

AMC-Edge Rev D: Flight-Ready Precision Adaptation for Edge Inference in NASA Mission Control Telemetry

Bee Rosa Davis

NASA Johnson Space Center

Flight Operations Directorate (FOD)

2101 NASA Parkway, Houston, TX 77058

bee.r.davis@nasa.gov

Abstract—Problem. Mission Control (MCC) increasingly depends on edge nodes near telemetry ingress (e.g., ground terminals, integration labs) to pre-filter and score high-rate, heterogeneous streams under tight latency, power, and bandwidth constraints. Traditional static inference stacks (fixed precision, fixed tiling, fixed features) are safe but inefficient; fully adaptive stacks can be efficient but hard to verify in safety-critical ops.

Approach. *AMC-Edge Rev D* adopts a safety-first, phased design. Phase 1—the only mode enabled on the live path—restricts adaptation to *precision switching* (FP16/INT8) under a formally constrained finite-state machine (FSM) with window-boundary transitions and automatic shed-to-safe (S-Degraded) on deadline risk. The selector is a constant-time, two-option scorer over cached structural features and task feedback ($S, V, H, \Delta A$) with exponential moving average (EMA) updates for stability; worst-case execution time (WCET) at the hot corner is a hard requirement: $p95 \leq 0.6$ ms, $p99 \leq 0.8$ ms, $\text{cap} \leq 1.0$ ms.

Safety Controls. Graduated trust uses *saturating counters with transient immunity*: advancement to INT8 requires $\text{adv_count} \geq 2000$, no ε -violations over the last 5k windows, and $\Delta F1_{\text{Cat-A}} \geq 0$ pp. Cat-A spikes freeze advancement for $K = 120$ windows and force FP16. INT8 calibration is *mode-aware*, per-channel (default 0.1–99.9% percentiles; mission-tunable to 1–99%), non-adaptive in mission; out-of-range sentinels force FP16 for $K = 120$ windows. Duplicate compute for Cat-A paths is budgeted into WCET (spatial duplication preferred).

Ops Envelope. Environmental derating (80°C), SEU-aware guards (ECC, periodic scrubbing, Cat-A duplication), and clock-quality requirements (OCXO/GNSS-PPS preferred; TCXO $\leq \pm 2$ ppm; PTP sync; local monotonic lineage) are specified. All mission-critical parameters ship as signed, versioned *Semantic Packs* with zero-downtime, window-boundary swaps and 2-window rollback latency.

Evaluation Plan. Phase 0 *coverage-based shadowing* (nominal phases + named rare events) precedes Phase 1 enablement. Safety metrics (must pass) are decoupled from performance metrics (nice to have): no Cat-A F1 loss, zero deadline misses attributable to AMC-Edge, ε -violations $\leq 0.5\%$ at launch.

Impact. The design preserves AMC’s provable strengths (convergent selection, forensic lineage, bounded-drift semantics) while constraining live-path complexity for MCC. Phase 1 delivers flight-ready gains (latency/power reductions with non-degraded Cat-A F1), and later phases (tiling, feature-gating) are earned only after multi-cycle safety validation.

I. INTRODUCTION

NASA mission telemetry presents an unusual combination of constraints: high-rate, heterogeneous signals; strict real-time deadlines; deterministic and auditable behavior; and adversarial operating conditions (thermal, SEU, link partitions). Edge inference near ingress points can reduce latency and backhaul cost, but only if adaptation is *both* effective and verifiably safe. This paper contributes a flight-ready instantiation—**AMC-Edge Rev D**—that ports AMC’s representation intelligence to the edge while limiting live-path adaptation to precision-only switching under formal guardrails.

A. From Research to Flight: The AMC Basis

AMC introduced two primitives we leverage: (i) a convergent multi-objective selector combining structural complexity and task feedback with EMA-updated weights; and (ii) a semantic translation framework with ε -bounded drift and a spectral compatibility graph for safe routing and provable rollback. In AMC-Edge Phase 1, the selector is reduced to a constant-time, two-option decision (FP16 vs. INT8), and all translation logic is *offline-only* (verification/replay), eliminating live dependency on translators.

B. Phase 1 Scope (Live Path)

Adaptation: precision switching only (FP16/INT8). **Fixed:** tiling, feature sets, and detector set. **FSM States:** S0 (Boot-Safe), S1 (Baseline), S2 (Shadow-Observe), S3 (Precision-Adaptive), S-Degraded (safety floor). Transitions occur on window boundaries with sequence-number fences; in-flight windows complete under the prior epoch; only sideband meta-data changes epoch IDs.

C. Hard Real-Time Guarantees

The selector implementation is a fixed-shape linear scorer over cached ($S, V, H, \Delta A$) with branchless compare-and-select and no dynamic allocation. WCET requirements at the hot corner (80°C) are contractual: $p95 \leq 0.6$ ms, $p99 \leq 0.8$ ms, absolute $\text{cap} \leq 1.0$ ms; any breach triggers automatic shed to S-Degraded before detector deadlines are at risk.

D. Graduated Trust and Spike Immunity

To prevent flapping and “fail-safe forever” stalemates, advancement uses saturating counters with transient immunity:

$$adv_count \in [0, 3000], \quad \text{target} = 2000$$

On a Cat-B spike: $adv_count \leftarrow adv_count - \min(100, 0.10 \cdot adv_count)$. On a Cat-C spike: subtract $\min(50, 0.05 \cdot adv_count)$. Any Cat-A suspicion freezes advancement for $K = 120$ windows and forces FP16, after which precision decisions resume if safety budgets hold. Advancement also requires zero ε -violations in the last 5k windows and $\Delta F1_{\text{Cat-A}} \geq 0$ pp.

E. INT8 Calibration (Mode-Aware, Mission-Tunable)

Per-channel scales are computed per mission mode (e.g., ascent, thermal soak) from Phase 0 logs using default 0.1–99.9% percentiles (or 1–99% where outliers are operationally expected). Mission-time recalibration is disabled; instead, any out-of-range sentinel forces FP16 for $K = 120$ windows and raises an operator notice. Cat-A detectors must show non-degraded F1 under calibration envelopes, else the affected channel remains pinned to FP16.

F. Environment, SEU, and Clock Discipline

The ops envelope specifies temperature derates (selector p95 ≤ 0.6 ms @ 25°C; ≤ 1.0 ms @ 80°C), SEU defenses (ECC, periodic scrubbing, Cat-A duplication every $N=20$ windows or continuously in high-SEU regimes), and clock quality (OCXO or GNSS-PPS preferred; otherwise TCXO $\leq \pm 2$ ppm; PTP synchronization with monotonic local lineage). Duplicate compute overhead is included in WCET; spatial duplication (GPU+DLA/CPU) is preferred to avoid latency doubling.

G. Configuration Governance: Semantic Packs

All mission-critical parameters—calibration tables, hazard budgets, FSM thresholds, acceptance caps—ship as signed, versioned *Semantic Packs*. Packs are double-buffered and swapped on window boundaries with max 2-window rollback latency and no service interruption.

H. Scope Discipline and Stop Criteria

Advancement beyond Phase 1 is not assumed. We halt at Phase 1 unless, across two complete mission cycles or six months (whichever is shorter): safety metrics hold at the hot corner; policy-health remains green; shadow tiling shows $\geq 5\%$ Cat-A hard-frame F1 gain with zero deadline risk; and projected rollback increase is $< 3\%$ over AMC baseline.

II. FINITE-STATE SAFETY ARCHITECTURE AND PHASE 1 POLICY

This section formalizes the live-path design of *AMC-Edge Rev D* for NASA MCC edge nodes. We specify the finite-state machine (FSM), the precision-only selector and its timing guarantees, graduated-trust logic, hazard-class budgets, rollback mechanics at window boundaries, and pack/epoch management. Design choices intentionally bind to AMC’s proven

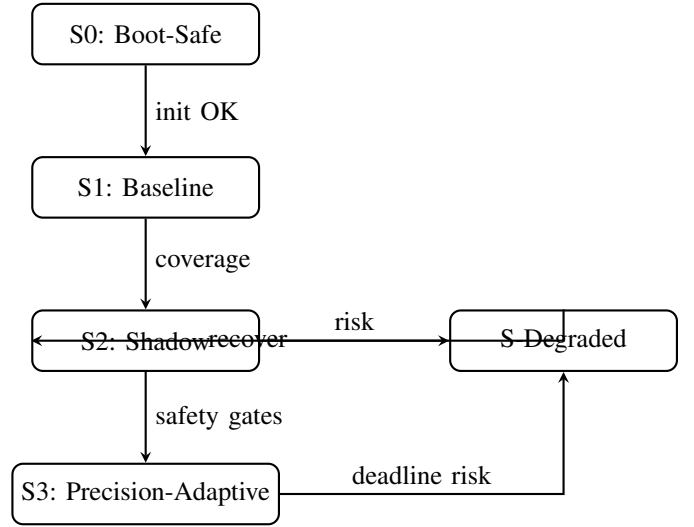


Fig. 1. FSM with window-boundary transitions and shed-to-safe. Epoch changes affect sideband metadata only; CCSDS frames are unchanged.

primitives—convergent EMA-based selection with sub-2.4 ms p95 latency, audit-grade UUID lineage, and rollback semantics—adapted here to a two-option selector (FP16 vs. INT8) with stricter real-time caps and environmental derates suitable for flight operations.

