

SECTOR ANALYSIS • GOVERNMENT, DEFENSE & AEROSPACE

Inside the Boundary

Commercial SaaS observability cannot follow classified data across the boundary. The result is that the estates with the highest consequence of failure frequently have the least visibility into it.

THALAMUS ADVISORY • ENTERPRISE RESILIENCE PRACTICE

01 — THE SIGNAL

A gap created by classification

There is a pattern in defense and intelligence estates that surprises people arriving from commercial engineering: the systems with the highest consequence of failure often have the *least* observability. Not because the organisations do not want it. Because the tooling could not follow the data.

US defense spending is around \$850 billion, inside roughly \$6.8 trillion in total federal outlays, with about \$1.5 trillion in federal contracting and state and local technology spending combined. Almost every leading observability platform is SaaS, and SaaS requires egress. Inside a classified enclave, egress is the thing that is not permitted.

02 — THE MECHANISM

Public accountability, constrained tooling

Two forces shape reliability practice here, and they pull in opposite directions.

The first is accountability. An outage becomes a congressional hearing, an inspector-general report, or a story about citizens unable to access benefits. Failure is political, and the record of what happened is scrutinised by people with subpoena power.

The second is constraint. Classified enclaves cannot use tooling that phones home. Systems integrators operate estates they did not design, under contract, with limited authority to change architecture. Fielded systems — tactical communications, deployed platforms — produce telemetry that arrives late, in batches, or not at all.

The organisations with the most at stake are frequently the least instrumented, and the reason is architectural rather than cultural.

Citizen-facing systems add a third pattern: load is driven by policy deadlines — enrolment windows, filing dates — which are known years ahead and absolutely immovable, and which fail publicly when they fail.

03 — THE COST

Where it actually lands

SURFACE	HOW IT PRESENTS
Public failure	Citizen-facing outages affect people with no alternative and become political events within hours.
Mission readiness	Sustainment system availability affects platform readiness directly, and the metric is reported upward.
Contractual SLAs	Integrators operate agency systems under contract with penalties, on architectures they cannot unilaterally change.
Late telemetry	Fielded systems report hours after the event, so any diagnostic tooling anchored to “recent” is structurally unable to investigate them.
Programme timescales	Multi-year programmes mean data integrity over years matters more than uptime in an hour, and silent corruption is the catastrophic case.

04 — THE AI STAKES

Adoption pressure against an evidence requirement

Federal AI adoption is accelerating under explicit policy direction, and simultaneously under explicit oversight requirements. That combination is unusual: the same organisations are being told to adopt faster and to explain more.

For operational AI the binding constraint is explainability. An automated action against a mission or benefits system must be reconstructable afterwards with the evidence available at the time — not summarised, but reproducible. Most implementations discard the probes once a conclusion is reached, which makes the reconstruction impossible precisely when it is demanded.

The adversary side is straightforward and well documented: state-aligned actors target this sector continuously, dwell times are long, and the automation of reconnaissance has made patient campaigns cheaper. An estate that cannot investigate a window that closed three weeks ago cannot answer the question that matters after a discovery.

05 — THE REMEDY

Bring the tooling inside

- **Self-hosted, with no egress.** This is an eligibility requirement rather than a preference. Tooling that cannot run inside the boundary cannot be used at all.
- **Anchor investigation to event onset** and retain error telemetry long enough to cover realistic discovery latency. Fielded systems report late by nature.
- **Treat escalation as the default for mission systems.** Guardrails that prevent automated remediation, enforced at the execution boundary rather than requested of a model.
- **Retain evidence, not conclusions.** Probes, results, confidence scores and the full agent-to-agent chain, reconstructable for audit.
- **Instrument without architectural change** where the estate belongs to someone else — OTLP collection observes without demanding the redesign a contract does not permit.
- **Pre-stage for policy deadlines.** Enrolment and filing peaks are known years ahead; configuration derived from expected load beats configuration set in a quiet quarter.

06 — IN PRACTICE

Hosted, or customized for your estate

Hosted on Thalamus AI Cloud

Appropriate for unclassified work: contractor corporate IT, proposal and programme management systems, commercial subsidiaries, and citizen-facing services that hold no classified data. Managed instances with dedicated tenancy remove the infrastructure lead time, which matters where standing up an internal platform is itself a programme with its own approval cycle.

Customized for your enterprise

For everything else — and in this sector that is most things — the platform deploys self-hosted inside your boundary, including air-gapped and classified enclaves, with no egress to any vendor service. This is the deployment model the architecture was designed around rather than an accommodation bolted on afterwards.

