GOVERNED AI WORKFORCE · SOC LEAD

The SOC Lead agent, under zero-trust governance.

A frontier-capable incident synthesizer that reads two upstream analyst feeds, decides severity, and escalates to opus only when the incident earns it — every call minted, policy-checked, and signed into an audit chain.

Asphodel + Tiresias-ZT · ACME Corporation

THE STAKES

An autonomous SOC role is exactly where "trust the model" breaks.

The SOC Lead ingests threat intelligence and analyst findings, then reaches for the most expensive frontier model on the roster the moment it smells a breach. Ungoverned, that is an unbounded blast radius: unbounded spend, unbounded data reach, and no defensible record of why any call was made.

WITHOUT GOVERNANCE

Model picks its own tier. Reads any tenant's memory. No signed evidence of the decision. A prompt-injected "critical" burns frontier budget on demand.

UNDER TIRESIAS-ZT

Deny-by-default policy at the PEP. Per-agent capability token. RLS-isolated data tenant. Every hand-off and escalation written to an Ed25519 audit chain.

THE WORKFORCE

The SOC Lead is one governed identity inside a fleet of 367.

This is a live instance — 367 agents registered as control-plane tenants, each mapped to a real principal in the ACME directory and run under the same zero-trust policy the SOC Lead is bound by.

367	7	360	325,154
GOVERNED AGENTS	SHOWCASE HAND-OFF AGENTS	FLEET AGENTS	DIRECTORY PRINCIPALS

Source: registry (367 agent tenants) · personas.json (7) · agent-manifest.json (360) · tzt_principal — 28 departments, 10 regions, 626 roles.

ARCHITECTURE SLICE

The SOC Lead sits at a junction in a 14-edge hand-off DAG.

Two upstream agents feed it. It synthesizes, decides, and passes a hardened brief downstream. Nothing about that flow is implicit — each edge is a governed call through the enforcement plane.

`threat-intel-researcher` → `security-analyst` → `soc-lead` → `research-lead` → `data-analyst` → `content-editor` → `compliance-officer`

upstream	Two feeds converge threat-intel-researcher findings + security-analyst triage — the SOC Lead's only inputs	2 upstreams
tiers	Home / ceiling / floor per agent keyword escalation (+incident / +breach / +critical) lifts toward the ceiling	home·ceiling·floor
degrade	Fallback-to-floor on LLMError a provider failure never stalls the SOC pipeline	safe-degrade

Source: personas.json · runner.py schedule — 7 showcase agents, 14-edge topological hand-off DAG.

LIVE PROOF

Every SOC Lead call takes the same governed path — no exceptions.

The runner mints a per-agent HS256 capability token from its 64-hex secret. From there the request is validated, policy-evaluated deny-by-default, and forwarded verbatim to the agent's own data tenant.

`runner mints cap-token` → `PEP :8343 — validate + policy eval` → `soul-svc tenant` → `PostgREST` → `acme_soul (RLS write)`

On allow, the PEP signs the call into its audit chain and ships it to the control plane. Usage lands in `tzt_usage` and surfaces on the Grafana ROI board.

Source: governed request flow — PEP host :8343, deny-by-default, forwards verbatim; audit chain → CP; usage → /cp/usage/report → tzt_usage → Grafana.

WALKTHROUGH · 1 OF 3

Synthesis: two upstream outputs become one governed judgment.

The SOC Lead does not gather its own intelligence. It reads what the threat-intel researcher and the security analyst already produced — both written to their own RLS-isolated tenants — and fuses them into a single severity call.

INPUT A — THREAT-INTEL-RESEARCHER

Upstream domain brief on emerging indicators, delivered as a governed CoT + memory write in its own soul-svc tenant.

INPUT B — SECURITY-ANALYST

Triaged findings from the analyst stage — the second and final upstream edge feeding the SOC Lead node.

> The SOC Lead reads across the hand-off — but only what policy grants. Cross-tenant reach requires an explicit service role; it is never the default.

Source: personas.json hand-off DAG (2 upstream edges into soc-lead) · Plane 4 — FORCE RLS, cross-tenant needs SET ROLE service_role.