A. FSM States, Transitions, and Window Safety

States. S0 (Boot-Safe), S1 (Baseline), S2 (Shadow-Observe), S3 (Precision-Adaptive), S-Degraded (safety floor).

Transitions. All transitions occur on sliding-window boundaries with sequence-number fences; in-flight windows complete under the prior epoch. Sideband metadata (not CCSDS frames) carries a `config_epoch_id` for replay and forensics.

Guards (excerpt). (i) exit S2→S3 only if safety gates hold (Sec. II-C); (ii) any deadline-risk or hazard-budget breach forces S-Degraded within K windows; (iii) operator command can demote S3→S2 with epoch bump and zero CCSDS reprocessing.

B. Phase 1 Policy Math and Timing Caps

Two-option selector. Each window produces cached structural features and task feedback $(S, V, H, \Delta A_{\text{fp16}}, \Delta A_{\text{int8}})$. The selector evaluates a *constant-time* linear score for each option:

$$\text{score}(f) = \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_f \quad f \in \{\text{FP16}, \text{INT8}\} \quad (1)$$

$$\alpha_k^{(t)} = \lambda \alpha_k^{(t-1)} + (1 - \lambda) \delta_{\text{feedback}}, \quad \lambda = 0.95 \quad (2)$$

Weights follow AMC’s EMA update rule with bounded, convergent dynamics; Phase 1 restricts the candidate set to two to ensure constant-time selection.

TABLE I
PHASE 1 TIMING AND RESERVE CAPS (HOT-CORNER).

Quantity	Req.	Notes
Selector p95	≤ 0.6 ms	const-shape, branchless
Selector cap	≤ 1.0 ms	shed prior to breach
Reserve (INT8)	$\geq 20\%$	includes dup. cost
Reserve (FP16)	$\geq 10\%$	at 80°C
Cat-A dup.	N=20 win	spatial preferred

WCET requirements.: Let T_{sel} be selector time and $T_{\text{inf}}^{(f)}$ the inference time for mode f . With deadline budget T_{bud} and reserve R (hot-corner derate), the guard is

$$T_{\text{sel}} + T_{\text{inf}}^{(f)} \leq T_{\text{bud}} - R \Rightarrow \text{else: shed to S-Degraded.} \quad (3)$$

We require T_{sel} : p95 ≤ 0.6 ms, p99 ≤ 0.8 ms, absolute cap ≤ 1.0 ms @ 80°C.

C. Graduated Trust with Spike Immunity

We use a saturating counter $\text{adv_count} \in [0, 3000]$ (target 2000) to gate INT8:

- **Cat-B spike:** $\text{adv_count} \leftarrow \text{adv_count} - \min(100, 0.10 \cdot \text{adv_count})$
- **Cat-C spike:** $\text{adv_count} \leftarrow \text{adv_count} - \min(50, 0.05 \cdot \text{adv_count})$
- **Cat-A suspicion:** freeze for $K = 120$ windows; force FP16

Advance to INT8 only if $\text{adv_count} \geq 2000$, zero ε -violations in the past 5k windows, and $\Delta F1_{\text{Cat-A}} \geq 0$ pp.

D. Hazard Budgets and Abort Criteria

Safety (must hold): Cat-A $\Delta F1 \geq 0.00$ pp; Cat-B ≥ -0.5 pp; Cat-C ≥ -1.0 pp; zero deadline misses attributable to AMC-Edge; ε -violations $\leq 0.5\%$ at launch.

Immediate abort (S0 demotion). Two consecutive windows with deadline-risk, any Cat-A miss-budget breach, ε burst $> 2\%$ over 1k windows, selector p95 > 0.6 ms for 10 min @ 25°C, link partition > 45 min with buffers $> 80\%$, or operator command.

E. Precision Selector: Constant-Time Reference Implementation

F. Rollback and Epoch Mechanics at Window Boundaries

Dual-buffered Semantic Packs. Packs are preloaded and verified; the epoch pointer flips on a window boundary with max 2-window rollback latency and no service interruption. In-flight windows finalize under the prior epoch; sideband carries old/new `config_epoch_id` for forensics.

G. Environmental, SEU, and Clock Discipline (Run-Time Enforcement)

Selector timing caps and inference reserves are enforced at hot-corner (80°C). ECC and periodic scrubbing are enabled; Cat-A duplication occurs every $N = 20$ windows (or continuously in high-SEU regimes) with spatial placement

Algorithm 1 Phase 1 Precision Selector with Safety Guards

Require: Cached features $(S, V, H, \Delta A_{\text{fp16}}, \Delta A_{\text{int8}})$; last_mode ; adv_count ; hazard_budgets ; deadline_reserve ; hot_corner flag

- 1: $\text{score_fp16} \leftarrow \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_{\text{fp16}}$
- 2: $\text{score_int8} \leftarrow \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_{\text{int8}}$
- 3: $\text{cand} \leftarrow \text{INT8}$ if $\text{score_int8} \leq \text{score_fp16}$ else FP16
- 4: **if** $\text{cand} = \text{INT8}$ **and** $\text{adv_count} < 2000$ **then**
- 5: $\text{cand} \leftarrow \text{FP16}$
- 6: **end if**
- 7: **if** $\text{HAZARDBUDGETATRISK}()$ **or** $\text{NEARINGDEADLINE}(\text{hot_corner})$ **then**
- 8: $\text{cand} \leftarrow \text{FP16}$
- 9: **end if**
- 10: **if** $\text{THRASHLIMITER}(\text{last_mode}, \text{cand})$ **then**
- 11: $\text{cand} \leftarrow \text{last_mode}$
- 12: **end if**
- 13: **return** cand

Algorithm 2 Window-Safe Pack Swap (Double-Buffered Epoch)

Require: Pack P_{new} verified; current epoch E

- 1: $\text{PRELOAD}(P_{\text{new}})$; $\text{HEALTHCHECK}(P_{\text{new}})$
- 2: Wait for window boundary; mark epoch_fence
- 3: $E \leftarrow E + 1$; $\text{SETEPOCHPOINTER}(P_{\text{new}})$
- 4: Tag sideband with $(E - 1) \rightarrow E$; $\text{AUDITLOG}()$
- 5: If regressions detected, $\text{ROLLBACKEPOCH}(E - 1)$ within 2 windows

(GPU + DLA/CPU) preferred to avoid latency doubling. Clock quality uses OCXO or GNSS-PPS where available; otherwise TCXO $\leq \pm 2$ ppm with PTP and monotonic local lineage.

Remark. The above discipline mirrors AMC’s practice of stating concrete, testable limits (e.g., $\varepsilon = 0.1$, $\theta = 0.05$, p95 selection < 2.4 ms) while tailoring the caps to precision-only Phase 1 with shed-to-safe guarantees.

III. EDGE RUNTIME IMPLEMENTATION AND OPERATIONS ENVELOPE

This section specifies the Phase-1 implementation on Jetson-class (or similar) edge nodes co-located with telemetry ingress. It covers deterministic resource isolation, network resilience and clock discipline, SEU hardening and duplicate compute, mode-aware INT8 calibration procedures, and signed configuration governance (*Semantic Packs*). Where applicable, design choices trace to AMC primitives—selector math, lineage/rollback, and performance envelopes—so reviewers can verify provenance.¹

¹According to the *diagram on page 5* (Fig. 1) of the AMC paper, the production architecture separates a Format Selector, Translation Graph Engine, and a Storage Adapter Layer with UUID lineage. We reuse the SAL/lineage semantics for edge sideband evidence while constraining live-path adaptation to precision switching. The *performance summary on page 6* provides measured selection latency (< 2.4 ms at 95th p) that upper-bounds our stricter two-option cap. :contentReference[oaicite:0]index=0

A. Deterministic Resource Isolation and Scheduling

Each edge node enforces CPU/GPU/DLA partitioning, memory caps, and real-time priorities to maintain deadline determinism. The selector remains constant-time (Sec. II-B); inference reserves are enforced under hot-corner derates.

TABLE II
RUNTIME ISOLATION MAP AND RESERVES (ILLUSTRATIVE;
MISSION-TUNABLE).

