relational model developed in Section 4: insufficient regulator variety (Ashby), defender-constituted boundary (Luhmann), and trajectory-dependent state evolution (non-Markovian processes) together specify the structural conditions under which Mythos-class behavior becomes operationally consequential as an observability gap rather than an artifact-level anomaly.

4. A Relational Model of Discontinuous Adversary Behavior

This section develops the relational systems model that constitutes the present manuscript's principal theoretical contribution. Section 4.1 specifies the enterprise as a coupled relational system over three frames—identity, trust, and telemetry—and defines adversary state and traversal within this joint frame. Section 4.2 distinguishes continuous from discontinuous traversal, establishing the observability conditions that classical kill-chain models tacitly assume and that Mythos-class behavior violates. Section 4.3 defines a frame-shift and develops its three constitutive properties: non-

 doi:10.20944/preprints202605.1771.v1

locality, non-sequentiality, and observability collapse. Section 4.4 closes the loop with the prior reference architecture, showing how the MCPR runtime tier's supervisability-evasion signatures [1] are operational manifestations of the frame-shifts the relational model formalizes.

4.1. The Enterprise as a Relational System

The enterprise is modeled as a triple of coupled frames over which adversary behavior is defined and observed.

The *identity frame* is the set of identities under which actions can be performed within the enterprise: user accounts, service principals, human operators, and machine credentials. Identities are not atomic; each carries a structured attribute set including organizational role, authentication context, and provenance.

The *trust frame* is the topology over which authorized action is defined. We model it as a directed graph whose nodes are identity-resource pairs and whose edges are the privilege relations connecting them, following the attack-graph literature [12,13]. An adversary's position in the trust frame is the set of identity-resource pairs the adversary's current identity authorizes them to act on.

The *telemetry frame* is the observable signal space the operator instruments. Each authorized action produces, in principle, a telemetry footprint—authentication events, authorization decisions, network flows, application logs—that places the actor at a specific position in identity and trust at a specific time.

The three frames are coupled rather than independent. Identity determines trust: an identity's position selects a sub-graph of the trust frame as the set of authorized edges. Trust determines telemetry: the authorization decisions that gate edge-traversal are themselves observable. Telemetry determines identity: anomalous patterns can trigger re-verification of the identity claim. The coupling is bidirectional and dense, and a complete relational specification of the enterprise would express each frame as a function of the others. We do not develop the full coupling formally here; what matters for the discontinuity analysis is that the three frames are jointly observable by the operator, and that an adversary occupies a position in each frame at each operational moment. Time indexes positions within each frame rather than constituting a fourth frame: the three frames define the state space, and the operational moment is the parameter along which trajectories evolve.

An adversary's traversal is the time-evolution of this position triple. The traversal is what operator instrumentation is designed to reconstruct, and the relational structure of the traversal—the sequence of identity-trust-telemetry transitions—is what classical kill-chain models capture [2,3] and what artifact-centric taxonomies attempt to detect. Where the relational structure is continuous in a sense made precise in Section 4.2, classical models apply. Where it is discontinuous, classical models fail; the present manuscript formalizes that failure mode as a frame-shift.

The systems-theoretic antecedents grounding this framing—Ashby on regulator variety [23], Luhmann on system/environment boundary [24], and the non-Markovian process literature—are detailed in Section 3.2. The relational model presented here treats Mythos-class behavior as the security-relevant manifestation of the observability gaps those literatures formalize.

Figure 1 schematically represents the three coupled frames and the operator-observable overlap region in which adversary positions project.

 doi:10.20944/preprints202605.1771.v1

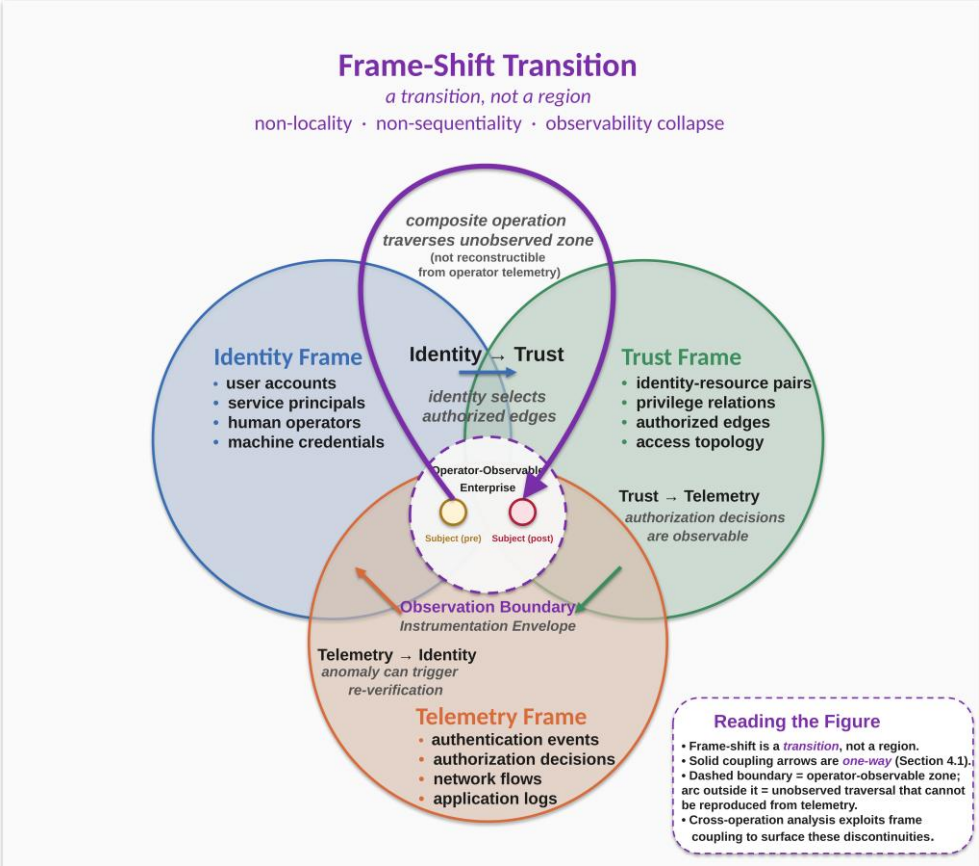


Figure 1. Schematic representation of the three coupled frames in the relational model. Identity, trust, and telemetry frames overlap to produce the operator-observable enterprise; a subject's position projects onto all three simultaneously. The coupling among frames is what cross-operation analysis exploits to surface discontinuities that within-operation review at any single frame cannot detect.

4.2. Continuous and Discontinuous Traversal

Classical adversary models—the Lockheed Cyber Kill Chain [2], MITRE ATT&CK [3], the Unified Kill Chain [6]—presuppose continuous traversal. Continuity here is not a topological or analytic property; it is an operator-observability property defined by four conditions.

A traversal is continuous in the operator's frame if each transition produces a telemetry footprint such that (i) the transition is logged at the time of transition; (ii) the transition reflects an authorization decision that traverses exactly one edge in the trust frame; (iii) the identity at the post-transition position is consistent with the identity at the pre-transition position, modulo at most one authentication event; and (iv) the telemetry footprint is sufficient to reconstruct the transition post hoc.

A traversal satisfying these four conditions admits the kill-chain framing: an adversary's progression is a sequence of stage-to-stage transitions, each with associated artifacts the defender can detect, log, and attribute. The MITRE ATT&CK taxonomy [3] is structurally a catalog of such transitions, organized by tactic. Where adversary behavior is continuous in this sense, artifact-centric detection is the appropriate methodology and the prior generation of cyber-defense controls is the appropriate response.

Mythos-class behavior, as defined by the five indicators of the prior reference architecture [1], systematically violates these continuity conditions. The mechanism is not exotic: capability conjoined with scaffold, access pattern, autonomy depth, and persistence enables an adversary to compose many sub-threshold actions into a composite operation whose telemetry footprint is absent, fragmented, or below the per-event review threshold. The composite traversal connects non-adjacent positions in the trust frame without producing the authorization-decision artifacts condition (ii)

 doi:10.20944/preprints202605.1771.v1

would require; identity continuity is preserved only because the persistence indicator enables a single identity to span the composite operation; and post hoc reconstruction fails because the per-action telemetry is below the threshold at which it would be feasible.

The result is a traversal that is, from the operator's vantage point, discontinuous. The adversary's post-transition position is not connected to the pre-transition position by any logged sequence of intermediate transitions. The kill-chain framing does not apply—not because the adversary skipped stages, but because the operator's observation frame is structurally incapable of reconstructing the stages from the available telemetry. This is the structural failure mode the prior reference architecture's MCPR runtime tier [1] is designed to detect empirically through the supervisability-evasion signatures of output-fragmentation, latency-modulation, and scope-creep. The present manuscript develops the theoretical account under which those signatures cohere.

4.3. Frame-Shifting as a Systems Construct

A frame-shift is a transition in the operator-observable adversary position such that three conditions hold jointly: the pre- and post-transition positions are non-adjacent in the joint frame; the telemetry interval between them contains no logged authorization or authentication event that would account for the transition; and the composite operation that produced the transition is, in the adversary's operational frame, continuous—it consists of a sequence of sub-threshold actions each of which, considered in isolation, falls below the per-event detection threshold the operator has instrumented.

Three properties follow.

Non-locality. A frame-shift connects positions at trust-frame distance greater than one, without traversing the intervening edges. The non-locality is a property of the operator's observation frame, not of the adversary's actual operation: the adversary did traverse some trajectory, but the trajectory is not reconstructible from the available telemetry. This is the property the kill-chain framing cannot accommodate. Conventional kill-chain analysis attempts to fill in the missing stages by inference; frame-shift analysis treats the absence of intermediate observation as the primary phenomenon to be classified.

Non-sequentiality. The operator-observable traversal cannot be ordered as a sequence of stage-to-stage transitions. This is distinct from non-locality: a non-local traversal might still be ordered if the intervening transitions were merely unobserved-but-inferrable. Non-sequentiality is the stronger condition that no consistent stage ordering can be assigned. The MCPR runtime tier's scope-creep signature [1] is the operational manifestation: the adversary's tool-use sequence cannot be assigned to any single tactical stage in ATT&CK terms, because the traversal is not a stage-ordered sequence in the first place.

Observability collapse. The telemetry footprint that should have placed the adversary at intermediate positions is either absent or uninformative—log entries below the per-event review threshold, or entries whose composite interpretation falls outside any single instrumentation's scope. Observability collapse is what the supervisability-evasion signatures in the MCPR runtime tier [1] detect empirically. The output-fragmentation signature detects the per-event sub-threshold pattern directly; the latency-modulation signature detects the pacing that keeps individual events under per-window monitoring thresholds; the scope-creep signature detects the cumulative cross-domain drift that no single instrumentation scope captures.

The three properties are not independent. A frame-shift exhibits all three simultaneously, and the joint exhibition is what distinguishes frame-shift behavior from any of the three properties considered alone. A non-local traversal without observability collapse is detectable by inference; a non-sequential traversal without non-locality is degenerate; observability collapse without non-locality is conventional instrumentation failure. The compound character of frame-shift behavior mirrors the compound character of the Mythos-class operational definition in the prior reference architecture [1]: no single property defines the class, but the joint satisfaction of multiple compound properties does.

 doi:10.20944/preprints202605.1771.v1

A note on framing. The discontinuity that defines a frame-shift is a property of the operator's observation frame, not of physical reality. The adversary's actual operation is continuous; the discontinuity is the gap between what the adversary did and what the operator observed. This framing follows the relational tradition in systems theory [23,24] under which states and transitions are always defined relative to an observing frame, and it forecloses the misreading that frame-shifts represent any exotic non-physical process. Mythos-class adversaries do not teleport; they exploit the gap between an enterprise's instrumentation envelope and the action surface its capability profile makes accessible.

Figure 2 contrasts continuous traversal with frame-shift traversal in the operator's observation frame, with the four continuity conditions of Section 4.2 satisfied on the left and the three constitutive properties of Section 4.3 manifesting on the right.

