TIRESIAS-ZT · CISO BRIEFING

Govern the AI workforce, don't trust it.

367 autonomous agents and 205,034 non-human identities are a live attack surface. Every call they make is authenticated, policy-checked, and signed into a tamper-evident record — or it does not execute.

Asphodel + Tiresias-ZT · ACME Corporation

THE STAKES

Your largest identity population no longer has a human behind it.

The directory holds 325,154 principals. Two-thirds are machines: service and agent identities that authenticate, call APIs, and write data around the clock. Ungoverned, each is an unattended credential inside your data plane.

205,034	120,120	367	28
NHI / SERVICE IDENTITIES	HUMAN PRINCIPALS	AUTONOMOUS AGENTS IN RUN	DEPARTMENTS · 10 REGIONS

Source: LDAP directory sync (Samba-AD, base DC=acme,DC=example) → tzt_principal · live count.

WHAT IS RUNNING

A real governed workforce — 367 agents doing real work.

Seven showcase agents run a 14-edge hand-off DAG; 360 fleet agents are mapped like-for-like to the Acme org across all 28 departments. Each is a registry tenant with its own capability secret and its own isolated data backend.

SHOWCASE	7 agents on a topological hand-off DAG threat-intel → security-analyst → soc-lead → research-lead → compliance-officer, with keyword escalation on +incident / +breach / +critical	14 edges
FLEET	360 agents mapped to real dept × role-family independent rolling batches, ~1–2 governed jobs/day/agent	360
REGISTRY	Every agent is a postgres-backed tenant durable registry, enrollment ledger, usage, audit, quarantine — survives restart	367

Source: personas.json + agent-manifest.json · registry store (postgres, TZT_REGISTRY_STORE).

THE CONTROL THAT MATTERS

Two factors on every call, then deny-by-default.

An agent cannot reach data on identity alone. The enforcement plane requires the tenant API key and a per-agent HS256 capability token the caller mints from its own 64-hex secret. Policy is deny-by-default; only on allow does the PEP forward — verbatim — to that agent's backend.

FACTOR 1 — TENANT API KEY

X-Tiresias-Api-Key identifies which of the 367 tenants is calling. No key, no evaluation.

FACTOR 2 — CAPABILITY TOKEN

X-Capability-Token: HS256, minted per call from the agent's 64-hex capability secret. Proves possession, not just identity.

> No key and matching token, no policy pass — no data. The default answer is no.

Source: PEP (acme-tiresias-zt, TZT_ROLE=proxy, host :8343) · deny-by-default policy eval.

LIVE PROOF

This is not a diagram. It is instrumented and running.

The enforcement plane pulls the 367-agent tenant slice from the control plane, evaluates every governed request, and writes a signed record. The numbers below are read from the live stores, not a spec.

367	deny	Ed25519	100%
GOVERNED AGENT TENANTS	DEFAULT POLICY POSTURE	SIGNED AUDIT WAL	CALLS TWO-FACTOR AUTHED

Source: PEP policy engine + SoulKey audit chain · durable WAL at /data/audit-wal · ships to control plane.

WALKTHROUGH · THE PATH

One job, end to end — every hop is a gate.

A runner mints the capability token; the PEP is the only door to data; the data plane re-validates the bearer; and Postgres enforces isolation at the row. Nothing reaches the store on a single unchecked assertion.

runner mints HS256 → PEP · authn + policy → soul-svc tenant → PostgREST → acme_soul · RLS write

> The PEP is the only path to data. Verbatim forward on allow; drop on deny; signed either way.

Source: governed request flow — PEP → acme-soul-svc → Caddy/PostgREST → acme_soul (Cloud SQL).

WALKTHROUGH · BLAST RADIUS

Forced row-level isolation — one agent cannot read another.

Each agent is a distinct soul-svc tenant keyed by sha512 of its bearer. Row-level security is FORCED on the memory and key tables, so even the database owner is subject to it. A cross-tenant read requires an explicit privilege escalation, not a bug.

FORCE RLS	Enforced on _memories and _soul_api_keys postgres role on Cloud SQL is itself subject to FORCE RLS	per-tenant
TENANT KEY	_soul_api_keys.key_hash = sha512(bearer) no shared credential across the 367 agents	sha512
CROSS-TENANT	Requires SET ROLE service_role explicitly auditable, deliberate — never the default path	deny

Source: acme_soul (Cloud SQL, RLS forced) · authorized-networks only · JWT-gated Cloud Run tier.

WALKTHROUGH · DRIFT

Behavioral drift becomes containment, automatically.

SoulWatch watches how each agent behaves against its baseline. When an agent drifts, it is auto-quarantined into the shared quarantine store — pulled out of the run without a human in the loop — while the event streams to the SIEM.

DETECT

SoulWatch scores behavioral drift per agent against its model baseline and trajectory.

CONTAIN

Drift → auto-quarantine in the postgres quarantine store. The agent stops transacting immediately.

> Anomaly detection that acts, not just alerts — quarantine is the default response to drift.

Source: SoulWatch → auto-quarantine (shared postgres store) · RFC5424 syslog → Promtail:1514 → Loki.

GOVERNANCE · AUDIT

Every call is signed into a tamper-evident chain.

The PEP signs each governed request into an Ed25519 audit chain with a durable write-ahead log, then ships it to the control plane. Because entries are chained and signed, an altered or missing record is detectable — the audit is evidence, not a log you have to trust.

Ed25519	WAL	→ CP	RFC5424
SOULKEY CHAIN SIGNATURE	DURABLE AT /DATA/AUDIT-WAL	SHIPPED TO CONTROL PLANE	SYSLOG → LOKI SIEM

Source: PEP signed audit chain (SoulKey Ed25519) · /admin audit/verify console (pdp.tiresias.watch).

GOVERNANCE · SEPARATION

Enforcement, control, and data are separate planes.

The control plane is in-network only; the PEP is the sole published data door; the data tier is a dedicated Cloud SQL instance reachable from authorized networks only. Compromising a runner does not hand an attacker the policy engine or the database.

PLANE	WHERE	SCOPE
Enforcement (PEP)	acme-demo :8343	published
Control plane	tiresias-cp :8343	in-network only
Identity store	postgres :5432	in-network only
Data tier (acme_soul)	Cloud SQL :5432	authorized-nets only
Directory	Samba-AD :636	VPC (LDAPS)

Source: port map · control plane X-Admin-Key gated · data tier JWT-gated behind PEP.

ECONOMICS OF GOVERNANCE

Governance routes work to the cheapest safe tier — visibly.

Because every call is metered into the usage store, the ROI board shows true per-model economics. Routing favors free and self-hosted tiers under policy; the expensive frontier models are reserved for escalation, and you can see the difference per call.

MODEL TIER	$/CALL OBSERVED
opus-4.8 (frontier)	$0.19
sonnet-4.5	$0.03
deepseek-v4-pro	$0.002
gpt-oss:120b (Ollama)	$0.0015
gemma-4-31b:free	$0.00

Source: ROI dashboard (uid acme-roi-economics) · tzt_usage → Grafana · observed $/call.

THE CLOSE

Provable governance, not trust.

367 agents and 205,034 machine identities operate under two-factor call authentication, deny-by-default policy, forced tenant isolation, automatic drift quarantine, and a signed audit chain. Every action is authenticated, checked, and provable — or it does not happen.

Asphodel + Tiresias-ZT · ACME Corporation