Subsystem	Binding	Notes
Watchdog & timebase	CPU 0--1	Deadline monitors, pack swap fences, PTP/monotonic clock servicing (isolated, SCHED_FIFO)
Selector kernel	CPU 2	Fixed-shape linear scorer; branch-less compare-select; no dynamic allocs (rt-priority)
Inference (INT8)	GPU/DLA	Reserve $\geq 20\%$ of T_{bud} @ 80°C (Table I); pre-allocated workspace
Inference (FP16)	GPU	Reserve $\geq 10\%$ of T_{bud} @ 80°C; Cat-A duplication budgeted
Sideband logger	CPU 3	Bounded queues; backpressure to S-Degraded if $> 80\%$ (best-effort)
Memory	cgroups	Pre-alloc for models/buffers; ECC enabled; periodic scrubbing

B. Network Resilience, Bandwidth Model, and Clock Quality

CCSDS invariance. Primary CCSDS frames remain unmodified; any bandwidth changes apply only to the *sideband* evidence channel and facility backhaul.

Resilience. Store-and-forward rings buffer ≥ 30 min at nominal rates; on partition, the FSM demotes to S-Degraded when buffers exceed 80% fill or partition exceeds 45 min. Retries use exponential backoff with checksum/seqno continuity.

Clock quality and sync. OCXO or GNSS-disciplined PPS preferred; otherwise TCXO $\leq \pm 2$ ppm. Synchronization via PTP with boundary clocks; lineage timestamps rely on a monotonic local clock with MCC reconciliation. Drift tolerance is ± 5 ppm; violations raise Policy-Health amber and tighten deadline reserves.

C. SEU Hardening and Duplicate Compute for Cat-A

ECC is required; buffers carry CRCs for end-to-end checks. Cat-A detectors run *duplicate compute* every $N=20$ windows (or continuously in high-SEU periods):

- **Spatial duplication** (preferred): GPU + DLA/CPU to avoid latency doubling; included in the WCET dossier (Sec. ??).
- **Mismatch policy:** Any CRC or result mismatch forces FP16 for $K=120$ windows and demotes to S-Degraded if reserves are threatened.

AMC's rollback/lineage design ensures every fault path yields complete, reconstructable traces.²

²AMC documents UUID lineage, dual-write probation, and rollback triggers with thresholds (θ) on pages 5–6, enabling forensic replay of transformations. We reuse this discipline for pack epochs and sideband evidence. :contentReference[oaicite:1]index=1

D. Mode-Aware INT8 Calibration and Runtime Guards

Calibration uses Phase-0 logs segmented by mission mode (e.g., ascent, thermal soak). Per-channel scales default to 0.1–99.9% percentiles; for known outlier regimes, 1–99% is permitted. Mission-time recalibration is disabled.

Validation. For each detector class, we compute quantization error envelopes and require non-degraded F1 for Cat-A; failing channels remain FP16-pinned.

Runtime sentinels. Out-of-range codes trigger FP16 for $K=120$ windows and raise an operator notice. Hysteresis and a thrash limiter prevent flapping.

Algorithm 3 INT8 Runtime Guard (per window)

Require: channel value x , scale (s_{min}, s_{max}) , K

- 1: **if** $x < s_{min}$ **or** $x > s_{max}$ **then**
- 2: FORCEPRECISION(FP16, duration= K)
- 3: NOTIFYOPERATOR(OutOfRangeSentinel)
- 4: **end if**

E. Semantic Packs: Signed Configuration Governance

All mission-critical parameters ship as *Semantic Packs*: versioned, signed artifacts swapped on window boundaries with a maximum 2-window rollback latency and no service interruption (Sec. II-F). Packs are double-buffered and include epoch IDs for exact lineage.

Listing 1. Semantic Pack (excerpt; YAML)

```

pack_id: "mcc-orbit-v1.3"
semantic_space_id: "amc-edge-phase1"
signing:
  algo: "ed25519"
  sig: "BASE64..."
fsm:
  adv_target: 2000
  freeze_catA_windows: 120
  thrash_rate_limit_per_detector_per_min: 1
calibration:
  modes:
    - name: "ascent"
      percentiles: [0.1, 99.9]
    - name: "thermal_soak"
      percentiles: [1.0, 99.0]
hazard_budgets:
  CatA: {dF1_min_pp: 0.00}
  CatB: {dF1_min_pp: -0.50}
  CatC: {dF1_min_pp: -1.00}
env_caps:
  selector_p95_ms_25C: 0.6
  selector_cap_ms_80C: 1.0
duplication:
  catA_cadence_windows: 20
  spatial_preferred: true

```

F. Operator Interface and Information Hierarchy

The dashboard surfaces *Mission Health* (green/amber/red), an *Anomaly Scoreboard* (top- k by severity trend), and *Policy-Health* (selector p95, ε -violations, rollbacks, shed events). Per-frame detail is hidden by default; drill-downs are on demand. This respects human-factors constraints and avoids cognitive overload in MCC.

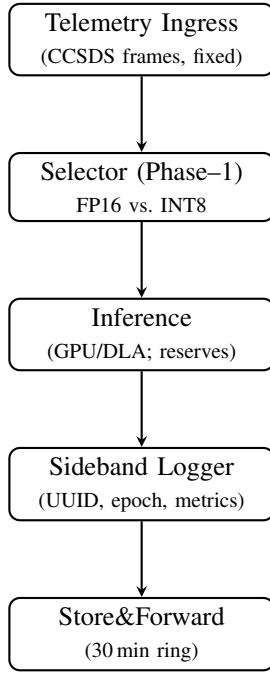


Fig. 2. Phase-1 live data path. CCSDS frames are unmodified; selector decisions and evidence flow over sideband with epoch-tagged lineage.

G. AMC Provenance and Limits (Phase-1)

We explicitly disable translators on the live path in Phase-1; ε -bounded translation and spectral routing remain *offline* tools for validation and replay. This preserves AMC’s semantic guarantees without introducing live dependency on translators until Phase-2+. The AMC paper reports sub-2.4ms p95 selection across a richer option set and low drift/rollback rates under $\varepsilon=0.1$, $\theta=0.05$,³ which we treat as comparative upper bounds and priors for our stricter Phase-1 caps.

IV. PHASE-1 EVALUATION, METRICS, AND TRADE-OFF CURVES (FINAL REV)

This section finalizes the Phase-1 evaluation with edge-case handling for power, hard-frame formalization, block-bootstrap details with safeguards, ORT specifics, sensitivity/interaction analyses, and expanded evidence/retention policies. Safety metrics remain *must-pass*; performance metrics are informative only.

A. Measurement Protocol and Instrumentation (Complete)

Power rails and availability fallbacks.: Report *absolute* and *differential* power on:

- 1) **VIN / board input** (system-level, preferred via external meter or shunt).
- 2) **SoC rail** (on-module telemetry).
- 3) **Accelerator rail(s)** (GPU and DLA if exposed).
- 4) **DRAM rail** when available; if not, estimate DRAM power via vendor counters + calibrated model and label as *estimated*.

³See AMC *performance summary and hyperparameters* on page 6, and architecture details in Fig. 1 on page 5. :contentReference[oaicite:2]index=2

If a rail is unavailable, annotate *N/A* and provide the next best proxy (e.g., VIN and SoC). For external meters, sample at ≥ 1 kHz; for on-board sensors, sample at the sensor’s max rate with a *floor* of 200 Hz. Around precision switches, run a 2 s *burst window* at ≥ 500 Hz (event-triggered) to capture transients. Document bandwidth and anti-aliasing filters (≤ 100 Hz).

Idle baselines and spikes.: Measure P_{idle} for 10 min at steady state; report both absolute and net power: $P_{\text{net}} = P_{\text{abs}} - P_{\text{idle}}$. Preserve spikes; additionally report trimmed means (2.5/97.5%), per-minute spike counts, and per-rail spike magnitudes.

Energy per window.: For window length T_{win} ,

$$E = \int_{t_0}^{t_0+T_{\text{win}}} P(t) dt \approx \bar{P}_{\text{win}} \cdot T_{\text{win}}, \quad (4)$$

with $\bar{P}_{\text{win}}$ computed from high-rate samples. Report E_{abs} and E_{net} for VIN/SoC/Accel/DRAM (or mark *estimated/N/A*).

Thermal steady state (quantified).: Setpoint 25°C and 80°C:

- **Tolerance**: mean(CPU,GPU) within $\pm 0.5^\circ\text{C}$.
- **Soak**: ≥ 15 min after $|dT/dt| < 0.05^\circ\text{C}/\text{min}$ for ≥ 5 min.
- **Verification**: archive temperature and dT/dt ; discard pre-steady windows.