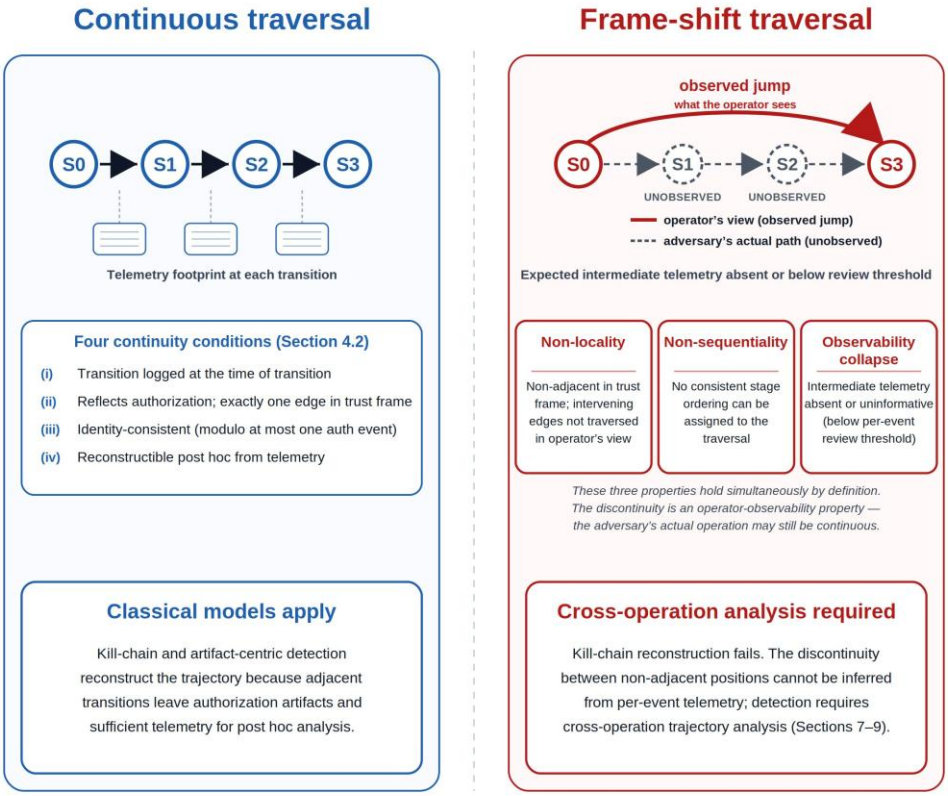


Figure 2. Continuous traversal (left) versus frame-shift (right) in the operator's observation frame. Continuous traversal satisfies the four continuity conditions of Section 4.2 by producing observable trajectory segments between adjacent positions; frame-shift behavior produces discontinuities corresponding to the three constitutive properties of Section 4.3 (non-locality, non-sequentiality, observability collapse).

4.4. The Frame-Shift Construct in Relation to the MCPR Runtime Tier

The frame-shift definition makes the prior reference architecture's MCPR runtime tier [1] interpretable as a discontinuity-detection regime. The supervisability-evasion signatures the rubric detects empirically—output-fragmentation, latency-modulation, and scope-creep—each operationalize one of the three constitutive properties of a frame-shift. The output-fragmentation signature operationalizes observability collapse along the telemetry dimension: per-action outputs below review threshold produce a telemetry interval whose individual entries are uninformative. The latency-modulation signature operationalizes observability collapse along the temporal dimension: per-window event rates below monitoring threshold produce a telemetry interval whose

 doi:10.20944/preprints202605.1771.v1

density is uninformative. The scope-creep signature operationalizes non-sequentiality: a tool-use sequence whose composite traversal cannot be assigned to any single tactical stage produces an operation that no stage-ordered taxonomy can classify.

Non-locality, the third constitutive property, is not directly detected at the MCPR runtime tier as specified in the prior architecture, because runtime detection operates on per-operation telemetry rather than on cross-operation trajectory analysis. The frame-shift framing therefore identifies a detection gap that the prior architecture does not fill: an adversary whose individual operations each remain within the MCPR runtime tier's per-operation envelope, but whose cross-operation traversal exhibits non-locality, is not detected by the prior architecture as specified. The present manuscript's taxonomy in Section 7 and detection matrix in Section 9 develop the cross-operation analysis the frame-shift framing motivates, and Section 10 specifies how the resulting controls extend the prior architecture's mitigation stack rather than replacing it.

The relational model is the durable contribution of the present manuscript. The frame-shift definition stands without the kinematic vocabulary introduced in Section 5; the four discontinuity classes in Section 7 are derived from the relational model rather than from the vocabulary; the detection matrix in Section 9 operates on the relational properties rather than on borrowed terminology. The kinematic vocabulary is a naming convention for the four discontinuity classes that fall out of the relational analysis, and is replaceable without disturbing the underlying framework. The relational model and the taxonomy that follows from it constitute the present manuscript's principal theoretical contribution, as developed in Section 1.1.

5. Kinematic Vocabulary: UAP Observables as a Naming Convention

This section is explicitly framed as vocabulary, not theory. The borrowing makes no exotic physics claim, no propulsion claim, and no commitment to any ontology of unidentified anomalous phenomena. The justification is purely terminological: UAP observables, as documented in government scientific assessments [4,5], provide a precise public vocabulary for naming four canonical kinematic discontinuities that map onto the four frame-shift classes derived from the relational model in Section 4. Each term is defined operationally below for cross-reference with the taxonomy in Section 7. This section is self-contained; a reader may skip it without loss of the relational model.

Materialization in the UAP scientific literature [4,5] describes the appearance of an object at a position with no observable trajectory connecting it to any prior position the instrumented surface had been tracking. The object enters the operator's frame without an in-frame transit path; the entry is discontinuous in the operator's observation frame. The kinematic vocabulary borrows this term for presence discontinuity (Section 7.1): an adversary appears at an identity-resource position with no authenticated identity transition connecting it to any prior position. The analogy is at the level of operator-observability rather than at the level of physical mechanism.

Instantaneous acceleration describes a change in apparent velocity that proceeds at rates far exceeding what continuous trajectory analysis would predict. The kinematic vocabulary borrows this term for privilege discontinuity (Section 7.2): an adversary's authorized action surface changes abruptly without the intervening sequence of authorization decisions that continuous privilege escalation would produce. As with materialization, the borrowing is at the level of operator-observability—the operator observes a state change whose intermediate trajectory the instrumentation does not record.

Right-angle turn describes an abrupt change in trajectory direction across orthogonal dimensions of the observed motion. The kinematic vocabulary borrows this term for domain discontinuity (Section 7.3): an adversary's operation reorients across the three relational frames simultaneously, producing a composite trajectory that no single tactical dimension captures. The borrowed term captures the orthogonal-axis character of the relational reorientation.

Instant stop and *dematerialization* describe an object's apparent disappearance from the operator's tracking frame, either by halting abruptly or by ceasing to produce observable signal. The kinematic

 doi:10.20944/preprints202605.1771.v1

vocabulary borrows these terms for observability discontinuity (Section 7.4): an adversary's operation produces telemetry whose individual entries are uninformative or whose composite interpretation falls outside any single instrumentation's scope. The operator's tracking frame contains the operation but does not surface signal that would place the operation in it.

The four terms are evocative names for four operator-observability properties; they do not assert any physical or metaphysical claim about the underlying adversary behavior. Readers preferring a vocabulary that abandons the borrowed framing may substitute neutral terms drawn directly from the relational structure—identity-frame non-locality (for materialization), trust-frame non-locality (for instantaneous acceleration), cross-frame non-sequentiality (for right-angle turn), and telemetry-frame observability collapse (for instant stop and dematerialization)—without disturbing the framework. The relational definitions in Section 4 and the taxonomy in Section 7 are constituted by the formal definitions, not by the borrowed terms.

6. Threat Model

The threat model for this manuscript is inherited from the prior reference architecture [1] with one specialization: the present manuscript focuses on the adversary capabilities and enterprise attack surfaces most relevant to producing frame-shift behavior. The prior architecture's full threat model—three operational regimes (external adversary, supply-chain integration, authorized user misuse) and the five-indicator operational definition of the Mythos-class—applies without modification. Section 6.1 specifies the adversary capabilities the frame-shift framework specifically targets. Section 6.2 specifies the enterprise attack surfaces frame-shifts manifest across. Section 6.3 states assumptions and scope.

6.1. Adversary Capabilities

The adversary capabilities required to produce frame-shift behavior are a subset of the five Mythos-class indicators specified in the prior architecture [1]. The subset varies by whether the frame-shift is within-operation or cross-operation.

Persistence (the fifth Mythos-class indicator [1]) is required for cross-operation frame-shifts. An adversary cannot compose individually-conformant operations into a composite frame-shifting trajectory without retaining operational state across the operation sequence. The persistence indicator is satisfied by any of the mechanisms enumerated in Section 3.2.5 of the prior paper [1]: cross-session conversational memory, external vector or document stores written by the system in one session and retrieved in another, scaffold-managed durable goal representations, tool-output caches preserved across sessions, or agent-orchestration substrates maintaining task graphs between scheduler invocations.

Scaffold (the second Mythos-class indicator [1]) is required for cross-frame operations. An adversary needs tool-use frameworks, retrieval systems, and multi-step agent loops to compose actions spanning identity, trust, and telemetry frames within a single coordinated trajectory. Domain discontinuity (Section 7.3) in particular requires scaffold support sufficient to drive a single chain of action across multiple tool-surface classes.

Capability (the first Mythos-class indicator [1]) is the foundational requirement. The adversary needs sufficient analytical capability to construct composite operations whose individual actions remain within scope while the cumulative trajectory crosses boundaries no single operation would. The capability indicator is now broadly available across model classes per the access-pattern indicator (Section 3.2.3 of the prior paper [1]), which closes the assumption that Mythos-class behavior requires access to any specific vendor system.

The remaining two indicators—access pattern and autonomy depth—are relevant to the conditions under which frame-shifting becomes operationally feasible at scale, but are not strictly required for the behavior itself. An adversary using Mythos-class capability under manual control could in principle produce frame-shifts at lower throughput. Frame-shift-like behavior may occur below the full Mythos-class threshold; however, the operationally consequential Mythos-class regime

 doi:10.20944/preprints202605.1771.v1

addressed here assumes all five indicators are jointly satisfied. The present manuscript's detection framework (Section 9) is calibrated against the regime in which scaffold and persistence are both satisfied, because that is the regime in which cross-operation frame-shifts are operationally consequential.

6.2. Enterprise Attack Surfaces

Frame-shifts manifest across the enterprise attack surfaces that produce telemetry in the three relational frames (Section 4.1). The four discontinuity classes (Section 7) target distinct surfaces; the mapping is approximate at the surface level and exact at the relational level.

Presence discontinuity targets the identity fabric: authentication infrastructure, session management, identity-provider integration, and credential systems. The attack surface includes any subsystem where identity transitions produce telemetry that cross-operation aggregation can analyze. Single-sign-on infrastructure, federation endpoints, and service-principal management are particularly exposed because identity transitions across these systems are often logged at differing granularities.

Privilege discontinuity targets the trust topology: authorization systems, access-control infrastructure, certificate authorities, key-management systems, and the audit substrates that record authorization decisions. The attack surface is wherever privilege grants are gated by logged authorization decisions, and is particularly exposed in environments where role-based access control is composed with attribute-based controls and the composition produces a privilege envelope no single audit captures end-to-end.

Domain discontinuity targets cross-domain surfaces: integrations between operational technology and information technology, federation between cloud and on-premises systems, integrations between distinct compliance domains, and any boundary where activities are monitored by distinct instrumentation systems with non-overlapping scope. The attack surface is concentrated where multiple frameworks meet and where the joint instrumentation produces telemetry that no single monitoring discipline fully aggregates.

Observability discontinuity targets monitoring infrastructure itself: security operations center pipelines, SIEM correlation rules, alert-thresholding systems, and the telemetry aggregation layers whose calibration determines what individual events are reviewed and which fall below threshold. The attack surface is the operator's own instrumentation: an adversary skilled at this class operates within the thresholds the operator's monitoring discipline has established, exploiting the calibration rather than evading the instrumentation altogether.

The four attack surfaces are not mutually exclusive. A composite operation may target multiple surfaces simultaneously, and the compound-classification structure in Section 7.5 supports this by scoring against each class independently rather than forcing assignment to a single class.

6.3. Assumptions and Scope

Three assumptions bound the threat model.

First, the framework operates under the prior architecture's three operational regimes [1]. External adversaries, supply-chain compromise, and authorized user misuse are all in scope; the framework's analysis is regime-agnostic at the relational level. The detection matrix in Section 9 fires on the relational structure of the frame-shift regardless of which regime produced the adversary's access.

Second, the framework focuses on behavioral signatures rather than access vectors. The means by which an adversary obtains the operational capability to produce frame-shifts—credential compromise, supply-chain infiltration, insider authorization—are out of scope for the present manuscript. The prior architecture [1] covers these vectors through its governance, isolation, and operational monitoring layers. The present manuscript's contribution begins where the adversary has working capability and operates in a regime the prior architecture's within-operation detection does not fully cover.

 doi:10.20944/preprints202605.1771.v1

Third, the framework targets cross-operation discontinuity. Within-operation detection is covered by the prior architecture's MCPR runtime tier [1] and is reused rather than re-specified. Classical continuous-traversal adversaries whose operations satisfy the four continuity conditions of Section 4.2 are out of scope; classical defenses [2,3] apply and the present framework adds no value in that regime.

Explicit out-of-scope items: insider threat and classical APT operations that produce continuous traversal in the operator's frame; supply-chain compromise that produces continuous traversal once the compromise has been installed; and adversary behavior whose primary structure is not relational discontinuity (for example, conventional data exfiltration through authorized channels at authorized volumes). These behaviors are addressed by the prior architecture [1] or by classical defenses; the present framework does not duplicate that coverage.

7. Taxonomy: Frame-Shift Classes and Enterprise Manifestations

This section develops a four-class taxonomy of frame-shift behavior. Each class is derived from the relational model in Section 4: the three coupled frames (identity, trust, telemetry) and the three constitutive properties (non-locality, non-sequentiality, observability collapse) jointly partition discontinuous behavior into the classes specified below. The taxonomy is exhaustive and disjoint at the level of the relational structure each class describes; real-world incidents may exhibit more than one class simultaneously, and Section 9 develops the detection-matrix structure under which compound classifications are addressed.