Customization typically covers: deployment inside the accreditation boundary with the artefacts your ATO process requires; retention tiering that keeps error telemetry long enough for realistic discovery latency on fielded systems; guardrail policy that makes mission-system remediation impossible rather than merely discouraged; per-programme and per-agency tenancy so each contract is measured against its own SLAs; and audit logging of every agent action and hand-off, retained for the period your oversight regime specifies rather than a vendor default.

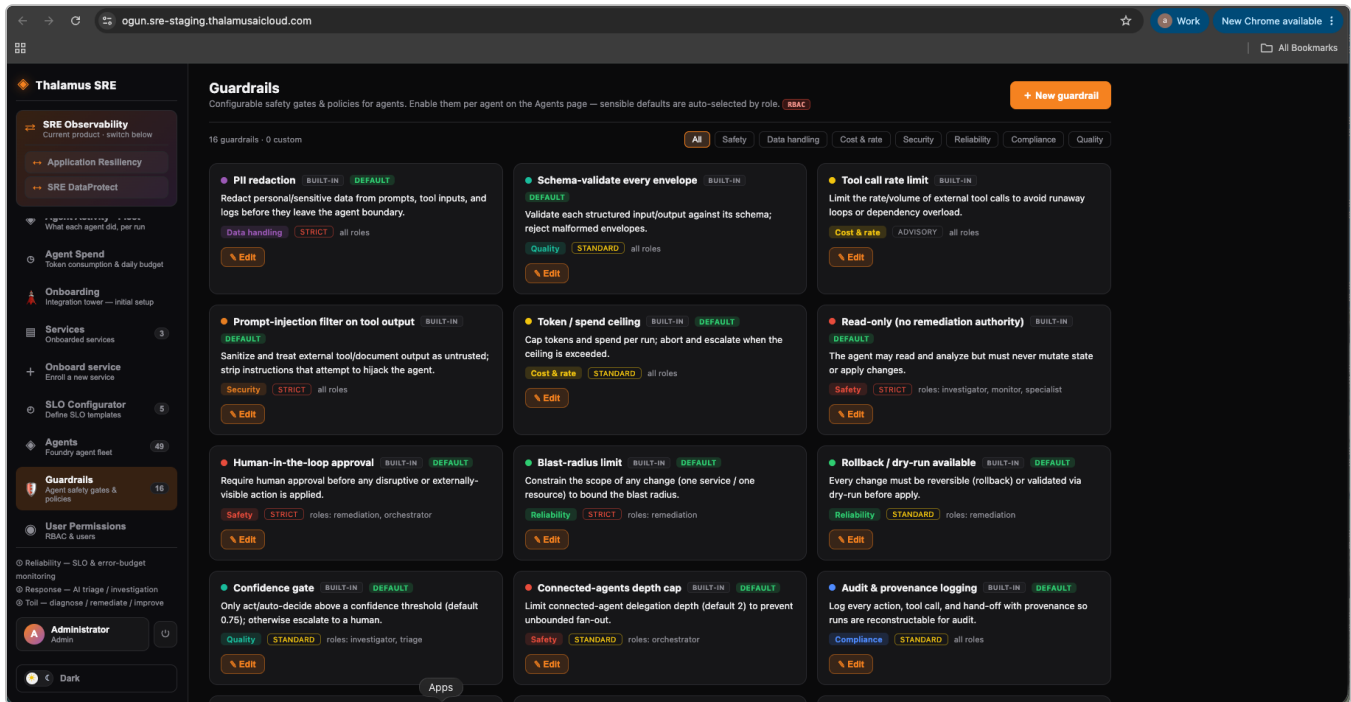


FIGURE 1 — THALAMUS SRE • GUARDRAILS

Authority bounded by policy, not by prompt. *Read-only (no remediation authority)* scoped to investigator, monitor and specialist roles; *human-in-the-loop approval* required before any disruptive action; *blast-radius limit* constraining change to a single service or resource; and *audit & provenance logging* recording every action, tool call and hand-off so a run is reconstructable for oversight. For mission systems the correct configuration is one where the agent analyses and a human decides.