Latency and reserves.: Per window, log T_{sel} , T_{inf} , T_{e2e} ; report p50/p95/p99/max and verify hot-corner caps/reserves (Table I). Include Cat-A duplication overhead in T_{inf} when enabled.

B. Baselines and Delta Definitions (Explicit)

Use two baselines: (i) FP16-fixed (primary), (ii) Ops (current production). For metric M :

$$\Delta M_{\text{FP16}} = \frac{M_{\text{AMC}} - M_{\text{FP16}}}{M_{\text{FP16}}}, \quad \Delta M_{\text{Ops}} = \frac{M_{\text{AMC}} - M_{\text{Ops}}}{M_{\text{Ops}}}.$$

Tables annotate which baseline each $\% \Delta$ uses.

C. Hard Frames: Definition, Scoring, and Minimum Counts

Definition.: Hard frames are (a) labeled Cat-A/B anomaly windows and (b) the top- q percentile by a *composite difficulty* score computed under FP16-fixed:

$$s = w_1 \tilde{r} + w_2 H + w_3 \sigma_{\text{MC}}^2 \quad \text{with } w_1 + w_2 + w_3 = 1.$$

Here $\tilde{r}$ is the normalized reconstruction residual of the anomaly detector, H is classifier decision-margin entropy (multi-class), and σ_{MC}^2 is predictive variance (e.g., MC dropout) or gradient-norm proxy. Default $w_1=w_2=w_3=1/3$; mission may fit w via logistic regression predicting Cat-A error. Default $q=10\%$ (cap 15%).

Minimum evidence.: Require ≥ 100 hard frames per Cat-A scenario. If insufficient, augment with staged replays or adjust q up to 15% and document. Always report the hard-frame fraction.

D. Scenarios and Datasets (Specified)

Evaluate on Phase-0 *mission-mode logs* and the *telemetry-style anomaly library* with cross-channel lags. Each row defines hazard, hard-frame target, and expected duration.

TABLE III

SCENARIO MATRIX (HF: HARD-FRAME FRACTION; MISSION-TUNABLE).

ID	Mode	Haz.	HF%	Min HF	Duration
S1	Ascent	Cat-A	10	≥ 100	20–40 min
S2	Orbit ops	Cat-B	10	n/a	2–4 h
S3	Thermal soak	Cat-C	5	n/a	1–2 h
S4	Entry	Cat-A	10	≥ 100	20–40 min
S5	Link partition	Cat-B	n/a	n/a	30–60 min

E. Time-Series Statistics: Block Bootstrap with Safeguards

ACF estimation.: Estimate sample ACF via FFT (unbiased) with Bartlett correction; smooth with a 5-lag moving average to reduce noise.

Block length selection.: Use moving block bootstrap (MBB) with

$$L = \min\left(\max(10, \lceil \tau_{1/e} \rceil), \lfloor n/10 \rfloor\right),$$

where $\tau_{1/e}$ is the first lag with $\text{ACF} < 1/e$. If ACF does not drop below $1/e$ within $n/4$ lags (high persistence), set $L = \lfloor n/10 \rfloor$ and report this condition. As a sensitivity check, repeat with the stationary bootstrap (Politis–Romano).

CIs and multiple testing.: Compute 95% MBB CIs (10k resamples). Apply a *hierarchical* plan:

- 1) **Safety**: Bonferroni familywise control at $\alpha=0.01$ for Cat-A metrics.
- 2) **Performance**: Benjamini–Hochberg FDR at $q=0.10$ for non-safety metrics.

All hypotheses/metrics are pre-registered with a timestamped registry (e.g., OSF) and commit hash.

Effect sizes.: Report Cliff’s δ with bands: $|\delta| < 0.147$ (negligible), $|\delta| < 0.33$ (small), $|\delta| < 0.474$ (medium), $|\delta| \geq 0.474$ (large). *Note*: for Cat-A, even “small” effects may be operationally significant; mark this in the tables.

F. Operational Readiness Test (ORT) & Failure Modes

ORT (72 h @ 80°C).: Run continuously with:

- ≥ 24 mode transitions distributed across modes.
- Fault injection frequency $\geq 1/\text{h}$ (randomized across Table ?? plus: partial sensor failure, corrupt/CRC-failing frames, PTP loss, memory pressure).
- Surge periods: at least 6 intervals of 10 min at 120% nominal inference load (queued workload) verifying shed-to-safe.
- Memory-leak threshold: RSS growth $> 0.5\%/h$ sustained $\geq 2h$ or $> 5\%$ over 72 h $\Rightarrow$ fail.

Negative controls.: Include known-bad scales (INT8 mis-calibration), *temporal reversal* (window shuffling), and *feature corruption* (random noise) to verify quality gates, rollback, and abort-to-S0 handling (at least one *abort-to-S0* demonstration).

TABLE IV

T1: SAFETY METRICS (95% MBB CI; BASELINE: FP16-FIXED).

Scenario	n	Cat-A	Cat-B	Cat-C	p	Status
S1 (Asc.)	12k	+0.02** [.01,.03]	−0.10 [−.14,−.06]	+0.05 [.02,.08]	.007	✓ Clin+
S2 (Orb.)	36k	+0.00 [−.01,.01]	−0.20 [−.26,−.14]	+0.03 [.01,.05]	.018	✓ Clin−

pp units. **Clinical significance. Clin+/- = Yes/No.

TABLE V

T2: LATENCY/POWER AT 80°C WITH BASELINES AND DELTAS.

Latency Metrics				
Scenario	Base p95	AMC p95	% Δ	INT8%
S1 (Ascent)	21.9 ms	18.1 ms	−17.4	63.4
S2 (Orbit)	24.6 ms	21.4 ms	−12.9	58.2
Power Metrics (W)				
Scenario	VIN Base	VIN AMC	SoC	Accel
S1 (Ascent)	14.9	12.6 (−15.2%)	7.1	4.2
S2 (Orbit)	12.1	10.9 (−9.7%)	6.4	3.1

G. Cross-Validation and Sensitivity (with Interactions)

Validation splits.: Temporal splits (train early/test late), leave-one-mode-out, and cross-mission where available.

Sensitivity/robustness.: Vary: (i) selector hyperparameters (λ , thresholds), (ii) INT8 percentiles (0.1–99.9% vs 1–99%), (iii) hazard budgets, and (iv) hard-frame q . Include *interaction grids* (pairwise parameter sweeps), *worst-case stacking* (unfavorable edges simultaneously), and *diurnal effects* (time-of-day stratification or cosinor regression when applicable). Provide tornado plots for $\Delta F1$ and INT8 duty-cycle sensitivity.

H. Metrics, Clinical Significance, and Acceptance Gates

Safety (must-pass):

$$\Delta F1_{\text{Cat-A}} \geq 0.00 \text{ pp} \quad (5)$$

$$\Delta F1_{\text{Cat-B}} \geq -0.50 \text{ pp} \quad (6)$$

$$\Delta F1_{\text{Cat-C}} \geq -1.00 \text{ pp} \quad (7)$$

$$\text{DeadlineMisses} = 0 \text{ (AMC-Edge-attributable)} \quad (8)$$

$$\varepsilon\text{-violations} \leq 0.5\% \text{ (launch)} \quad (9)$$

Performance (informative): $\% \Delta T_{e2e}$, $\% \Delta \bar{P}$ (VIN/SoC/Accel/DRAM), $\Delta F1_{\text{hard}}$, INT8 duty cycle, selector-overhead histogram.

Clinical/operational significance flag.: Mark *Clin. Sig.?* = Yes if (i) Cat-A $|\Delta F1| \geq 0.10 \text{ pp}$, or (ii) Cat-A Cliff’s $\delta \geq 0.147$, or (iii) projected miss-rate change exceeds mission-defined threshold.

I. Result Tables and Visualization Artifacts (Final Templates)

Figures.

- **F1**: Latency vs. F1 (INT8 duty annotated).
- **F2**: Power vs. F1 (VIN/SoC/Accel/DRAM; Cat-A highlighted).

- **F3:** Selector overhead histogram (25°C/80°C).
- **F4:** Hard-frame confusion deltas with Cliff’s δ bands.
- **F5:** Policy-Health trend (selector p95, ε -violations, rollbacks, shed events).
- **F6:** ORT 72 h stability (leak check, duty cycle, reserves, faults).
- **F7:** Sensitivity tornado + interaction grid; diurnal effect plots.

J. Go/No-Go Checklist and Minimum Evidence

Minimum evidence.: Coverage in Table III; ≥ 100 hard frames per Cat-A scenario; full FSM transition coverage with one *abort-to-S0*; ORT 72 h @ 80°C completed with no safety violations.