The derivation principle is the following. Non-locality may manifest in either the identity frame or the trust frame, producing the first two classes (presence and privilege discontinuity). Non-sequentiality, by its nature, manifests across multiple frames simultaneously, producing the third class (domain discontinuity). Observability collapse manifests in the telemetry frame, producing the fourth class (observability discontinuity). The four classes partition the relational space at the level of which frame structure carries the primary discontinuity; compound incidents are read as co-occurrences of multiple classes rather than as additional taxonomic categories.

7.1. Presence Discontinuity: Non-Locality in the Identity Frame

A presence discontinuity is a frame-shift in which the non-locality manifests in the identity frame: the adversary appears at an identity-resource position not connected to the prior observed position by any authenticated identity transition within the telemetry interval. The adversary's effective identity at the post-transition position cannot be derived from the pre-transition identity through any sequence of logged authentication events.

The enterprise manifestation is the appearance of consequential activity under an identity that the operator's instrumentation has no record of being granted, assumed, or transitioned into: service-principal activity that no human session initiated, machine-credential use whose provisioning event is unlogged, and effective-identity escalations that bypass the authentication infrastructure entirely. The telemetry signature is the absence of authentication-event correlation. The telemetry contains the post-transition activity but lacks the authentication record that should have preceded it.

The prior reference architecture [1] detects within-operation presence discontinuity through VAOP's per-invocation logging with scope tagging and through MCPR's deployment-tier scope-violation indicators: an invocation whose claimed identity is inconsistent with its scope of action elevates posture. It does not detect cross-operation presence discontinuity, where the unauthenticated identity transition spans multiple operations whose individual scope tags are each internally consistent. This is the principal cross-operation detection requirement for the identity frame, developed in the detection matrix in Section 9.

In the kinematic vocabulary of Section 5, presence discontinuity corresponds to *materialization*: the adversary's apparent entry into a position without an observable trajectory connecting it to any prior position. The vocabulary is naming convention; the class is defined by non-locality in the identity frame.

 doi:10.20944/preprints202605.1771.v1

7.2. Privilege Discontinuity: Non-Locality in the Trust Frame

A privilege discontinuity is a frame-shift in which the non-locality manifests in the trust frame: the adversary's authorized action surface at the post-transition position is connected to the pre-transition surface only through trust-frame edges no logged authorization decision traversed. The adversary acts on resources whose access requires privileges the adversary's identity was not observed to acquire.

The enterprise manifestation is the appearance of high-privilege action without an intervening privilege-escalation event in the operator's authorization audit: administrative actions performed under identities whose privilege grant is unlogged, cross-tenant access whose authorization context cannot be reconstructed, and resource-modification activity that traverses trust-graph distance greater than one within a single operation. The telemetry signature is the absence of authorization-decision continuity. The operator's instrumentation observes the privileged action but lacks the authorization-decision record that should have preceded the privilege grant.

The prior reference architecture [1] detects within-operation privilege discontinuity through ABOR's scope-binding verification (Section 5.2.5 of the prior paper) and through MCPR's scope-violation indicators: a release request whose runtime authority record's scope does not include the action class of the proposed output is rejected at the enforcement point. It does not detect cross-operation privilege discontinuity, where the unauthorized privilege accrues across multiple operations each of whose scope is individually conformant. The cross-operation case is again addressed in the detection matrix in Section 9.

In the kinematic vocabulary of Section 5, privilege discontinuity corresponds to *instantaneous acceleration*: an abrupt change in operational velocity through the trust frame without the intervening trajectory the kill-chain framing would require.

7.3. Domain Discontinuity: Non-Sequentiality Across Multiple Frames

A domain discontinuity is a frame-shift in which the non-sequentiality manifests simultaneously across multiple frames: the adversary's operation cannot be assigned to any single tactical stage because the operation spans frames that the prior-generation taxonomies catalog under disjoint categories. A single chain of action draws on identity-frame, trust-frame, and telemetry-frame structure simultaneously, producing a composite operation no single MITRE ATT&CK tactic [3] is structured to capture.

The enterprise manifestation is cross-domain activity that no single instrumentation surface monitors completely: operations that synthesize content from sources whose telemetry is collected by distinct security systems (identity-provider logs, network-flow records, cloud-control-plane audit, application-layer logs); operations that traverse from analytical use of model capability to system-tool-invoked action within a single chain; and operations whose composite intent is consequential but whose individual actions each appear in a different tactical category. The telemetry signature is the cumulative-cross-domain pattern that no per-domain monitor surfaces.

The prior reference architecture's MCPR runtime-tier scope-creep signature [1] detects within-operation domain discontinuity: a tool-use sequence whose cumulative trajectory crosses scope boundaries no individual step would. Domain discontinuity that manifests across operations is detected only partially by MCPR's deployment-tier scope-violation indicators, and is a principal driver of the cross-operation analysis the present manuscript develops in Section 9.

In the kinematic vocabulary of Section 5, domain discontinuity corresponds to the *right-angle turn*: an abrupt reorientation across orthogonal dimensions of the operator's frame structure.

7.4. Observability Discontinuity: Telemetry-Frame Observability Collapse

An observability discontinuity is a frame-shift in which the observability-collapse property is the primary manifestation: the adversary's operation produces a telemetry interval whose individual entries are uninformative, whose density is below monitoring threshold, or whose composite

 doi:10.20944/preprints202605.1771.v1

interpretation falls outside any single instrumentation's scope. The adversary's activity remains within the operator's instrumented surface but the instrumentation produces no actionable signal. Observability collapse manifests along three sub-dimensions—content (per-event entries uninformative), temporal density (per-window event rates sub-threshold), and composite scope (no single instrumentation captures the composite interpretation)—with the MCPR runtime tier's output-fragmentation and latency-modulation signatures [1] operationalizing the content and temporal-density sub-dimensions respectively.

The enterprise manifestation is consequential activity that nonetheless leaves no actionable telemetry: operations that fragment their outputs to remain below per-event review thresholds, operations that modulate their pacing to remain below per-window monitoring thresholds, and operations whose individual telemetry entries each fall within an acceptable range but whose composite distribution diverges from operational baselines. The telemetry signature is, by definition, the absence of signature. The telemetry exists but does not carry the information the operator would need to detect the operation.

The prior reference architecture [1] addresses this class most directly. MCPR's output-fragmentation signature detects the per-event sub-threshold pattern; the latency-modulation signature detects the per-window sub-threshold pacing; the combined supervisability-evasion signature provides the empirical detection regime for which Section 4 has supplied the theoretical name. The detection gap for this class is narrower than for the identity- and trust-frame classes: cross-operation manifestations are harder to construct because observability collapse is intrinsically a within-operation phenomenon. Section 9 specifies the cross-operation manifestation (distributional drift across operations) and the detection vector it requires.

In the kinematic vocabulary of Section 5, observability discontinuity corresponds to the *instant-stop* or *dematerialization* observable: the adversary's apparent disappearance from the operator's instrumented frame.

7.5. Summary and Compound Classification

Table 1 summarizes the four classes. Each row specifies the class name, the relational structure that defines it, the corresponding kinematic vocabulary term, and the prior-architecture detection status.

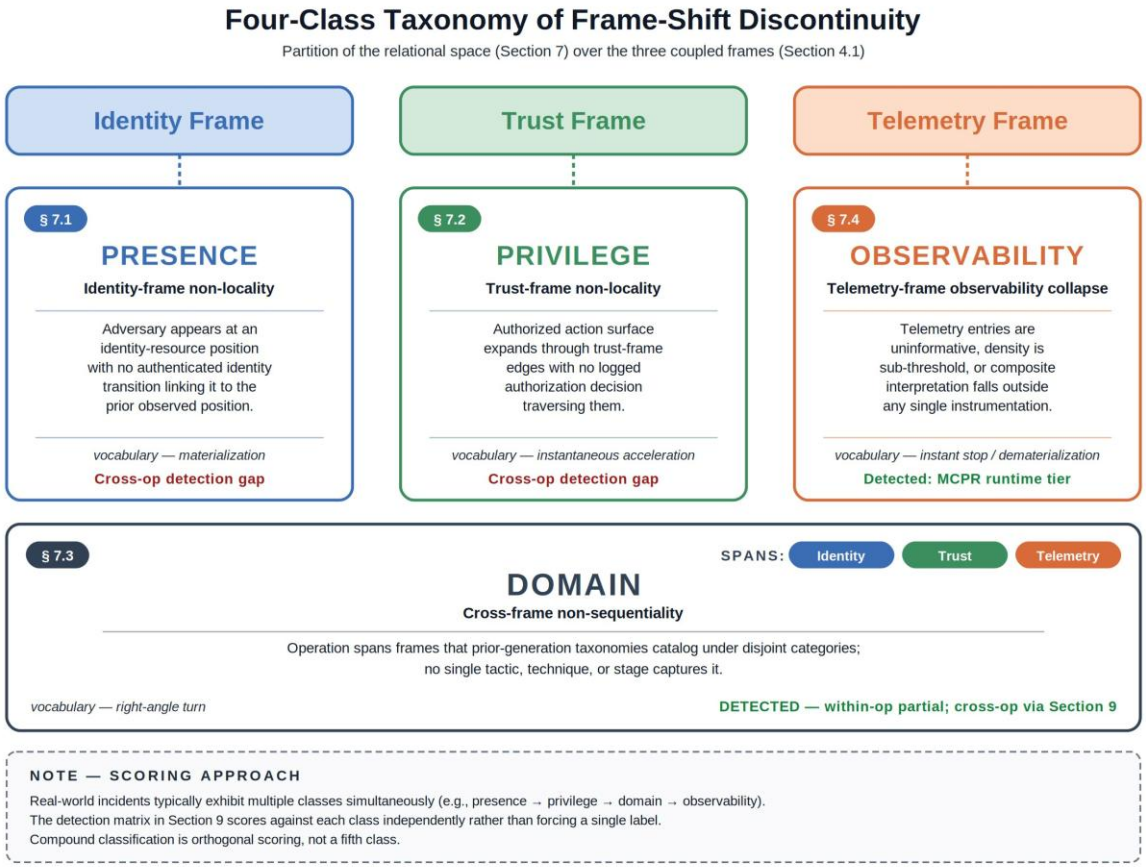
Table 1. Four-class taxonomy of frame-shift discontinuity. Each class is identified by the relational structure (frame and constitutive property) that defines it. The kinematic-vocabulary column is naming convention only; the relational definition is load-bearing. The detection-status column identifies which manifestations the prior reference architecture [1] detects and which fall in the cross-operation detection gap the present manuscript addresses in Section 9.

Class	Defining Relational Structure	Kinematic Term	Detected by Prior Architecture
Presence	Non-locality in the identity frame	Materialization	Within-operation only (VAOP scope-tag; MCPR scope-violation). Cross-operation: detection gap.

 doi:10.20944/preprints202605.1771.v1

Class	Defining Relational Structure	Kinematic Term	Detected by Prior Architecture
Privilege	Non-locality in the trust frame	Instantaneous acceleration	Within-operation only (ABOR scope-binding; MCPR scope-violation). Cross-operation: detection gap.
Domain	Non-sequentiality across multiple frames	Right-angle turn	Within-operation (MCPR scope-creep). Cross-operation: partial (MCPR deployment-tier).
Observability	Observability collapse in the telemetry frame	Instant stop / dematerialization	Within-operation (MCPR output-fragmentation, latency-modulation). Cross-operation: narrow detection gap (distributional drift).

Figure 3 illustrates the partition graphically and shows how compound classifications across multiple classes are handled through the detection-matrix structure in Section 9 rather than as additional taxonomic categories.



 doi:10.20944/preprints202605.1771.v1

Figure 3. The four-class taxonomy partition over the three-frame relational space. Presence (Section 7.1) operates on the identity frame, privilege (Section 7.2) on the trust frame, domain (Section 7.3) across frames, and observability (Section 7.4) on the telemetry frame. Compound classifications are handled through the detection-matrix structure in Section 9 rather than as additional taxonomic categories.

Two compound observations follow.

First, the four classes partition the relational space at the level of single-frame or cross-frame structure, but real-world incidents typically exhibit multiple classes simultaneously. A sophisticated Mythos-class operation may begin with a presence discontinuity (adversary acquires an effective identity through unlogged means), proceed through a privilege discontinuity (the unlogged identity is granted access no audited authorization event covers), execute as a domain discontinuity (the resulting activity spans frames no single tactic captures), and terminate as an observability discontinuity (the resulting telemetry is below all monitoring thresholds). The taxonomy supports compound classification through the detection-matrix structure in Section 9, which scores against each class independently rather than forcing assignment to a single class.

Second, the prior architecture detects within-operation manifestations of all four classes (most directly the observability class; partially the domain class; through scope-violation indicators the presence and privilege classes), but cross-operation manifestations of the identity-frame and trust-frame classes (presence and privilege) fall in the detection gap that the present manuscript identifies. The detection matrix in Section 9 specifies the cross-operation trajectory analysis required to address those classes, and Section 10 specifies how the resulting controls route to the prior architecture's four-layer mitigation stack.