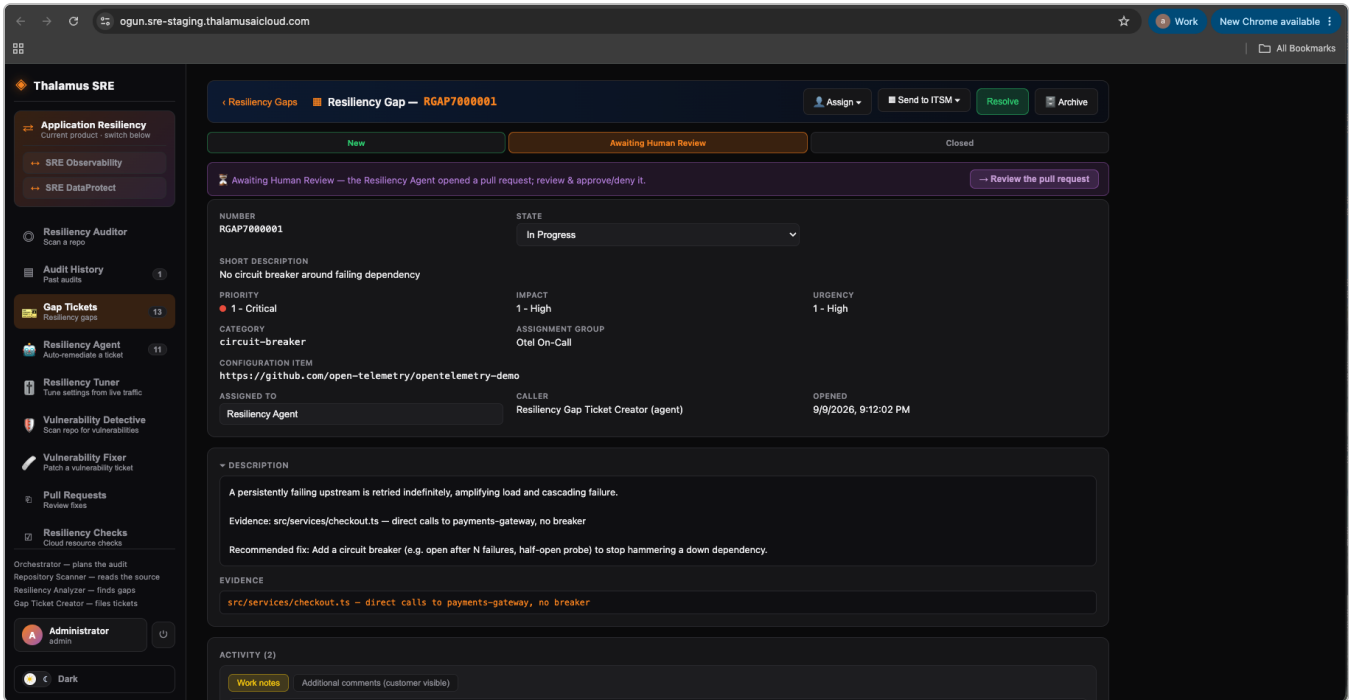


FIGURE 2 — THALAMUS SRE • HUMAN-IN-THE-LOOP

Analysis completed, decision withheld. The agent diagnosed the gap, produced the fix and raised it for review — then stopped. The evidence supporting the finding travels with the record rather than being summarised away, which is the property that makes the decision defensible to an inspector general months later.

07 — GETTING THERE

How Thalamus Advisory helps

ENGAGEMENT

WHAT IT PRODUCES

Enclave Observability Design
4–6 weeks

A deployment architecture that runs entirely inside the accreditation boundary, with the documentation your ATO process requires produced alongside rather than afterwards.

Late-Telemetry Readiness
3–4 weeks

Onset anchoring and retention tiering sized to the real discovery latency of fielded systems, so an event reported three weeks late remains investigable.

Contract SLA Instrumentation
4–6 weeks

For integrators: per-programme objectives and automated evidence generation as a contract deliverable rather than a separate reporting effort.

Deadline Readiness
6–8 weeks before the window

For citizen-facing services: traffic-derived configuration and rehearsal ahead of an enrolment or filing peak that cannot move.

THE COUNTER-ARGUMENT, MADE HONESTLY

Accreditation is slow, and it should be. Introducing any component into a classified enclave is a substantial undertaking with a timeline measured in quarters, and no vendor timeline should compress it.

The realistic sequence is to prove the approach on unclassified systems first — contractor IT, programme management, citizen-facing services — and carry that evidence into the accreditation conversation. Anyone proposing to start inside the boundary is either inexperienced in this sector or not being straight with you about the timeline.

ONE NUMBER

Zero. Bytes of telemetry that may leave a classified enclave for a vendor's SaaS platform.

That single constraint disqualifies most of the observability market, and it is the reason these estates run with less visibility than a mid-sized retailer. Self-hosted is not a deployment preference here — it is the entire question.

Where to start on Monday

For one mission or benefits system, ask: **could we reconstruct a defensible timeline for an event discovered thirty days from now?**

Check the retention window and whether any diagnostic tooling can address a closed window at all. In most enclaves the honest answer is no, and that gap is the work.

Enterprise resilience for the age of autonomous systems. We help organisations build operational practice that stays trustworthy when the systems making decisions are no longer entirely human.

THE RESILIENCE NEWSLETTER · ISSUE 12 · SEPTEMBER 2026

Every figure we print carries its assumptions. Where we model, we say so.

Written for engineering leaders. Forward it to one.