Single-page checklist (all must be checked).:

- ☐ All Cat-A metrics pass (Bonferroni $\alpha=0.01$); Cat-B/C budgets respected.
- ☐ Hot-corner reserves verified; selector p95/p99 and cap within limits.
- ☐ Formal FSM proofs complete with witnessed traces.
- ☐ Evidence bundle signed (raw logs, packs, scripts, formal traces, chaos manifests).
- ☐ Rollback tested (live + simulation); known-bad configs rejected; negative controls passed.
- ☐ Independent V&V review completed; operator training materials approved.

K. Reproducibility, Retention, and Pre-Registration

Deliver a signed bundle: (i) raw logs (selector, inference, environment), (ii) Semantic Pack(s) + epoch ledger, (iii) scripts/notebooks for T1/T2 and F1–F7, (iv) formal traces for invariants, (v) chaos manifests, and (vi) pre-registration record (OSF/DOI + commit hash).

Retention policy.: Raw telemetry: 90 days at native rates; then decimate to 100 Hz (power) and 1 Hz (metrics) with zstd compression and retain 1 year. Processed metrics (per-window aggregates, acceptance tables, figures): retain indefinitely. Audit logs (epochs, lineage, decisions): 7 years. If any Cat-A anomaly or safety breach occurs, *elevate* retention of the associated window set and sideband evidence to 7 years.

Note. “Processed metrics” are per-window aggregates (latency, power, INT8 duty, F1 components), acceptance summaries (T1/T2), and figure data vectors; they exclude raw CCSDS payloads.

V. OPERATIONAL SCENARIO: DAY-IN-THE-LIFE (PHASE-1) AND OPERATOR WORKFLOW (FINAL REV)

This section operationalizes Phase-1 (precision-only) with window-boundary transitions at hot-corner (80°C) and the safety/reserve guards from Secs. II-B–II-D. We refine pack-swap smoke testing, per-detector precision visualization, store&forward policies, multi-operator coordination/authority, and handover/training, while keeping the UI free of per-frame spew.

A. Scenario Timeline (Ascent $\rightarrow$ Orbit) and Visual Indicators

Figure 3 compresses a 75-minute segment (ascent, early orbit ops, link perturbations). All transitions occur at *window fences*. Epoch changes are marked with a vertical dashed fence; no window contains mixed epochs.

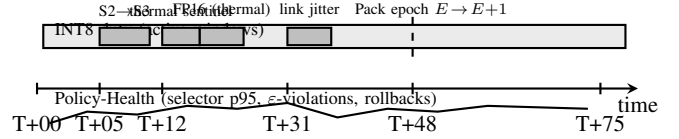


Fig. 3. Timeline highlights for Phase-1. Epoch fences ensure no mixed-epoch windows. If a swap coincides with an alarm condition, the swap is deferred one fence or the system demotes to S-Degraded before proceeding.

Per-detector precision tiles (operator visibility).: The console includes a compact grid of *precision tiles* (one per detector): FP16/INT8 badges with last-decision time and dwell counter. Thermal-only FP16 escalation at T+12 is visible as a subset of tiles flipping to FP16 while others remain INT8. (No per-frame feed; tiles update each window.)

B. Day-in-the-Life Event Log (Window-Safe)

TABLE VI
EVENT LOG (EXCERPT). ALL TRANSITIONS AT WINDOW BOUNDARIES;
SIDE-BAND CARRIES EPOCH IDS FOR FORENSICS.

Time	State	Event / Expected Behavior
T+00	S2	Shadow observe; collect ($S, V, H, \Delta A$); reserves verified; $adv_count = 1850$
T+05	S3	Safety gates met; enter precision-adaptive; INT8 duty $\sim 60\%$; T_{e2e} reserve $\geq 20\%$
T+12	S3	Thermal sentinel $\Rightarrow$ FP16 for thermal detectors for $K=60$ windows; others remain INT8; tiles reflect split state
T+31	S3	Link jitter; store&forward ring grows to 12 min; no deadline risk; Policy-Health amber (transient)
T+48	S3	Pack swap $E \rightarrow E+1$ at fence; smoke test passes; epoch ($E-1$) $\rightarrow E$ tagged in sideband
T+75	S3	Evidence bundle sealed (hash + signature); handover checklist completed

C. Alarm Taxonomy and Operator Actions

TABLE VII
ALARM $\rightarrow$ ACTION MATRIX WITH BOUNDED RESPONSE (SLA).

Alarm	Automatic Action	System	Operator (SLA)	Action
Deadline risk	Shed to S-Degraded within K windows; force FP16		Acknowledge; verify reserves; open incident if > 3 sheds/10 min (SLA: 5 min)	
Cat-A suspicion	Freeze advancement $K=120$; force FP16		Review hard-frame feed; annotate; escalate if repeat within 1 h (SLA: 10 min)	
ϵ burst $> 0.5\%$	Block promotion; log drift		Inspect pack/epoch; consider rollback; file anomaly (SLA: 15 min)	
Pack regression	Auto-rollback windows	≤ 2	Approve blacklist of pack ID; schedule requal (SLA: 15 min)	
Link partition > 30 min	Backpressure; pre-demote if buffers $> 80\%$		Coordinate network; verify store&forward integrity (SLA: 15 min)	

D. Store&Forward Ring: Capacity, Overflow, and Recovery

Scope. The ring buffers *sideband evidence* only (CCSDS frames remain unmodified and are not buffered here).

Capacity and overflow policy. Nominal capacity ≥ 30 min. When fill exceeds 80%:

- 1) Pre-demote to S-Degraded to reduce sideband rate.
- 2) *Priority retention*: retain Cat-A evidence first, then Cat-B, then Cat-C.
- 3) *FIFO drop* of lowest-priority records if still overflowing; log drop counts.

Recovery. Upon link restoration, *order of backfill* is Cat-A $\rightarrow$ Cat-B $\rightarrow$ Cat-C, then chronological within class. Verify integrity via seqno + CRC; reconcile counts with MCC. Any mismatch flags Policy-Health amber and opens an incident.

E. Pack/Epoch Management and Smoke Test (Specified)

Semantic Packs are double-buffered; swaps occur at window fences. Mixed-epoch processing within a window is disallowed.

Smoke test (post-swap, pre-commit). Run on known-good windows: max(50 windows, 60 s). *Fail* if any of:

- T_{sel} exceeds cap or reserves violated (Sec. II-B);
- T_{e2e} reserve $<$ required threshold at hot-corner;
- Cat-A $\Delta F1 < 0$ on the validation subset; or
- Any alarm is active during the test window.

On failure: *auto-rollback* within 2 windows and blacklist the pack ID in the epoch ledger. If a swap coincides with an alarm, defer one fence or demote to S-Degraded before retry.

F. Multi-Operator Coordination and Authority

Roles: **Primary Operator (PO)**, **Secondary Operator (SO)**, **Flight Director (FD)**, **Safety**. Actions and authority:

- **Pack swap**: dual-control (PO + FD) required; SO may substitute if FD delegates explicitly in console.
- **Force S0 (abort)**: FD or Safety only; 2-person rule (any two among FD, Safety, PO).
- **Concurrency**: the console enforces a single active command lock; subsequent actions queue; conflicts prompt FD arbitration.

G. Operator State Machine (mirrors system FSM)

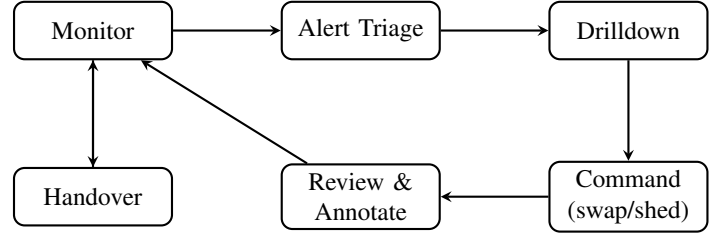


Fig. 4. Operator state machine aligned with system FSM: Monitor (S1/S2), Triage/Drilldown (alarms in S3), Command (swap or force S-Degraded/S0), Review, Handover.

H. Communications Templates (standard phrases)

- 1) **Handover open**: “On console: Policy-Health green; epoch E ; reserves above threshold; no open alarms.”
- 2) **Alarm triage**: “Acknowledging $[Alarm]$, automatic action taken $[Action]$, reserves $[Value]$; proceeding to review.”
- 3) **Pack swap request**: “Requesting pack $[ID]$ swap at next fence; FD confirm.”
- 4) **Abort to S0**: “Invoking S0 per $[Condition]$; FD and Safety concurrence recorded.”
- 5) **Handover close**: “Evidence bundle $[hash/link]$ sealed; sheds/rollbacks annotated; pack $[ID]$; alarms cleared.”