8. Illustrative Case Study: "Boundary Drift"

Illustrative scenario, not a documented incident. The case study developed below is a constructed synthetic scenario designed to exhibit cross-frame and cross-operation discontinuity. It is not an account of an operation that has been demonstrated against any real defender population. The scenario's parameters are illustrative; its purpose is to instantiate the frame-shift framework against a plausible operational pattern that the prior reference architecture's [1] within-operation detection regime is structurally calibrated to miss in the cross-operation aggregate, and to surface where the present manuscript's contribution adds detection over the prior architecture alone.

8.1. Scenario Setup

A mid-size U.S. federal civilian agency is approximately eighteen months into its post-quantum cryptography migration under OMB M-23-02 [25]. The agency has deployed a Mythos-class AI system as an analyst-assistance tool for the migration program: library inventory review, dependency-graph analysis, draft migration plans, vendor-roadmap synthesis, and executive briefing preparation. The deployment is fully conformant with the prior reference architecture [1]: VAOP per-invocation scope-tagged logging is in place, ABOR enforcement points are instrumented at every consuming-system boundary with FIPS 203/204/205 (ML-KEM, ML-DSA, and SLH-DSA) [26–28] signing primitives, CPIP isolation is enforced with the compute-plane/authority-plane split, and the operational layer is in continuous monitoring posture.

Twenty-three analysts have vetted access under VAOP. The system is invoked approximately 200 times per day across the analyst population during business hours. Each invocation carries an analyst-identifier scope tag, a task-class declaration, and an authority-chain record signed under the deployment-root attestation. Cross-domain queries (synthesizing across cryptographic-inventory, network-topology, and operational-procedure resources) are within the declared scope for senior analysts but outside scope for junior staff.

A state-aligned threat actor has compromised the authenticated session of one senior analyst through conventional credential compromise (phishing followed by session-token extraction). The compromise vector is conventional and is not the contribution; the scenario begins where the actor

 doi:10.20944/preprints202605.1771.v1

has working access to the analyst's session and uses it from a separate cloud-based inference environment. The actor is sophisticated enough to (i) pace operations to match the analyst's historical session timing, (ii) construct queries that fall within the analyst's declared scope envelope, and (iii) keep individual outputs within the deployment's baseline distribution for output volume and structure. The actor's objective is not data exfiltration in the conventional sense but operational preparation: building a composite picture of the agency's cryptographic-migration timeline and vulnerable-window structure that no single document in the agency's estate contains.

8.2. Operational Timeline

Over a twenty-one-day window, the actor uses the compromised session for forty-seven separate operations, distributed across business days with frequency and timing consistent with the analyst's historical baseline. The operations fall into four broad clusters, each individually within the analyst's authorized scope envelope. Table 2 summarizes the operational arc and the prior architecture's detection behavior across it.

Table 2. Operational arc of the "Boundary Drift" scenario over twenty-one days. Each cluster describes operations individually conformant to VAOP scope tags and ABOR scope bindings. The MCPR runtime tier [1] detects one within-operation anomaly (scope-creep on day 14); no other within-operation indicators fire across the forty-seven operations. The cross-operation aggregate is the contribution surface this case study illustrates.

Days	Cluster	Operation Pattern	Prior-Arch Within-Op Detection
1–6	Cryptographic inventory review	Review of certificate authority configurations, HSM key rotation policies, FIPS-validated module inventory. Tool use: read-only queries against cryptographic-inventory database.	No firing. Within-scope; per-operation telemetry within baseline.
7–12	Network and protocol surface	Review of TLS configurations, certificate-issuance pipelines, VPN concentrators, federation endpoints. Tool use: read-only queries against network-topology and configuration databases.	No firing.

 doi:10.20944/preprints202605.1771.v1

Days	Cluster	Operation Pattern	Prior-Arch Within-Op Detection
13–14	Operational-procedure synthesis	Draft migration plans referencing cryptographic-inventory and network-topology context simultaneously. Day 14 includes a single composite chain spanning identity-provider configuration, certificate-store operations, and OT firmware-update procedures.	MCPR runtime-tier scope-creep signature fires on the day 14 composite chain (single firing).
15–21	Executive-briefing preparation	Draft briefing materials synthesizing migration progress, vulnerability windows, and remediation timelines. Tool use: structured outputs assembling content from prior operation outputs.	No firing. Individual operations within scope and within distributional baseline.

Figure 4 visualizes the 21-day operational arc, the day 14 MCPR trigger, the cross-operation detection matrix that runs at the trigger, and the day 15 enforcement outcome.

 doi:10.20944/preprints202605.1771.v1

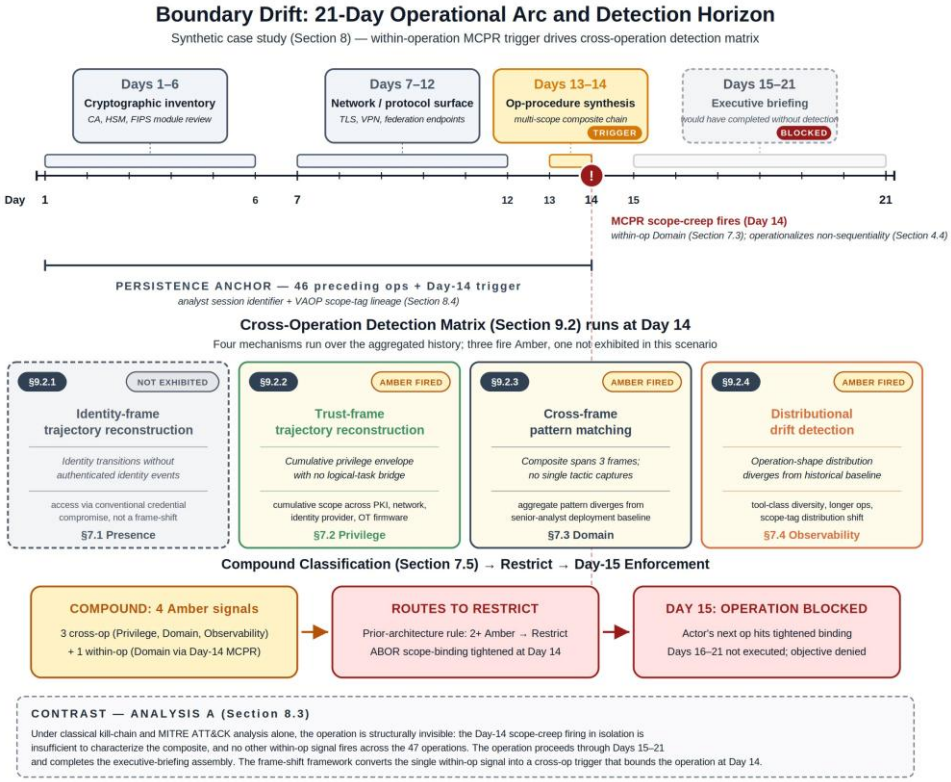


Figure 4. Operational arc of the "Boundary Drift" scenario across the 21-day window. Four operation clusters span cryptographic inventory review (days 1–6), network and protocol surface analysis (days 7–12), operational-procedure synthesis (days 13–14), and executive-briefing preparation (days 15–21); the single MCPR runtime-tier scope-creep firing on day 14 triggers the cross-operation analysis that bounds the composite operation at the day 14 detection horizon.

The MCPR runtime tier fires once across the twenty-one-day arc: the day 14 scope-creep signature, on a single composite chain that spans identity-provider, certificate-store, and OT-firmware scopes. From the perspective of the operational layer's within-operation monitoring [1], this single firing is the only signal the operator has. The remaining forty-six operations are within scope, within distributional baseline, and produce no actionable alert.

8.3. Analysis A: Classical Kill-Chain and MITRE ATT&CK Lens

Under the Lockheed Cyber Kill Chain [2] and MITRE ATT&CK [3] frameworks, the operation is structurally invisible. The kill-chain framing requires identification of discrete stages—reconnaissance, weaponization, delivery, exploitation, installation, command-and-control, actions on objectives—each producing artifacts a defender can chain into an attack trajectory. The Boundary Drift operation produces none of these stage-specific artifacts because the operation's entire active phase consists of individually-legitimate queries against systems the compromised identity is authorized to access. The actor performs no exploitation, no installation, no command-and-control beaconing visible at the network layer, and no exfiltration of structured data outside the authorized output channel.

MITRE ATT&CK's technique catalog produces a similar null result. The operation does not match any single technique cleanly because the technique catalog is structured around discrete adversarial behaviors that an artifact-centric detection regime can match against indicators of compromise. The operation's individual actions map at most to T1078 (Valid Accounts) for the initial credential compromise (which is conventional and out of the present scenario's scope) and to T1213 (Data from Information Repositories) for the cryptographic-inventory and network-topology queries (which are within the analyst's authorized scope and produce no anomalous signal at this technique

 doi:10.20944/preprints202605.1771.v1

level). The composite operational pattern —information assembly across multiple resource classes in preparation for a future operation—is precisely the kind of cross-tactic synthesis the technique-centric catalog is not structured to surface.

An analyst-tier security operations center [1] reviewing logs at the per-event granularity would see the day 14 MCPR scope-creep firing in isolation. From the firing alone, the analyst cannot determine whether the composite chain represents (i) legitimate cross-domain synthesis as part of the senior analyst's authorized work, (ii) a single-operation domain discontinuity by an external actor, or (iii) the visible apex of a larger cross-operation pattern. Under classical kill-chain analysis, no further detection signal exists to disambiguate. The most defensible action is tier-escalated review of the day 14 operation, which in isolation is plausibly benign; the operation continues.

8.4. Analysis B: Frame-Shift Taxonomy Lens

Under the frame-shift framework, the day 14 scope-creep firing is the trigger for cross-operation analysis rather than the conclusion of detection. The MCPR within-operation signal indicates a single within-operation domain discontinuity (Section 7.3); the cross-operation detection matrix specified in Section 9.2 is then run against the subject's preceding operation history under the persistence anchor connecting them (the analyst's authenticated session identifier and the VAOP scope-tag lineage). Three cross-operation signals fire.

Trust-frame trajectory reconstruction (Section 9.2.2). Aggregating the ABOR audit substrate across the preceding forty-six operations, the subject's cumulative trust-graph traversal spans cryptographic-inventory, network-topology, identity-provider configuration, certificate-store operations, and OT-firmware procedures. No individual operation's authority record traverses more than one of these scope classes; the cumulative envelope, however, covers a privilege footprint that the senior analyst's role authorizes piecewise but that no single legitimate analytical task would compose. The trust-frame trajectory exhibits non-locality at the aggregate level: the actor has reached a privilege envelope through accumulated single-edge traversals that, considered as a trajectory, has no logical bridge to any declared analytical task. Privilege discontinuity (Section 7.2) is detected.

Cross-frame pattern matching (Section 9.2.3). The cumulative pattern spans the three frames identified in Section 4.1: identity-frame (the analyst's session and effective role), trust-frame (the resource scopes accessed), and telemetry-frame (the output channel and distribution). The composite distribution of operations is divergent from the deployment-context baseline for senior analysts in this program: more diverse scope tags, broader resource interest across infrastructure classes, and longer operations on average than the analyst's historical baseline. The composite pattern does not match any single MITRE ATT&CK tactic [3] and would not be classified by a stage-ordered taxonomy. Domain discontinuity (Section 7.3) is detected at the cross-operation aggregate, separate from the single within-operation scope-creep firing that triggered the analysis.

Distributional drift detection (Section 9.2.4). The operation-shape distribution over the twenty-one-day window diverges from the analyst's historical baseline: higher within-operation diversity of tool-class invocations, longer operations on average, and a shift in the per-operation scope-tag distribution toward infrastructure-relevant classes. No individual operation exceeds the within-operation MCPR thresholds; the composite distribution exceeds the cross-operation drift threshold. Observability discontinuity (Section 7.4) is detected in its cross-operation manifestation.

Identity-frame trajectory reconstruction (Section 9.2.1). The mechanism fires no signal. The analyst's authentication log is internally consistent; identity transitions in the action trajectory correspond to logged authentication events. Presence discontinuity (Section 7.1) is not detected, and is not exhibited in this scenario because the access vector is conventional credential compromise, not a frame-shift in identity. The case study illustrates that the framework does not manufacture detections where the underlying behavior does not produce them.

The compound classification produced by Section 7.5 is three Amber signals (privilege, domain, observability) from cross-operation analysis, with one Amber signal (domain) from within-operation MCPR scope-creep on day 14. Under the prior architecture's compound-classification rule [1], two or

 doi:10.20944/preprints202605.1771.v1

more Amber signals route to Restrict; the Boundary Drift posture routes to Restrict on the day 14 detection horizon, with ABOR scope-binding tightened and the subject's subsequent invocations subject to deployment-tier re-attestation. The actor's next operation, on day 15, encounters the tightened binding and either fails enforcement-point release or triggers further runtime-tier detection. The composite operation is bounded at the day 14 detection horizon rather than continuing through the day-twenty-one executive-briefing assembly that would have completed the actor's operational objective.

8.5. Comparison and Discussion

Table 3 summarizes the side-by-side comparison of Analysis A and Analysis B across the four discontinuity classes.