WALKTHROUGH · 2 OF 3

Escalation: "critical" is a governed decision, not a free reach for opus.

Keyword escalation is the mechanism — when the synthesized incident carries +incident, +breach, or +critical, the agent lifts from its home tier toward its ceiling. The frontier model is reachable, but bounded by the ceiling policy.

trigger	Keyword escalation +incident / +breach / +critical in the synthesized brief	rule-based
ceiling	Frontier tier — opus-4.8 reached only for high-severity, and only up to the agent's ceiling	opus-4.8
home	Routine severity stays on cheaper tiers sonnet-4.5 and free/Ollama tiers carry the everyday load	sonnet · free
floor	Fallback-to-floor on LLMError a degraded escalation still completes rather than failing open	safe-degrade

Source: personas.json — per-agent home/ceiling/floor tiers + keyword escalation, fallback-to-floor on LLMError · config.json roster (opus-4.8, sonnet-4.5).

WALKTHROUGH · 3 OF 3

The hand-off: a signed, isolated write — then downstream to research-lead.

The SOC Lead's judgment is not a chat message. It is a governed CoT and memory write into its own tenant, validated end to end, and it becomes the input to the next stage of the DAG.

mint	Runner mints per-agent HS256 capability token from the agent's 64-hex capability secret	HS256
enforce	PEP validates api-key + cap-token, evaluates policy deny-by-default; on allow, forwarded verbatim	:8343
write	CoT capture + memory write to acme_soul RLS-forced, tenant-isolated — POST /v1/cot/capture, /v1/memory/write	RLS
sign	PEP signs the call into its audit chain SoulKey Ed25519, durable WAL at /data/audit-wal, ships to CP	Ed25519
pass	Output flows downstream to research-lead the next of the 14 edges in the hand-off DAG	next edge

Source: Plane 3/4/5 — governed request flow · signed audit chain (SoulKey Ed25519) · endpoints /v1/cot/capture, /v1/memory/write.

GOVERNANCE & AUDIT

Every decision is reviewable, and drift is contained automatically.

The SOC Lead operates under continuous behavioral monitoring. If its trajectory drifts, SoulWatch quarantines it without waiting for a human — and an analyst can verify the whole audit chain after the fact.

BEHAVIORAL CONTAINMENT

SoulWatch scores behavioral drift and auto-quarantines the agent into a shared postgres store. RFC5424 syslog streams to Promtail:1514 → Loki as SIEM.

AFTER-THE-FACT REVIEW

The PEP `/admin/*` console exposes audit/verify, soulwatch/status, quarantine, and analytics — trajectory, model_baseline, campaign, drift. Role-token gated at pdp.tiresias.watch.

Source: Plane 3 SoulWatch → auto-quarantine, SIEM RFC5424 → Promtail:1514 → Loki · Review console = PEP /admin/* (pdp.tiresias.watch).

ECONOMICS

Governing the tier is governing the bill.

Because escalation is bounded by policy, opus is reserved for the incidents that warrant it. The observed per-call economics show why that boundary matters — the frontier tier costs two orders of magnitude more than the floor.

MODEL TIER	ROLE FOR SOC LEAD	OBSERVED $/CALL
opus-4.8	Ceiling — high-severity only	$0.19
sonnet-4.5	Routine synthesis	$0.03
deepseek-v4-pro	Cheap capable tier	$0.002
gpt-oss:120b	Self-host tier	$0.0015
gemma:free	Free floor	$0.00

Source: ROI dashboard uid acme-roi-economics — observed per-call cost from tzt_usage.

CLOSE

A frontier-capable SOC Lead you can actually put in production.

Two upstreams synthesized, severity decided, opus reached only when it is earned — and every step minted, policy-checked, RLS-isolated, and signed. The capability is the easy part. The governance is what makes it deployable.

2	deny	Ed25519	$0.19
UPSTREAM FEEDS SYNTHESIZED	DEFAULT POLICY POSTURE	SIGNED AUDIT CHAIN	BOUNDED FRONTIER CALL

Asphodel + Tiresias-ZT · ACME Corporation · governed, live, on GCP