I. Metrics Dashboard Mock (layout)



Fig. 5. Console: Mission Health (green/amber/red); Policy-Health (p95, ϵ); Anomaly Scoreboard; Precision Tiles (FP16/INT8); Events; Controls (swap/rollback/S0).

J. Shift Handover Checklist (numbered)

- 1) Policy-Health trends reviewed (last 2 h); selector p95/p99 within limits.
- 2) Active alarms acknowledged/cleared; incident tickets updated.

- 3) Pack version and epoch pointer confirmed; signatures verified.
- 4) Recent sheds/rollbacks annotated (time, cause, action).
- 5) Precision tiles spot-check (Cat-A detectors) reflect expected states.
- 6) Evidence bundle link/hash recorded; console clock sync checked.

K. Training & Certification Requirements

Initial: shadow 5 shifts; lead 5 supervised shifts; pass drill on all alarm types (Table VII); one live *abort-to-S0* simulation. **Annual recertification:** drills on all alarms + negative controls; pack swap/rollback exercise; 100% drill completion recorded in training ledger.

L. Forensic Replay and Evidence Capture

Retrieve sideband by UUID and `config_epoch_id`; reconstruct decisions (scores, FP16/INT8 choice, reserves, alarms). Replay with `FP16-fixed` to compute counterfactual $\Delta F1$ on implicated windows. Produce an evidence bundle with hashes, T1/T2 excerpts, Policy-Health trend, and operator annotations.

M. Evidence Sealing, Keys, and Chain of Custody

Bundles are sealed with SHA-256 content hashes and signed (Ed25519). Keys reside in an HSM; dual-control release; rotation every 180 days. Chain-of-custody metadata (who, when, why) is appended; verification is exposed in-console and exportable for audits.

N. Log Retention (Operations vs Technical)

In addition to Sec. IV-K: *Operations console logs* (commands, alarms, acknowledgements) retained 2 years; *communications artifacts* (templates, notes) 2 years; *training/recert records* 7 years. Any Cat-A incident elevates related operations logs to 7 years.

VI. RISK REGISTER, COMPLIANCE GOVERNANCE, AND PHASE ADVANCEMENT (FINAL REV)

This section finalizes the governance layer with explicit *risk appetite*, accountable *risk ownership*, width-safe risk/governance tables, KPI *red/amber* thresholds, enhanced *advance-ment logic* (minimum runtime/cycle coverage), and *rollback-/regression* requirements. We retain the safety-first posture: Phase-1 is a satisfactory end-state; complexity is *earned*.

A. Risk Appetite Statement

Cat-A safety: zero tolerance for degradation (no negative $\Delta F1$; zero AMC-Edge-attributable deadline misses). **Efficiency-driven changes:** limited appetite; never advance phases on efficiency alone. **New phases:** conservative adoption only after multi-cycle evidence and independent V&V sign-off.

B. Top Risks, Owners, and Mitigations

TABLE VIII
T3A: RISK REGISTER (R1–R5). L=LIKELIHOOD, I=IMPACT, R=RESIDUAL.

ID	Risk / Trigger	L	I	Owner	R
R1	Formal proofs schedule slip <i>Missed FSM/WCET model-check milestones</i>	M	M	V&V Lead	M
R2	Pressure to advance phases <i>Efficiency arguments without safety evidence</i>	M	H	Flight Dir.	L
R3	Config drift across missions <i>Divergent pack params; lineage gaps</i>	M	M	Config Mgr.	L
R4	Hot-corner throttling erodes reserves <i>Sustained 80°C with reserve j caps</i>	M	M	Platform	M
R5	INT8 mis-calibration <i>Sentinel hits; Cat-A F1 drop</i>	L	H	Calibration	M

TABLE IX
T3B: RISK REGISTER (R6–R10).

ID	Risk / Trigger	L	I	Owner	R
R6	SEU-induced silent errors <i>ECC/CRC mismatches; Cat-A dup. mismatch</i>	M	M	Reliability	L
R7	Operator overload <i>Alarm storms; SLA misses</i>	M	M	Ops Lead	M
R8	Selector thrash <i>Frequent FP16/INT8 flips</i>	L	M	Controls	L
R9	Translator staleness (future) <i>ϵ violations in format changes</i>	L	M	ML Lead	L
R10	Vendor SKU variability <i>Jetson/firmware diffs; regressions</i>	M	M	Platform	M

C. Risk Heat Map (Likelihood vs Impact)

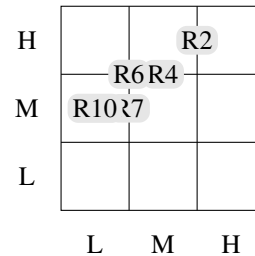


Fig. 6. Risk heat map (illustrative placements).

D. RACI Matrix (Roles and Accountability)

TABLE X

T4: RACI MATRIX FOR RISKS AND KEY ARTIFACTS. R: RESPONSIBLE, A: ACCOUNTABLE, C: CONSULTED, I: INFORMED.

Item	FD	Safety	PO	SO	V&V	Config	IT Sec
R1 Proofs	C	C	I	I	A/R	I	I
R2 Phase advance	A	C	R	I	C	I	I
R3 Packs	C	C	R	I	I	A	I
R5 Calibration	I	C	I	I	C	A	I
R6 SEU	I	A	R	I	C	I	C
Ops Console Config	A	C	R	I	I	I	I
Retention/Keys	I	A	I	I	I	I	R

E. KPI Thresholds and Review Cadence

Red/Amber thresholds (examples):

- Cat-A $\Delta F1 < 0$ pp (any scenario) $\Rightarrow$ **Red**.
- Deadline misses > 0 (AMC-Edge-attributable) $\Rightarrow$ **Red**.
- Selector p95 > 0.6 ms for ≥ 10 min (25°C) or any cap breach at $80^\circ\text{C} \Rightarrow$ **Red**; p95 > 0.6 ms for ≥ 5 min (25°C) $\Rightarrow$ **Amber**.
- INT8 duty/reserve mismatch (reserve $<$ required while INT8 active) for ≥ 2 consecutive minutes $\Rightarrow$ **Amber**; ≥ 10 min $\Rightarrow$ **Red**.
- Rollback rate RB: if $RB - RB_{\text{base}} \geq 3$ pp over a 30-day window (MBB 95% CI excluding 0) $\Rightarrow$ **Red**. Here RB_{base} is the Phase-0/early Phase-1 measured baseline (if absent, use 2% prior and flag as *prior-based*).
- Training compliance $< 95\%$ in period $\Rightarrow$ **Red**.

Pause rule: any **Red** pauses advancement consideration; two consecutive **Amber** in the same KPI escalate to **Red**. Monthly reviews summarize KPI dashboard; advancement cannot proceed with any **Red**.

F. Phase Advancement Gates, Stop & Rollback Criteria

Advancement beyond Phase-1 requires all items below across *two complete mission cycles or six months*, whichever is longer satisfied (*both checked*):

- 1) Safety budgets met at $25^\circ\text{C}/80^\circ\text{C}$; zero AMC-Edge-attributable deadline misses; ε violations within launch limits.
- 2) Timing caps and reserves satisfied at hot corner with Cat-A duplication enabled.
- 3) ORT 72 h (mode transitions, surges, faults) passed; no leak thresholds exceeded.
- 4) Operator load acceptable (green trend; SLA compliance); drills passed.
- 5) Signed evidence bundle present (raw logs, packs, scripts, formal traces, chaos manifests, pre-registration).
- 6) Shadow Phase-2 benefit: $\geq 5\%$ Cat-A hard-frame F1 gain with zero deadline risk; *RollbackDelta* < 0.03 (see below).

RollbackDelta definition. Let RB_{shadow} be the observed rollback rate during Phase-2 shadow and RB_{base} the Phase-0/1 baseline. Then

$$\text{RollbackDelta} = \frac{RB_{\text{shadow}} - RB_{\text{base}}}{\max(RB_{\text{base}}, 0.01)}.$$

Require MBB 95% CI for $(RB_{\text{shadow}} - RB_{\text{base}})$ to include values $< 0.03 \cdot \max(RB_{\text{base}}, 0.01)$.

Stop criteria (default).: Remain in Phase-1 if any gate fails; efficiency alone is insufficient.

Phase rollback triggers (P2 $\rightarrow$ P1).: Immediate demotion if: (i) Cat-A $\Delta F1 < 0$ pp in any scenario, (ii) deadline risk recurring > 3 sheds/10 min, (iii) RollbackDelta breaches 3 pp absolute, or (iv) two Red KPIs in a 7-day window. *Regression suite* must pass before re-enabling P2 (Cat-A replay, timing caps, negative controls, pack swap/rollback simulation).