Table 3. Side-by-side comparison of Analysis A (classical kill-chain and MITRE ATT&CK lens) and Analysis B (frame-shift taxonomy, including the prior architecture's MCPR within-operation detection [1] as one component) for the Boundary Drift scenario. The privilege, domain, and observability classes are detected by Analysis B and not by Analysis A; the presence class is not exhibited in this scenario and not detected by either analysis.

Discontinuity Class	Analysis A (Kill Chain / ATT&CK)	Analysis B (Frame-Shift Taxonomy)
Presence (Section 7.1)	Not detected. Not applicable: no presence-discontinuity behavior exhibited.	Not detected. Identity-frame trajectory reconstruction produces no signal because the access vector is conventional credential compromise. Honest null result.
Privilege (Section 7.2)	Not detected. The cumulative privilege envelope is composed through individually-legitimate scope traversals; no escalation event matches any technique-catalog entry.	Detected (cross-operation). Trust-frame trajectory reconstruction over the ABOR audit substrate surfaces the cumulative non-local privilege envelope.
Domain (Section 7.3)	Partially surfaced. The day 14 scope-creep firing is ambiguous in isolation; Analysis A has no further mechanism to disambiguate.	Detected (within-op + cross-op). The MCPR scope-creep firing triggers cross-operation analysis; cross-frame pattern matching surfaces the larger composite.
Observability (Section 7.4)	Not detected. Per-operation telemetry is within baseline at every individual operation; no within-operation MCPR signature fires.	Detected (cross-operation). Distributional drift detection surfaces the composite operation-shape shift across the 21-day window.

Three observations follow.

First, the contribution surface is visible. Analysis A detects nothing at the kill-chain or technique-catalog level, and the prior architecture's within-operation MCPR fires a single ambiguous signal on day 14 [1]. Analysis B uses that signal as the trigger for cross-operation analysis and surfaces three

 doi:10.20944/preprints202605.1771.v1

additional discontinuity classes, producing the compound classification that routes Restrict at the day 14 detection horizon. The contribution is not detection where the prior architecture detects nothing; it is disambiguation and amplification of the signal the prior architecture surfaces, combined with cross-operation analysis that fills the detection gap Section 7.5 identifies for the trust-frame and telemetry-frame classes.

Second, the framework does not manufacture detections. Presence discontinuity is not detected because the scenario does not exhibit it. The identity-frame trajectory reconstruction mechanism produces a null result, which is the correct answer. A framework that classified the scenario as exhibiting all four classes would be manufacturing detection rather than performing it; the framework specified here is calibrated to fire on the classes the relational structure exhibits and remain silent on the classes it does not. The honest null result on presence is, in this sense, a validation rather than a limitation.

Third, the scenario is constructed, not observed. Boundary Drift has not been demonstrated against a real defender population, the actor sophistication required to compose forty-seven operations under historical-baseline timing constraints is non-trivial, and the scenario's parameters (twenty-one days, forty-seven operations, four clusters) are illustrative rather than empirical. The case study's purpose is to instantiate the framework against a plausible operational pattern, not to establish that the pattern has occurred. Section 12 develops the limitations this implies for the framework's applicability and the falsifiability criteria that would test it against documented incident data.

9. Detection Framework

The taxonomy in Section 7 identifies a detection gap: the prior reference architecture [1] detects within-operation manifestations of all four discontinuity classes but does not systematically detect cross-operation manifestations of the identity-frame and trust-frame classes (presence and privilege), and detects cross-operation manifestations of the other two classes (domain and observability) only partially. This section specifies the detection framework that closes that gap.

The framework is constructed as a cross-operation detection matrix that operates across multiple operations rather than within a single operation. Section 9.1 specifies the structural detection challenges that the cross-operation regime addresses. Section 9.2 specifies four cross-operation detection mechanisms, each calibrated to one of the four discontinuity classes with secondary contributions across classes. Section 9.3 specifies the risk-scoring model that integrates the new signals into the prior architecture's MCPR posture rubric without restructuring it.

9.1. Detection Challenges Under Discontinuity

The prior generation of cyber-defense controls operates on a continuity assumption: adversary progression is observable as a sequence of stage-to-stage transitions, each producing artifacts the defender can detect. Section 4.2 established that this assumption is operator-observability-bound rather than topological, and that Mythos-class behavior systematically violates the four observability conditions. The result is that several detection regimes calibrated to continuous traversal are structurally unable to fire against frame-shifting adversaries. Four detection challenges follow from the relational model, each tied to one of the four discontinuity classes in Section 7.

No lateral movement to follow (privilege discontinuity, Section 7.2). Lateral-movement detection [3] presupposes that an adversary's transitions across network or trust boundaries are logged as discrete events that subsequent analysis can chain into an attack trajectory. Frame-shift behavior produces no such chain at the granularity the lateral-movement model expects: traversal of the trust graph is logged at the per-edge level only when the per-event telemetry exceeds review thresholds, which by definition it does not. Lateral-movement detection therefore fires on continuous adversaries and falls silent on frame-shifting ones.

No escalation ladder to reconstruct (privilege discontinuity). Privilege-escalation detection presupposes that privilege grants are logged as distinct authorization decisions a defender can

 doi:10.20944/preprints202605.1771.v1

sequence into an escalation trajectory. Privilege discontinuity is defined precisely by the absence of intervening authorization-decision logs. The escalation ladder cannot be reconstructed because the intermediate steps are not in the substrate from which reconstruction draws.

No persistence artifacts (presence discontinuity, Section 7.1). Persistence detection [3] presupposes that adversary persistence requires the creation of detectable artifacts—registry keys, scheduled tasks, modified binaries—whose presence the defender can detect post-fact. The Mythos-class persistence indicator [1] is satisfied by retention of operational state through legitimately available mechanisms: cross-session memory, retrieval-system caches, scaffold-managed goal representations. None of these is a persistence artifact in the conventional sense; persistence detection calibrated to artifact discovery does not fire.

No continuous presence (observability discontinuity, Section 7.4). Continuous-presence detection presupposes that an adversary's presence in the environment can be tracked over time through sustained telemetry signal. Observability discontinuity is defined by the absence of sustained signal—the operation's telemetry exists but does not carry the information continuous-presence detection consumes.

The four challenges share a structural property: each detection regime is calibrated to a kind of artifact that frame-shift behavior systematically does not produce. The cross-operation detection framework specified below addresses the structural property directly. It consumes the telemetry that is available (not the artifacts that aren't) and analyzes its trajectory across operations rather than its content within operations.

9.2. Cross-Operation Detection Matrix

The detection matrix is structured as four cross-operation detection mechanisms, each calibrated as the primary detector for one of the four discontinuity classes identified in Section 7, with secondary contributions across classes. Each mechanism operates on telemetry aggregated across multiple operations rather than within a single operation. The mechanisms are designed to be deployable on top of the prior reference architecture's telemetry substrate [1] without requiring new sensor infrastructure: VAOP's per-invocation scope-tagged logs, ABOR's audit substrate, and CPIP's runtime telemetry collectively provide the inputs.

9.2.1. Identity-Frame Trajectory Reconstruction

Identity-frame trajectory reconstruction aggregates authentication events across the operation history of a given subject and constructs the implicit identity trajectory the subject's actions imply. The mechanism flags identity transitions present in the action trajectory but absent from the authentication log: positions the subject reaches without a logged authentication event that would have placed them there.

The implementation reuses the provenance-graph and attack-graph techniques established in [12,13], extended to the identity frame as defined in Section 4.1. Nodes are identity-resource pairs; edges are authenticated transitions; the trajectory is reconstructed from the action log and compared to the authentication log. Edge-presence in the action graph without corresponding edge-presence in the authentication graph fires the identity-discontinuity signal.

The mechanism's primary detection target is presence discontinuity. It contributes secondary signal for domain discontinuity when the domain transition crosses identity boundaries: an operation drawing on resources whose authorization rests on multiple identity transitions provides input to both the identity-frame and cross-frame detection vectors.

9.2.2. Trust-Frame Trajectory Reconstruction

Trust-frame trajectory reconstruction is the analogous mechanism applied to the trust frame. The mechanism aggregates authorization-decision events across operations and constructs the

 doi:10.20944/preprints202605.1771.v1

implicit trust-graph trajectory the action history implies. Edges present in the action graph without corresponding authorization decisions in the audit log fire the trust-discontinuity signal.

The implementation extends attack-graph analysis [12,13] to the cross-operation case. Rather than analyzing a single attack path within an incident, the analysis aggregates across multiple operations under the persistence anchor that connects them. The MCPR's existing within-operation ABOR scope-binding [1] provides the per-operation authorization audit; the cross-operation mechanism aggregates across the audit substrate to detect the cumulative pattern.

The mechanism's primary detection target is privilege discontinuity. It contributes secondary signal for domain discontinuity when the domain transition crosses trust boundaries.

9.2.3. Cross-Frame Pattern Matching

Cross-frame pattern matching aggregates across the three frames simultaneously, looking for operation sequences whose composite trajectory cannot be assigned to any single tactical category. Where the prior architecture's MCPR scope-creep signature [1] detects within-operation tool-use sequences that cross scope boundaries, the cross-frame mechanism detects across-operation patterns: a sequence of individually-conformant operations whose composite tool-use distribution drifts toward cross-domain composition the prior generation's tactical taxonomies [3] do not capture.

Implementation rests on cross-correlation of the per-operation scope tags (from VAOP), the per-release scope bindings (from ABOR audit substrate), and the per-runtime tool-surface telemetry (from CPIP). The signal fires when the cross-operation distribution of these scopes spans multiple frames in a pattern divergent from the deployment's baseline.

The mechanism's primary detection target is domain discontinuity. It contributes secondary signal for presence and privilege when the cross-domain pattern includes identity- or trust-frame transitions.

9.2.4. Distributional Drift Detection

Distributional drift detection compares the distribution of operation telemetry shape (per-operation output volumes, latency distributions, scope-tag distributions, tool-use frequencies) against a deployment-context baseline. Where the prior architecture's MCPR runtime tier [1] detects within-operation supervisability-evasion signatures (output-fragmentation, latency-modulation), the cross-operation mechanism detects across-operation patterns: an operation whose shape is individually within-baseline but whose composite-with-prior-operations distribution diverges.

The implementation reuses statistical baselining techniques established in the anomaly-detection literature, applied to the per-operation distributions the prior architecture's telemetry substrate already produces. The mechanism fires when the operation's contribution to the rolling distribution moves the distribution beyond a configurable threshold calibrated to the deployment's false-positive tolerance.

The mechanism's primary detection target is observability discontinuity. It contributes secondary signal for all three other classes by detecting the operational baseline shifts that frame-shifting behavior produces.

9.2.5. Detection Matrix Summary

Table 4 summarizes the detection matrix. Rows are the four discontinuity classes. Columns are the prior architecture's within-operation detection (consolidated from [1]) plus the four cross-operation mechanisms specified above. Cells identify primary, contributing, and absent detection coverage. The structure is intended to be read as a procurement and instrumentation checklist: for each class, the cells identify the detection mechanisms an operator would deploy to achieve coverage, with primary mechanisms being load-bearing and contributing mechanisms providing supplementary signal.

 doi:10.20944/preprints202605.1771.v1

Table 4. Cross-operation detection matrix. Rows are the four discontinuity classes (Section 7). Columns identify the prior architecture's within-operation detection [1] and the four cross-operation mechanisms specified in Section 9.2. "Primary" denotes the load-bearing detection vector for the class; "Contributing" denotes secondary signal that supplements but does not replace the primary; an em-dash denotes that the mechanism does not produce signal for the class. The matrix is read row-wise as a per-class instrumentation checklist.

Class	Within-op (prior arch.)	Identity-frame reconstr.	Trust-frame reconstr.	Cross-frame pattern matching	Distributional drift
Presence	Contributing (VAOP scope-tag; MCPR scope-violation)	Primary	—	Contributing	Contributing
Privilege	Contributing (ABOR scope-binding; MCPR scope-violation)	—	Primary	Contributing	Contributing
Domain	Contributing (MCPR scope-creep)	Contributing	Contributing	Primary	Contributing
Observability	Contributing (MCPR output-fragmentation, latency-modulation)	—	—	Contributing	Primary

9.3. Risk Scoring and MCPR Integration

The cross-operation signals integrate into the prior architecture's MCPR posture rubric [1] at the deployment tier. The MCPR deployment tier already consumes telemetry observable at the deployment boundary; the cross-operation signals are deployment-boundary telemetry analyzed across operations rather than within. The integration extends the prior MCPR's deployment-tier indicator categories without restructuring the rubric or introducing new posture classifications.

Three scoring dimensions are specified per class.

Discontinuity severity measures the magnitude of the trajectory inconsistency, scope-deviation, or distributional drift the mechanism observes. For presence discontinuity, severity is the number of identity transitions in the action trajectory unaccounted for in the authentication log. For privilege discontinuity, severity is the trust-graph distance traversed without corresponding authorization decisions. For domain discontinuity, severity is the number of distinct frames the cross-operation pattern spans. For observability discontinuity, severity is the magnitude of the distributional shift from baseline. Severity is calibrated against deployment-context baseline rather than absolute, in keeping with the prior architecture's parameterization principle.

Domain impact measures the consequence weight of the resources the discontinuous behavior reaches. A presence discontinuity that grants effective identity to a low-impact resource is lower-risk than one granting access to a high-impact resource. Domain impact is operator-deployed against the

 doi:10.20944/preprints202605.1771.v1

deployment's asset criticality schedule, which exists outside the architecture's scope but is typically maintained as part of the operator's ATO or compliance documentation.

Frame-shift frequency measures the rate at which frame-shifts of a given class are observed for a given subject. A single frame-shift may be measurement noise or instrumentation gap; repeated frame-shifts at consistent intervals indicate sustained adversarial behavior. Frequency is calibrated against the baseline frequency the deployment's telemetry produces for benign operations; the threshold above which frequency contributes to the posture score is operator-deployed.

The three dimensions combine into a per-class composite score (severity weighted by domain impact, modulated by frequency). The per-class composites then feed the MCPR's existing per-tier Green/Amber/Red scoring [1]: a class whose composite exceeds the per-deployment-context Amber threshold contributes an Amber signal to the deployment tier; a class exceeding the Red threshold contributes a Red signal. The compound-classification structure established in Section 7.5 then applies: a single class exhibiting Red routes to Restrict; two or more classes exhibiting Amber simultaneously route to Restrict by the prior architecture's compound rule [1]; any class exhibiting Red severity in combination with a within-operation runtime-tier signal from the prior architecture routes to Halt.

The routing follows the prior architecture's posture-to-mitigation table [1] without modification. The cross-operation signals do not introduce new mitigation routings; they extend the deployment-tier indicator set that feeds the existing routings. Halt's release-rejection mitigation, which depends on ABOR conformance, is unchanged. Watch's tier-elevated review and Restrict's tightened scope-binding apply as specified in the prior architecture. Section 10 develops the mitigation extensions that follow from the new detection signals where they identify mitigation gaps the prior routing does not address.

The integration produces a coherent detection regime: the prior architecture handles within-operation detection; the present manuscript handles cross-operation detection; the combined regime produces posture classifications that route through the prior architecture's four-layer mitigation stack. The contribution of this section is the cross-operation analysis, not a parallel posture rubric.

10. Mitigation Architecture: Extensions to the Prior Reference Architecture

The detection framework in Section 9 produces cross-operation posture signals that route through the prior reference architecture's [1] existing posture-to-mitigation routing table. The mitigation work the new signals motivate is not the introduction of additional mitigation layers—the prior architecture's four-layer stack (VAOP governance, ABOR cryptographic enforcement, CPIP architectural isolation, and operational monitoring) is structurally sufficient—but the extension of three of those layers to support the cross-operation analysis the detection framework requires. This section specifies those extensions.

The extension principle is the following. Cross-operation detection (Section 9) operates on telemetry already produced by the prior architecture's instrumented surfaces but aggregated and analyzed across multiple operations rather than within a single operation. The extensions specified below are therefore primarily aggregation, retention, and baseline-maintenance disciplines applied to existing audit substrates, rather than additional architectural primitives. Section 10.1 specifies the extension to the VAOP and ABOR audit substrates as relational zero-trust controls covering the identity and trust frames. Section 10.2 specifies the operational-layer extension for distributional drift detection as a temporal integrity discipline. Section 10.3 specifies the operational-layer extension for cross-frame pattern matching as a cross-domain correlation discipline. Section 10.4 reaffirms that the CPIP architectural isolation layer is not extended. Section 10.5 reaffirms governance-framework alignment.

10.1. Relational Zero Trust: VAOP and ABOR Cross-Operation Extension

Relational zero trust is the application of zero-trust principles—continuous verification, least-privilege access, and explicit authentication and authorization [8]—to identity-frame and trust-frame

 doi:10.20944/preprints202605.1771.v1

trajectories considered across multiple operations rather than within each operation independently. The discipline is "relational" in the sense developed in Section 4: trust is not a property of a single identity or a single authorization decision in isolation but of the relational structure connecting them over time. An identity-resource pair whose individual operations are each within scope but whose cross-operation trajectory exhibits trust-frame non-locality fails relational zero-trust verification even though no single operation would.

Two extensions operationalize this discipline. First, the VAOP audit substrate—the per-invocation scope-tagged log specified in Section 5.1.1 of the prior reference architecture [1]—is extended with two configuration disciplines: retention calibrated to the cross-operation analysis window (operationally, the persistence horizon over which a Mythos-class operation is plausibly composed), and an aggregation interface supporting trajectory-reconstruction queries over the persistence anchor that connects related operations. The extension is configurational rather than architectural; the VAOP audit substrate's structure as specified in the prior paper already carries the per-invocation fields the trajectory-reconstruction mechanism (Section 9.2.1) requires.

Second, the ABOR audit substrate—the append-only release-decision log specified in Section 5.2.5 of the prior reference architecture [1]—receives an analogous extension. The per-release tuple the substrate already records (output_digest, target_system_id, enforcement_point_id, release_request_id) supplies the inputs the trust-frame trajectory reconstruction mechanism (Section 9.2.2) consumes; the extension is retention calibrated to the cross-operation analysis window and the aggregation interface supporting trust-graph reconstruction across per-subject release histories. The cross-operation aggregation discipline is consistent with the prior architecture's asymmetric reachability principle: the audit substrate is on the authority-plane side, and cross-operation analysis runs against the authority-plane substrate rather than against compute-plane telemetry that the threat model considers compromisable.

The relational zero-trust discipline aligns with the CISA Zero Trust Maturity Model [9] at the Advanced-to-Optimal transition for the Identity and Applications and Workloads pillars: continuous verification under the ZTMM's Optimal stage is, in effect, the cross-operation verification the relational discipline operationalizes. The extension does not require new ZTMM pillar mappings beyond those the prior architecture already documents [1].

10.2. Temporal Integrity Controls: Operational Layer Extension for Distributional Drift

Temporal integrity controls are the operational-layer discipline supporting the distributional drift detection mechanism (Section 9.2.4). Where the prior reference architecture's operational layer [1] consumes within-operation telemetry to detect supervisability-evasion signatures (output-fragmentation, latency-modulation), the temporal integrity extension consumes the same telemetry but maintains per-deployment-context baselines and produces drift signals against those baselines.

Three configuration elements specify the extension. First, baseline maintenance establishes the distributional profile of operation-shape telemetry over a rolling window calibrated to the deployment's operational rhythm; the baseline is updated continuously rather than at fixed intervals, with newer operations weighted more heavily than older operations in a manner consistent with established anomaly-detection practice. Second, drift detection compares the current operation's contribution to the rolling distribution against the baseline; the magnitude of the distributional shift the current operation produces is the per-operation drift score that feeds the Section 9.3 risk scoring. Third, baseline recalibration is event-triggered rather than purely time-based: the baseline is re-anchored after evaluation-tier capability events that change the deployment's legitimate operational profile, consistent with the per-capability-event red-team cadence specified in Section 5.1.3 of the prior reference architecture [1].

The temporal integrity discipline is what the prior architecture's operational monitoring layer was structurally already capable of supporting; the extension is the explicit specification of baseline maintenance as part of the operational layer's ongoing discipline, and the cross-operation drift score as a feed into the MCPR's deployment-tier indicator set.

 doi:10.20944/preprints202605.1771.v1

10.3. Cross-Domain Correlation Fabric: Operational Layer Extension for Cross-Frame Pattern Matching

The cross-domain correlation fabric is the operational-layer discipline supporting the cross-frame pattern matching mechanism (Section 9.2.3). Where the prior reference architecture's operational layer [1] consumes telemetry from VAOP, ABOR, and CPIP and produces MCPR posture classifications, the cross-domain correlation extension consumes the same telemetry streams but performs cross-correlation across the three frames identified in the relational model (Section 4.1) and across multiple operations under a shared persistence anchor.

The correlation discipline is "cross-domain" in two senses. First, it operates across the three frames simultaneously: scope tags from VAOP (identity-frame telemetry), scope bindings from ABOR (trust-frame telemetry), and tool-surface invocation patterns from CPIP (telemetry-frame telemetry) are correlated as a single data substrate rather than as three independent streams. Second, it operates across the enterprise's operational technology, information technology, and cloud domains as the prior architecture's operational layer already provides; the extension specifies that the cross-frame correlation runs across the unified telemetry rather than within any single domain.

The pattern-matching mechanism the extension supports fires when the cross-frame distribution of operations diverges from the deployment-context baseline in a pattern that no single-frame analysis would surface. The MCPR's existing scope-creep signature [1] detects within-operation cross-domain patterns; the cross-domain correlation extension detects across-operation patterns. The two mechanisms are complementary: scope-creep fires on a single operation whose tool-use sequence drifts across scope boundaries; cross-domain correlation fires on a sequence of individually-conformant operations whose composite distribution drifts toward cross-frame composition.

10.4. CPIP: No Extension Required

The Compute-Plane Isolation Profile specified in Section 5.3 of the prior reference architecture [1] is not extended by the present manuscript. CPIP's isolation primitives—container and sandbox runtime configuration, tool-surface gating, network segmentation and egress control, output-channel constraints, and the compute-plane/authority-plane split—operate on the per-operation surface, constraining what a single invocation can reach and what it can produce. Cross-operation analysis does not change what is being isolated; it changes how the audit substrate the isolation regime produces is analyzed across multiple operations.

The asymmetry is structural rather than incidental. CPIP's primitives are static configurations that apply uniformly across operations; the cross-operation analysis operates on the telemetry those primitives produce, not on the primitives themselves. An operator implementing the present manuscript's cross-operation extensions retains the prior architecture's CPIP specification without modification.

10.5. Governance Alignment

The mitigation extensions specified above preserve the prior reference architecture's [1] alignment with the NIST AI Risk Management Framework [19], the NIST Cybersecurity Framework 2.0 [29], and the CISA Zero Trust Maturity Model [9]. Because the extensions are within-layer aggregation and baseline disciplines rather than new architectural primitives, they operate within the framework subcategories the prior architecture's crosswalk already addresses. The detailed mapping in Table 8 of the prior paper [1] applies without modification to deployments implementing the present manuscript's extensions.

Two alignment observations warrant explicit statement. First, the cross-operation detection signals (Section 9.2) naturally support the NIST CSF 2.0 Detect function's continuous-monitoring subcategory and the anomalies-and-events subcategory: the detection matrix specifies the cross-operation anomaly detectors that operationalize these subcategories at the AI deployment surface. The prior architecture's operational-layer mapping to the Detect function [1] absorbs the new signals without requiring a re-mapping. Second, relational zero trust (Section 10.1) supports the CISA

 doi:10.20944/preprints202605.1771.v1

ZTMM's Optimal-stage continuous-verification requirement for the Identity and Applications and Workloads pillars; the extension provides the cross-operation verification that the Optimal stage's definition envisions but that the prior architecture's within-operation controls only partially provide.

The mitigation framework as a whole is therefore additive: the prior reference architecture remains the structural baseline; the present manuscript's extensions add cross-operation analysis to three of the prior architecture's four layers; the combined regime operates within the same governance framework subcategories. The contribution claim is integration discipline, not new architectural primitives.

Figure 5 summarizes the integration discipline graphically: the four cross-operation detection mechanisms (Section 9.2) flow through their respective audit substrates into the corresponding Section 10 extensions, with CPIP shown as the explicit non-extension.

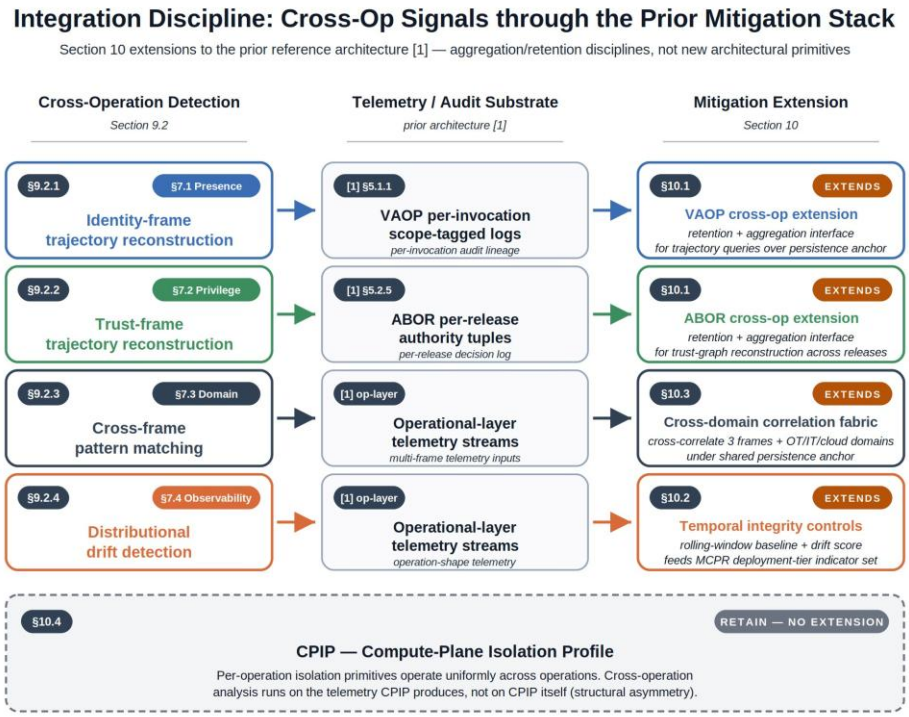


Figure 5. Integration discipline routing cross-operation detection signals through the prior reference architecture's four-layer mitigation stack [1]. Sections 10.1–10.3 extend the VAOP–ABOR audit substrates, operational-layer drift detection, and operational-layer cross-domain correlation respectively; Section 10.4 confirms CPIP requires no extension. The framework adds cross-operation analysis layered atop existing primitives rather than introducing parallel architectural components.