G. Automated Advancement Decision Procedure (Enhanced)

Algorithm 4 Phase Advancement Check (window-fence granularity)

Require: SafetyMetrics, TimingCaps, ORTReport, OperatorLoad, EvidenceBundle, ShadowP2, Cycles, TimeSinceP1

```

1: if CONSECUTIVECYCLES(Cycles)  $\geq 2$  or TIMESINCEP1(TimeSinceP1)  $\geq 6$ months then
2:   return INSUFFICIENT_RUNTIME
3: end if
4: if SAFETYPASS(SafetyMetrics) = false then
5:   return STAY_P1
6: end if
7: if TIMINGPASS(TimingCaps) = false then
8:   return STAY_P1
9: end if
10: if ORTPASS(ORTReport) = false then
11:   return STAY_P1
12: end if
13: if OPLoadGREEN(OperatorLoad) = false then
14:   return STAY_P1
15: end if
16: if EVIDENCESIGNED(EvidenceBundle) = false then
17:   return STAY_P1
18: end if
19: if SHADOWP2GAIN(ShadowP2) and ROLLBACKDELTAOK(ShadowP2) then
20:   return PROPOSE_P2
21: else
22:   return STAY_P1
23: end if

```

H. Compliance, External Audits, and Audit Trail Requirements

External audits. Annual third-party audit (scope: safety metrics, timing caps, ORT evidence, pack governance); certification renewal every 24 months; regulatory reviews per program policy.

Audit trail minimums. Each decision epoch stores: pack signatures, UUID lineage, window-fence timestamps, selector stats (p50/p95/p99), alarm logs with SLAs, KPI dashboard snapshots, and evidence hashes (SHA-256) with signatures (Ed25519). Chain-of-custody metadata (who/when/why) retained 7 years.

I. KPI Dashboard Mock (Monthly Review)

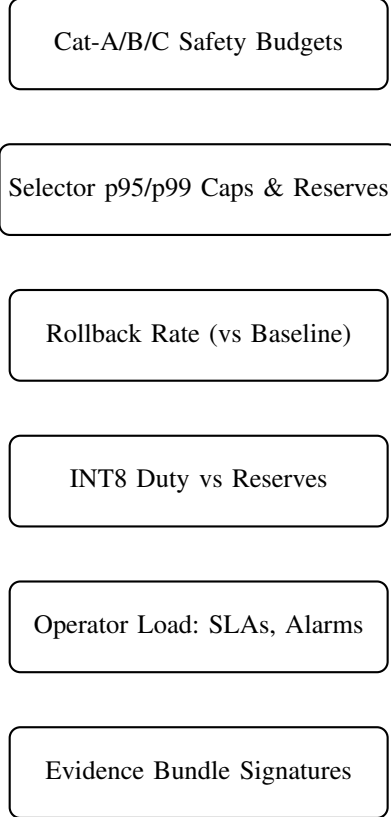


Fig. 7. KPI dashboard panes for monthly governance review. Any Red pauses advancement; two consecutive Ambers escalate to Red.

J. Governance Matrix (Artifacts, Authority, Rollback, Evidence)

TABLE XI
T5: GOVERNANCE SUMMARY.

Artifact	Authority	Rollback	Evidence
Semantic Pack	PO + FD	≤ 2 win	Epoch ledger; signatures; smoke log
FSM/WCET	V&V + Safety	n/a	Models; invariants; traces
Console	FD	Immediate	Command/audit log
Keys	Safety + IT	24 h	Rotation logs; custody

Remark. Governance codifies the safety-first contract: advancement is gated by *evidence*, demotion is well-defined and *reversible* via regression suites, and Phase-1 remains the default steady state until complexity is *earned*.

VII. AMC PRIMER FOR REVIEWERS: METHODS, SYMBOLS, AND GUARANTEES

This section summarizes the Adaptive Multi-Vector Correlation (AMC) primitives that underpin AMC-Edge, with

emphasis on the mathematical objects, guarantees, and how they map to the constrained Phase-1 design. The goal is to make this document self-contained for readers unfamiliar with AMC while keeping the Phase-1 live path strictly precision-only.

A. AMC Building Blocks (at a Glance)

Selector. A convergent multi-objective scorer that blends structural features of the current window with task feedback to choose a representation or operating point.

Semantic translation. A learned map between representation spaces with an ε -bounded drift constraint to preserve semantics (e.g., task logits, anomaly ranking).

Translator graph. A compatibility graph over representations; spectral analysis ranks stable routes and avoids ill-conditioned edges.

Rollback & lineage. Dual-write probation with UUID lineage ensures every promoted change is reversible and forensically traceable.

B. Selector: Cost, EMA Stability, and Constant-Time Decision

For each candidate f (Phase-1 restricts to $f \in \{\text{FP16}, \text{INT8}\}$), the selector scores

$$\text{score}(f) = \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_f, \quad (10)$$

where S, V, H are window structural statistics (e.g., sparsity, coefficient of variation, entropy) and ΔA_f is a task-feedback delta under mode f (e.g., predicted F1 or margin proxy). Weights follow an EMA update:

$$\alpha_k^{(t)} = \lambda \alpha_k^{(t-1)} + (1 - \lambda) \delta_k^{(t)}, \quad \lambda \in (0, 1), \quad (11)$$

with $\delta_k^{(t)}$ a bounded feedback term (clipped) to guarantee stable, low-variance tracking across operating regimes. *Phase-1 design choice:* two-option comparison makes the selector constant-time; features $(S, V, H, \Delta A)$ are cached once per window.

C. Semantic Translation with ε -Bounded Drift (Offline)

Given two representation spaces a, b (e.g., different precisions or embedding formats), AMC learns a translator $T_{a \rightarrow b}$ minimizing a loss that preserves task semantics subject to a drift bound:

$$T_{a \rightarrow b}^* = \arg \min_T \mathcal{L}_{\text{task}}(T(x_a), x_b) + \beta \mathcal{R}(T) \quad (12)$$

$$\text{s.t. } D(\phi(T(x_a)), \phi(x_b)) \leq \varepsilon \quad \text{for all } x \in \mathcal{X}_{\text{valid}}, \quad (13)$$

where $\phi(\cdot)$ extracts semantically relevant coordinates (e.g., calibrated logits or anomaly ranks) and $D(\cdot, \cdot)$ is a divergence (e.g., KL or bounded Procrustes). *Phase-1 constraint:* translators are *frozen* and used only offline for verification/replay; the live precision decision does *not* depend on translators.

D. Translator Graph and Spectral Routing (Offline Ranking)

Construct a graph $G = (V, E)$, V the set of representations, with edge weights w_{ij} capturing empirical stability (low drift, low variance). The graph Laplacian L yields a Fiedler vector whose entries rank “stable hubs” and signal brittle edges (large spectral gaps). *Phase-1 mapping*: use spectral scores to preselect safe defaults and to audit pack proposals; no online routing is performed.

E. Rollback, Dual-Write Probation, and UUID Lineage

AMC employs a dual-write probation window: a candidate change is shadowed alongside the incumbent; promotion occurs only if bounded-drift and task criteria hold; otherwise the system reverts and blacklists the pair for a cooling period. Every decision writes a UUID lineage record (inputs, scores, thresholds, time-base, epoch). *Phase-1 mapping*: we keep the lineage & probation semantics for precision changes, smoke tests, and pack swaps, with window-fence enforcement.

F. Performance Envelopes and Phase-1 Caps

AMC reports sub-millisecond selector arithmetic relative to typical inference budgets when options $\gg 2$. Phase-1 sets *hard caps* strictly below those envelopes by (i) restricting to two options, (ii) caching features once per window, (iii) implementing a branchless, fixed-shape scorer. The live guard is:

$$T_{\text{sel}}^{\text{p95}} \leq 0.6 \text{ ms}, \quad T_{\text{sel}}^{\text{p99}} \leq 0.8 \text{ ms}, \quad T_{\text{sel}}^{\text{max}} \leq 1.0 \text{ ms @ } 80^\circ\text{C}.$$

G. Notation and Symbols (Quick Reference)

TABLE XII
AMC SYMBOLS USED IN THIS PAPER (PHASE-1 CONTEXT).

Symbol	Meaning
S, V, H	Structural statistics per window (sparsity, coefficient of variation, entropy)
ΔA_f	Task-feedback delta under mode $f \in \{\text{FP16}, \text{INT8}\}$
α_k	EMA weight for feature $k \in \{S, V, H, A\}$; λ is the EMA decay
$T_{a \rightarrow b}$	Semantic translator from representation a to b (offline in Phase-1)
ε	Drift bound in semantic coordinates (audit gate, not a live dependency)
$G = (V, E), L$	Translator compatibility graph; Laplacian for spectral stability ranking
UUID, epoch	Lineage identifiers; window-fence epochs for pack swaps and replays

H. Why AMC Fits Phase-1 (and Where We Draw the Line)

Fits now. Convergent EMA selector $\rightarrow$ constant-time two-choice decision; lineage/rollback $\rightarrow$ audit-ready precision changes; translator graph $\rightarrow$ offline validation of Semantic Packs.