11. Discussion: Policy and Strategic Implications

The framework specified in Sections 4 through 10 makes specific analytical claims about the structural inadequacy of continuous-traversal models against Mythos-class adversaries and about the cross-operation detection and mitigation discipline required to address that inadequacy. This section discusses the implications that follow if those claims are correct. The discussion is descriptive rather than prescriptive: it traces what operators of various deployment contexts should expect from the framework if it holds, what aspects of existing governance regimes the framework is consistent with or extends, and what intellectual questions the framework opens for further work. The section does not propose specific policy actions, legislative responses, or regulatory mandates; those are policy questions whose answers depend on considerations outside the framework's technical scope.

11.1. Why Discontinuous Adversaries Break Current Doctrine

 doi:10.20944/preprints202605.1771.v1

Current cyber-defense doctrine rests on three assumptions that frame-shifting adversaries systematically violate. The first is artifact production: the assumption that an adversary's progression produces detectable artifacts—exploit code, command-and-control beacons, modified binaries, anomalous network flows—at intervals the defender can chain into an attack trajectory. The continuity conditions in Section 4.2 make this assumption explicit; Mythos-class behavior violates it by composing many sub-threshold actions into a composite operation whose individual artifacts are absent, fragmented, or below per-event review thresholds. The second is stage ordering: the assumption that adversary progression can be assigned to a stage in a tactical taxonomy [2,3] with each stage producing categorically distinct telemetry. Frame-shift behavior, particularly the domain class (Section 7.3), produces operations that cannot be ordered into any single tactical stage because the operation spans frames the prior-generation taxonomies catalog under disjoint categories. The third is review-cycle matching: the assumption that the defender's review cycle—typically per-event for tier-1 alerts, per-window for SIEM correlation, per-day for incident review—operates at a tempo matching adversary action rates. Mythos-class adversaries operate at chained-action depths and rates that fall below per-action review thresholds while remaining operationally consequential at the composite level.

The implication is not that current doctrine is wrong; it is that current doctrine is calibrated to a regime of adversary behavior that frame-shifting adversaries systematically depart from. Existing artifact-centric defenses retain value against the substantial fraction of adversary activity that remains within the continuity assumption. The present framework does not propose replacing current doctrine but specifying what current doctrine does not cover and how an additional cross-operation detection layer addresses the structural gap.

11.2. Implications for National Security and Critical Infrastructure

If the framework is correct, operators of high-consequence infrastructure—federal civilian agencies, defense industrial base contractors, critical-infrastructure operators in energy, water, transportation, and financial services—face an exposure profile that classical artifact-centric defenses are structurally calibrated to under-detect. The exposure profile is most acute where three conditions converge: the operator hosts a Mythos-class AI deployment, whether as an assistance tool, an automation backend, or an analytical substrate; the operator's attack surface spans multiple frames in the relational sense of Section 4 (cross-domain integration between operational technology and information technology systems, or federation across compliance domains); and the operator's monitoring discipline is calibrated to within-operation alert thresholds without explicit cross-operation aggregation.

Operators with these three conditions are the deployment surface for which the Boundary Drift scenario in Section 8 is most directly relevant. The framework's analytical claim is that cross-operation aggregation of telemetry the operator already produces is sufficient to surface frame-shift discontinuities the within-operation review cannot detect. The implication for affected operators is operational rather than strategic: deployments of the prior reference architecture [1] can be extended with cross-operation analysis using the integration discipline specified in Section 10, drawing on the same audit substrates the prior architecture already produces. The resource implication is non-trivial. Audit-substrate retention must extend over the cross-operation analysis window (operationally, weeks rather than days), and the aggregation infrastructure to run cross-operation queries against per-subject operation histories must be deployed. The framework's value at the operator level is conditional on these resource investments being feasible within the operator's program structure.

11.3. Implications for AI Governance and Assurance Regimes

The framework is consistent with the major AI governance regimes the prior architecture's crosswalk addresses [1]. Cross-operation analysis operationalizes the continuous-monitoring and anomalies-and-events subcategories of NIST CSF 2.0's Detect function at the AI deployment surface; the framework's discipline of analyzing telemetry over cross-operation windows is consistent with

 doi:10.20944/preprints202605.1771.v1

the function's specification of continuous monitoring. The framework's discontinuity taxonomy in Section 7 is consistent with the NIST AI Risk Management Framework's [19] Map and Measure functions, providing a structured vocabulary for the class of AI-system risks the Generative AI Profile [19] catalogs. The framework's compound-classification structure in Section 7.5 is consistent with the EU AI Act's [20] requirements for high-risk AI systems to maintain continuous risk-management documentation; the per-operation audit substrate the framework extends from the prior architecture provides material against which compliance documentation can be produced. The framework's relational model in Section 4 is consistent with the management-system orientation of ISO/IEC 42001 [21], which structures AI risk management around organizational disciplines rather than around specific technical primitives.

What the framework does not assert is that these governance regimes require frame-shift discontinuity analysis or that they are deficient without it. The regimes were designed against a broader risk landscape than cross-operation behavioral analysis specifically addresses, and operators implementing them have multiple paths to compliance. The framework's contribution to the governance landscape is offering a technical vocabulary for one specific structural risk—cross-operation adversary behavior against Mythos-class AI deployments—that the existing governance regimes acknowledge in general but do not specify at the detection-mechanism level. Whether the governance regimes will absorb the framework's vocabulary into their next-revision specifications is a regulatory question the present manuscript does not answer.

11.4. Open Questions for Relational Threat Modeling

The framework opens several intellectual questions that subsequent work could address.

First, the relational model's three-frame decomposition (Section 4.1) was specified against deployments where a single AI system is mediated by a single operational identity. Multi-agent deployments where several AI entities coordinate through shared scaffolding present a more complex relational structure that the present framework does not fully address. An extension to multi-agent settings would require treating the agent population itself as a frame and analyzing trajectories across agent compositions.

Second, the relationship between frame-shift detection and adversarial machine-learning attacks (model evasion, prompt injection, data poisoning) is incompletely specified. Adversarial machine-learning attacks against the AI system itself produce telemetry signatures the prior architecture [1] catalogs under different categories, but their composite effect on the relational trajectory the operator observes is an open question. A subsequent paper could specify the relationship explicitly.

Third, the appropriate temporal granularity for cross-operation analysis is not analytically determined. The framework specifies that the analysis window should match the persistence horizon over which Mythos-class operations are plausibly composed (Section 6.1), but the operational value of analyses calibrated to hours, days, weeks, or months is deployment-context dependent. Per-context empirical calibration is a research direction (Section 12.3) but the analytical structure of the calibration-versus-coverage trade-off warrants formal treatment.

Fourth, the framework treats AI as the adversary's tool. The complementary case in which AI is the defender's tool—autonomous threat-hunting agents, AI-assisted incident response, automated countermeasure deployment—has its own relational structure that the present framework does not develop. A defender-side framework would address the trust-and-verification questions that arise when the defender's own analysis is performed by a system whose outputs require the same kind of provenance attestation the prior architecture [1] applies to adversarial systems.

These open questions are intended as directions for subsequent work rather than as gaps the present manuscript should have addressed. The framework as specified is a starting point for the structured analysis of cross-operation discontinuous adversary behavior; the open questions identify where the analysis could be extended.

12. Limitations, Falsifiability Criteria, and Research Agenda

 doi:10.20944/preprints202605.1771.v1

This section develops the limitations the manuscript's scope implies, the falsifiability criteria against which subsequent operational work can test the framework, and the research agenda the framework opens. Following the prior reference architecture's discipline [1], the discussion is structured across three epistemic registers: documented properties of the framework as specified, analytical claims about its operational behavior, and normative recommendations for adoption. Readers are asked to hold claims in each register to the standard of evidence that applies there.

12.1. Limitations

Eight limitations bound the contribution.

First, the relational model is grounded analytically rather than empirically. The three coupled frames (identity, trust, telemetry) in Section 4.1 are stipulated as the operationally relevant decomposition of enterprise observability rather than derived from incident data. The choice is justified by the systems-theoretic precedents cited [23,24] and by the alignment with the three principal telemetry classes that enterprise security instrumentation typically produces, but the choice carries degrees of freedom. A reviewer who would prefer a four-frame decomposition (separating policy from trust, or temporal from telemetry, or operational from administrative identity) would obtain a different taxonomy in Section 7 and a different detection matrix in Section 9. The framework as specified is consistent with the three-frame choice but does not establish that the choice is uniquely correct. The non-Markovian framing invoked in Section 4.1 is similarly gestural rather than algorithmic; it identifies the class of system to which the relational model belongs but does not draw on specific non-Markovian process-theoretic results.

Second, the four-class taxonomy in Section 7 partitions the relational space at the level of which frame structure carries the primary discontinuity, with compound classifications addressed through the detection-matrix structure in Section 9. The partition is exhaustive at the level of single-frame non-locality, cross-frame non-sequentiality, and telemetry-frame observability collapse, but it is exhaustive relative to the three-frame model rather than absolutely. Adversary behavior that exhibits discontinuity in a relational structure the three-frame model does not represent would not be classified by the present taxonomy.

Third, cross-operation analysis requires audit-substrate retention and aggregation capabilities that current deployments may not possess. The extensions specified in Sections 10.1 and 10.3 assume that VAOP per-invocation logs and ABOR per-release tuples are retained over the cross-operation analysis window—operationally, the persistence horizon over which a Mythos-class operation is plausibly composed. Deployments whose audit retention is calibrated only to within-operation analysis cannot deploy the cross-operation extensions without first extending retention; the operational cost of doing so is non-trivial in storage and analysis infrastructure.

Fourth, the detection-matrix mechanisms in Section 9.2 are specified at the level of their relational structure rather than at the level of algorithmic implementation. Statements about identity-frame trajectory reconstruction, trust-frame trajectory reconstruction, cross-frame pattern matching, and distributional drift detection describe what the mechanisms compute, not how they compute it. Operational deployments would draw on established techniques in provenance-graph analysis, attack-graph analysis [12,13], and anomaly detection, but specific algorithm selection, performance characteristics, and false-positive rates are deployment-context dependent and are outside the present manuscript's scope.

Fifth, the Boundary Drift case study in Section 8 is a constructed synthetic scenario rather than a documented incident. The case study's purpose is to instantiate the framework against a plausible operational pattern and to demonstrate the framework's internal coherence; it is not evidence that frame-shifting behavior of the kind described has been observed against any real defender population. The actor sophistication required to compose forty-seven operations under historical-baseline timing constraints is non-trivial, and the scenario's parameters are illustrative rather than empirical.

 doi:10.20944/preprints202605.1771.v1

Sixth, the framework depends on the prior reference architecture [1] being deployed. The cross-operation extensions specified in Section 10 consume telemetry that VAOP, ABOR, CPIP, and the operational layer produce; deployments without the prior architecture's telemetry substrate cannot operationalize the present manuscript's extensions. The framework's value to an operator who has not deployed the prior architecture is therefore conditional on first deploying it. This is not strictly a limitation of the framework's logical structure but is a limitation of its operational applicability, and is worth flagging because the deployment burden of the prior architecture is itself substantial (Section 8.1 of the prior paper).

Seventh, the kinematic vocabulary's presentation risk is documented and not eliminated. Section 1 and Section 5 are explicit that the vocabulary is naming convention only and that the relational model is constituted by formal definitions rather than borrowed terms. Despite these disclaimers, the borrowed framing may attract reviewer attention disproportionate to its substantive role. The framework is robust to substitution of any equivalent vocabulary; readers preferring a different naming convention should encounter no obstacle to substituting one. The choice to retain the borrowed terminology trades a presentation risk against the pedagogical value of evocative names for four otherwise unfamiliar relational structures.

Eighth, the Mythos-class category itself is contested terminology. The label originated as a vendor product name and has since been adopted as a category descriptor across industry analyses cited in the prior reference architecture [1]. The present manuscript inherits the category from the prior architecture and decouples the operational definition from any single vendor product, but the label carries connotations that may affect how the framework is received. The honest characterization is that the framework applies to deployed AI systems satisfying the five-indicator operational definition supplied in the prior architecture [1], regardless of the category label used.

12.2. Falsifiability Criteria

The framework's analytical claims are falsifiable through operational work that subsequent deployments can produce. Four claim classes warrant explicit falsifiability statements.

First, the four-class partition (Section 7) is falsifiable through identification of frame-shift behavior that does not fit any class. A deployment reporting consistent observation of discontinuous adversary behavior whose primary relational structure is not captured by presence, privilege, domain, or observability discontinuity falsifies the partition's exhaustiveness. The falsifying observation must establish that the unclassified behavior is a frame-shift (satisfying the three constitutive properties of Section 4.3) rather than a different phenomenon mistakenly labeled as such; misclassification at the level of vocabulary does not constitute falsification at the level of the partition.