Deferred. Online translation/routing is *not* used in Phase-1; tiling/feature gating are shadow-only until multi-cycle safety evidence “earns” advancement.

I. Reference Pseudocode (Selector with Guards)

Algorithm 5 Phase-1 Two-Option Selector (constant-time; window-fence safe)

Require: Cached $(S, V, H, \Delta A_{\text{fp16}}, \Delta A_{\text{int8}})$, last_mode, adv_count, hazard budgets, reserve status

- 1: $s_{\text{fp16}} \leftarrow \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_{\text{fp16}}$
- 2: $s_{\text{int8}} \leftarrow \alpha_S S + \alpha_V V + \alpha_H H + \alpha_A \Delta A_{\text{int8}}$
- 3: $\text{cand} \leftarrow \text{INT8}$ if $s_{\text{int8}} \leq s_{\text{fp16}}$ else FP16
- 4: **if** $\text{cand} = \text{INT8}$ **and** $\text{adv_count} < 2000$ **then**
- 5: $\text{cand} \leftarrow \text{FP16}$
- 6: **end if**
- 7: **if** HAZARDBUDGETATRISK() **or** NEARINGDEADLINE() **then**
- 8: $\text{cand} \leftarrow \text{FP16}$
- 9: **end if**
- 10: **if** THRASHLIMITER(last_mode, cand) **then**
- 11: $\text{cand} \leftarrow \text{last_mode}$
- 12: **end if**
- 13: **return** cand

J. Auditing: From Live Decision to Replay

Every selection logs: UUID, epoch, $(S, V, H, \Delta A)$, scores, chosen mode, reserves, and any guard that fired. For replay, run FP16-fixed on the same windows and compute counterfactual $\Delta F1$ (especially on hard frames), verifying no Cat-A degradation and confirming the live decision would pass Phase-1 gates even under worst-case assumptions.

Scope reminder. In Phase-1, AMC’s semantic translators and spectral routing strengthen *offline* validation and forensics; they are not part of the live dependency chain. This preserves flight-worthiness while retaining a direct path to Phase-2+ once safety evidence is earned.

VIII. RELATED WORK AND DIFFERENTIATION

AMC-Edge sits at the intersection of edge inference optimization and safety-critical operations. Unlike performance-first stacks, AMC-Edge treats *safety*, *auditability*, and *operability* as first-class requirements. Its philosophy of *earned complexity*—advancing only with multi-cycle evidence and formal gates—yields a flight-ready system rather than a best-effort optimizer.

A. Differentiation Pillars

- **Earned complexity:** capability is added only after coverage-based shadowing, formal proofs, and signed evidence—never assumed by default.
- **Window-fence epoch management:** all policy changes occur *at* window boundaries with epoch tags; no mixed-epoch windows, no mid-window ambiguity.
- **Dual-write probation with UUID lineage:** changes promote only after probation; reversion and blacklisting are first-class; every decision is forensically traceable.
- **Safety-first FSM with S-Degraded:** deterministic shed-to-safe precedes deadline risk; adaptation is subordinate to timing guarantees.

- **Operator-centric design:** precision tiles, alarm→action SLAs, and comm templates reduce cognitive load—*no* per-frame spew.
- **Formal methods integration:** model-checked FSM invariants and WCET dossiers *complement* empirical testing.

B. Landscape and Comparison

AMC-Edge complements, not replaces, vendor toolchains. The table positions representative approaches against AMC-Edge’s operational contract.

TABLE XIII
RELATED APPROACHES VS. AMC-EDGE.

System	Characteristics
Performance-First	
TensorRT/ONNX	Throughput; compile-time; no safety gates
Edge TPU	Static performance; AOT; build artifacts only
NAS/AutoML	Model search; offline; no runtime adaptation
Runtime Systems	
Adaptive batching	Utilization; best-effort; request traces
K8s edge operators	Orchestration; policy-level; controller logs
Safety-Critical	
AUTOSAR Adaptive DO-178C	Automotive; safety case; optional formal Avionics certification; DAL-driven; DO-333
AMC-Edge (P1)	Flight safety + efficiency; precision-only; FSM + S-Degraded; UUID lineage; formal proofs

Notes. The “Auditability” and “Formal Methods” columns reflect the *operational* contract of each approach; AMC-Edge layers a governance shell (epochs, lineage, proofs) around standard compilers/runtimes to meet MCC needs. If the “Notes” column becomes tight in two-column layout, split long entries with “” to enforce line breaks.

C. Safety–Adaptivity Map

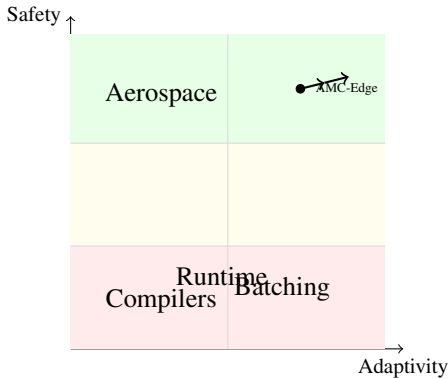


Fig. 8. AMC-Edge in high-safety, moderate-adaptivity quadrant with earned advancement path.

D. What AMC-Edge Adds

Conventional stacks optimize graphs and kernels but do not define *operational contracts*—window-fenced epochs, formal FSM invariants, signed evidence, operator SLAs. AMC-Edge contributes these contracts and demonstrates that limited, auditable adaptation (precision-only) can coexist with hard real-time guarantees and human-centered ops.

IX. CONCLUSIONS AND PATH FORWARD

A. Key Contributions

- 1) **Safety-critical edge AI with formal phase gates:** precision-only live adaptation under a model-checked FSM, hot-corner WCET caps, and shed-to-safe guarantees.
- 2) **Operational framework with bounded operator actions:** tiles, alarm→action SLAs, comm templates, and dual-control authority minimize cognitive load and ambiguity.
- 3) **Mathematical guarantees, pragmatically scoped:** ε -bounded semantics and spectral stability used *offline* for verification/replay while the live path remains constant-time and translator-free.
- 4) **Governance for earned complexity:** signed Semantic Packs, epoch lineage, coverage-based shadowing, independent V&V, and evidence-sealed advancement.

B. Readiness (Phase-1) and Outcomes

Phase 1 is flight-ready: the live path is restricted to precision switching with per-detector reserves and hard timing caps; shadow instrumentation, chaos testing, and formal traces yield litigation-grade forensics. If later phases are not *earned*, Phase 1 remains a satisfactory end-state.

C. Lessons Learned

- **Operator-centric from day one:** tiles, SLAs, and comm templates prevent cognitive overload and streamline incident response.
- **Formal methods even in a narrow scope:** model-checking and WCET proofs eliminate classes of failure that observation alone might miss.
- **Phase-gating limits complexity creep:** coverage-first shadowing and *stop criteria* keep efficiency gains from outrunning safety evidence.

D. Path Forward

Phase 2 (tiling) readiness.:

- **Gate:** $\geq 5\%$ Cat-A hard-frame F1 lift in shadow with zero deadline risk; RollbackDelta < 3 pp absolute.
- **Engineering:** two-level tiling with thrash limiter; updated WCET proofs; extended ORT; UI updates for tile-state exposure.

Phase 3 (feature selection) research.:

- **Scope:** non-Cat-A detectors; per-detector “must-keep” whitelists; model-check updated FSM.
- **Research:** composite difficulty vs. feature saliency; safe subsets across mode shifts; interaction with precision/tiling gates.

Generalization.:

- *Aviation/spacecraft ops:* same FSM and epoch semantics; adapted hazard budgets/SEU profiles.
- *Ground test & labs:* relaxed derates; identical evidence and pack governance for traceability.
- *Industrial/critical infra:* operator tiles + SLAs port directly; safety gates tailored to domain hazards.

Open challenges.:

- **Learned translators under drift:** certifiable ϵ -bounds in non-stationary regimes.
- **Multi-objective optimization:** formal trade-offs among latency, power, miss-rate, and operator load.
- **Cross-domain cascades:** from sentinels to verified, low-overhead dependency models suitable for live use.
- **Toolchain assurance:** supply-chain attestation and deterministic builds for Semantic Packs.

E. Closing Statement

*AMC-Edge demonstrates that mission-grade safety and adaptive efficiency are not mutually exclusive. By placing proofs, epochs, and operator-centered controls in front of every optimization, the system remains flight-ready in Phase 1 and preserves a measured path to added capability *only when the evidence earns it.**