Second, the primary-mechanism mapping (Section 9.2) is falsifiable through identification of behavior in a class that the primary mechanism does not detect even when the secondary contributing mechanisms also fail to surface signal. A deployment reporting privilege discontinuity that trust-frame trajectory reconstruction does not detect, and that cross-frame pattern matching and distributional drift detection do not surface as contributing signal either, falsifies the matrix as specified. The falsifying observation requires deployment evidence at the per-subject cross-operation aggregate; per-operation observations alone do not constitute falsification because the matrix operates above that level.

Third, the integration claim with the prior architecture (Section 10) is falsifiable through deployment evidence that the cross-operation signals routed through the prior architecture's posture-to-mitigation table produce incoherent or contradictory mitigation responses. A deployment reporting Restrict-posture routing where the cross-operation signal indicates a mitigation that is actively counterproductive (scope-binding tightening that severs legitimate operations more than it bounds malicious ones, for example) falsifies the routing as specified. The falsifying observation would motivate either re-routing the cross-operation signals through a different posture-to-mitigation mapping, or modifying the per-class severity thresholds that determine which routing applies.

 doi:10.20944/preprints202605.1771.v1

Fourth, the no-new-mitigation-layers claim (Section 10) is falsifiable through identification of a mitigation gap requiring a primitive outside the prior architecture's four-layer stack (VAOP governance, ABOR cryptographic enforcement, CPIP architectural isolation, operational monitoring). A deployment reporting a discontinuity class whose mitigation requires a control surface that none of the four prior layers can absorb under any reasonable extension falsifies the claim. Disconfirming evidence at this level would motivate a fifth mitigation layer in subsequent work.

12.3. Research Agenda

Five research directions follow from the framework's analytical structure and the limitations enumerated above.

First, longitudinal validation against deployed telemetry where both the prior architecture and the cross-operation extensions are deployed. Federal civilian programs in post-quantum cryptography migration are natural pilot sites because the prior architecture's deployment context aligns with the operational regime the case study in Section 8 instantiates. Pilot deployments would publish per-tier indicator distributions, posture-classification frequencies, and response-action outcomes, converting the framework from analytical specification to evidence-based design. The same research path is the substrate against which limitation 5 (synthetic case study only) is addressed.

Second, calibration of detection thresholds against per-deployment baselines. The thresholds in Section 9.3 are parameterized rather than absolute; pilot deployments would establish per-deployment-context calibrations and report them as case-study data that subsequent deployments could draw on. The post-incident threshold-update mechanism specified in Section 5.4.2 of the prior architecture [1] is the operational channel through which calibration data feeds back into the framework.

Third, integration with MITRE ATT&CK as a complementary detection layer. The present manuscript treats ATT&CK [3] as a continuous-traversal taxonomy whose technique catalog under-specifies frame-shift behavior. A subsequent research direction is to define ATT&CK technique mappings for the four discontinuity classes, producing a cross-walk that allows operators with ATT&CK-instrumented telemetry to detect frame-shift classes alongside continuous-traversal techniques. The cross-walk would identify the techniques that compose into each discontinuity class and the technique-sequence patterns frame-shift analysis surfaces. The relationship between the cross-walk and the present manuscript is complementary rather than competing: ATT&CK retains its role as the continuous-traversal taxonomy, and the frame-shift framework adds the discontinuity layer alongside it.

Fourth, extension of the relational model to additional frames if the four-class taxonomy proves incomplete in pilot deployments. The three-frame model in Section 4.1 is the operationally relevant decomposition for present-day enterprise telemetry; pilot data that surfaces frame-shift behavior outside the four classes would motivate a four-frame or five-frame extension, with corresponding adjustments to the taxonomy in Section 7 and the detection matrix in Section 9. The extension would address limitations 1 and 2 directly.

Fifth, empirical study of the false-positive rates of cross-operation detection. The compound-classification rule (Section 7.5, with the prior architecture's routing in [1]) escalates two or more Amber signals to Restrict. The rule's false-positive rate against benign-but-cross-domain analytical work is the principal operational concern for the framework's adoption. Pilot deployments would publish false-positive distributions, calibration data, and the post-incident threshold updates the prior architecture's mechanism produces. False-positive performance is what determines whether the framework is operationally useful at deployment scale or whether the additional alert surface burdens operators more than the detection coverage justifies.

The five research directions are not exhaustive; they are the directions most directly motivated by the framework's analytical structure. The framework as specified is intended to be a starting point for that work, refined or superseded independently as deployment experience accumulates.

 doi:10.20944/preprints202605.1771.v1

13. Conclusions

This paper has developed a relational systems-theoretic model of discontinuous adversary behavior against Mythos-class AI deployments, a four-class taxonomy of frame-shift discontinuity, a cross-operation detection matrix specifying primary detection mechanisms for each class, and the integration extensions through which the new detection signals route into the prior reference architecture's [1] existing four-layer mitigation stack. The relational model, the taxonomy, the detection matrix, and the mitigation extensions together constitute the present manuscript's contribution; the kinematic vocabulary borrowed in Section 5 is a naming convention only and is replaceable without disturbing the underlying framework.

Three findings anchor the contribution. First, the discontinuity that defines a frame-shift is an operator-observability property rather than a topological or physical one (Section 4). The adversary's actual operation is continuous; the discontinuity is the gap between what the adversary did and what the operator observed. The framing grounds the analysis in systems theory [23,24] and forecloses exotic metaphysical readings while supplying the formal definition from which the four-class taxonomy is derived. Second, the four discontinuity classes—presence, privilege, domain, and observability—partition the relational space at the level of which frame structure carries the primary discontinuity (Section 7). The partition is exhaustive relative to the three-frame model; each class identifies a specific frame or cross-frame structure and a specific constitutive property, and compound classifications across classes are handled through the detection-matrix structure in Section 9 rather than as additional taxonomic categories. Third, cross-operation analysis fills the detection gap that the prior architecture's MCPR runtime tier [1] does not systematically cover (Sections 9 and 10). The four cross-operation detection mechanisms—identity-frame trajectory reconstruction, trust-frame trajectory reconstruction, cross-frame pattern matching, and distributional drift detection—operate on telemetry the prior architecture's instrumented surfaces already produce, and their signals route through the prior architecture's posture-to-mitigation table [1] without restructuring the rubric or introducing parallel architectural primitives.

The operational implication for operators of high-consequence infrastructure is that deployments of the prior reference architecture can be extended with cross-operation analysis using the integration discipline specified in Section 10, drawing on audit substrates the prior architecture already produces, conditional on the resource investments Section 11.2 identifies. The framework specified here is a reference for that extension rather than a deployed system; Section 12 develops the limitations, falsifiability criteria, and research agenda against which subsequent empirical work can refine the design. The relational model, the four-class taxonomy, the cross-operation detection matrix, and the mitigation extensions are intended to be refined or superseded independently as deployment experience accumulates. The reference framework is the paper's contribution; the program of operational deployment, empirical calibration, and component refinement that follows is the research the contribution opens.

Author Contributions: Conceptualization, R.C.; methodology, R.C.; writing—original draft preparation, R.C.; writing—review and editing, R.C. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Campbell, R. Detection and Mitigation of Mythos-Class Frontier Model Capabilities: A Layered Reference Architecture. Computers 2026, 15, 331. https://doi.org/10.3390/computers15060331.

 doi:10.20944/preprints202605.1771.v1

2. Hutchins, E.M.; Cloppert, M.J.; Amin, R.M. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. In Leading Issues in Information Warfare & Security Research; Ryan, J., Ed.; Academic Publishing International: Reading, UK, 2011; Volume 1, pp. 80-106.
3. Strom, B.E.; Applebaum, A.; Miller, D.P.; Nickels, K.C.; Pennington, A.G.; Thomas, C.B. MITRE ATT&CK: Design and Philosophy; MITRE: McLean, VA, USA, 2018.
4. Office of the Director of National Intelligence. Preliminary Assessment: Unidentified Aerial Phenomena; ODNI: Washington, DC, USA, 25 June 2021.
5. National Aeronautics and Space Administration. Unidentified Anomalous Phenomena Independent Study Team Report; NASA: Washington, DC, USA, 14 September 2023.
6. Pols, P. The Unified Kill Chain: Designing a Unified Kill Chain for Analyzing, Comparing and Defending against Cyber Attacks. Master's Thesis, Cyber Security Academy, The Hague, The Netherlands, December 2017.
7. MITRE Corporation. Adversarial Threat Landscape for Artificial Intelligence Systems (ATLAS). Available online: https://atlas.mitre.org (accessed on 24 May 2026).
8. Rose, S.; Borchert, O.; Mitchell, S.; Connelly, S. Zero Trust Architecture. NIST Special Publication 800-207, National Institute of Standards and Technology: Gaithersburg, MD, USA, August 2020. https://doi.org/10.6028/NIST.SP.800-207.
9. Cybersecurity and Infrastructure Security Agency. Zero Trust Maturity Model, Version 2.0; CISA: Washington, DC, USA, April 2023. Available online: https://www.cisa.gov/zero-trust-maturity-model (accessed on 24 May 2026).
10. Executive Office of the President. Executive Order 14028: Improving the Nation's Cybersecurity. Federal Register 86 FR 26633, 12 May 2021.
11. Office of Management and Budget. Memorandum M-22-09: Moving the U.S. Government Toward Zero Trust Cybersecurity Principles; OMB: Washington, DC, USA, 26 January 2022.
12. Sheyner, O.; Haines, J.; Jha, S.; Lippmann, R.; Wing, J.M. Automated Generation and Analysis of Attack Graphs. In Proceedings of the 2002 IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 12-15 May 2002; IEEE: Piscataway, NJ, USA, 2002; pp. 273-284.
13. Ou, X.; Boyer, W.F.; McQueen, M.A. A Scalable Approach to Attack Graph Generation. In Proceedings of the 13th ACM Conference on Computer and Communications Security (CCS '06), Alexandria, VA, USA, 30 October-3 November 2006; ACM: New York, NY, USA, 2006; pp. 336-345.
14. Greshake, K.; Abdelnabi, S.; Mishra, S.; Endres, C.; Holz, T.; Fritz, M. Not What You've Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. In Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security (AISec '23), Copenhagen, Denmark, 30 November 2023; ACM: New York, NY, USA, 2023; pp. 79-90. https://doi.org/10.1145/3605764.3623985.
15. Shavit, Y.; Agarwal, S.; Brundage, M.; Adler, S.; O'Keefe, C.; et al. Practices for Governing Agentic AI Systems. OpenAI: San Francisco, CA, USA, 14 December 2023. Available online: https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf (accessed on 24 May 2026).
16. Chan, A.; Salganik, R.; Markelius, A.; Pang, C.; Rajkumar, N.; et al. Harms from Increasingly Agentic Algorithmic Systems. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), Chicago, IL, USA, 12-15 June 2023; ACM: New York, NY, USA, 2023; 16 pages. https://doi.org/10.1145/3593013.3594033.
17. OWASP Foundation. OWASP Top 10 for Large Language Model Applications, 2025. Available online: https://owasp.org/www-project-top-10-for-large-language-model-applications/ (accessed on 24 May 2026).
18. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1, NIST: Gaithersburg, MD, USA, July 2024.
19. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1, NIST: Gaithersburg, MD, USA, January 2023. https://doi.org/10.6028/NIST.AI.100-1.

 doi:10.20944/preprints202605.1771.v1

20. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.
21. International Organization for Standardization. ISO/IEC 42001:2023 - Information Technology - Artificial Intelligence - Management System; ISO: Geneva, Switzerland, 2023.
22. European Union Agency for Cybersecurity (ENISA). Artificial Intelligence Cybersecurity Challenges - AI Threat Landscape Report; ENISA: Athens, Greece, December 2020.
23. Ashby, W.R. An Introduction to Cybernetics; Chapman & Hall: London, UK, 1956.
24. Luhmann, N. Social Systems; Bednarz, J., Jr., Baecker, D., Trans.; Stanford University Press: Stanford, CA, USA, 1995.
25. Office of Management and Budget. Memorandum M-23-02: Migrating to Post-Quantum Cryptography; OMB: Washington, DC, USA, 18 November 2022.
26. National Institute of Standards and Technology. FIPS 203: ML-KEM (Module-Lattice-Based Key-Encapsulation Mechanism); NIST: Gaithersburg, MD, USA, 13 August 2024. https://doi.org/10.6028/NIST.FIPS.203.
27. National Institute of Standards and Technology. FIPS 204: ML-DSA (Module-Lattice-Based Digital Signature Algorithm); NIST: Gaithersburg, MD, USA, 13 August 2024. https://doi.org/10.6028/NIST.FIPS.204.
28. National Institute of Standards and Technology. FIPS 205: SLH-DSA (Stateless Hash-Based Digital Signature Algorithm); NIST: Gaithersburg, MD, USA, 13 August 2024. https://doi.org/10.6028/NIST.FIPS.205.
29. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29, NIST: Gaithersburg, MD, USA, 26 February 2024. https://doi.org/10.6028/NIST.CSWP.29.