CASE STUDY · FORTUNE 1000 ENTERPRISE

How a Fortune 1000 enterprise built one record of its AI estate

9.5M observations and 6M traces a month from 170+ agents, all in one record.

9.5M
observations a month: individual model calls, tool calls and steps

6M
traces a month: end-to-end AI requests and agent runs

170+
AI agents in production, each with an owner, a cost and a risk score

AT A GLANCE

CUSTOMER
Fortune 1000 enterprise

INDUSTRY
IT consulting

SCALE
125,000 employees

PRODUCTS
Discover AI, Gauge AI, AI Agent Center

DATA IN
Endpoint AI Agent, browser extension, SDK / OTel

IDENTITY
Okta

"We went from guessing to knowing. In four weeks we had a full inventory of our AI and a named owner for every one of our 170+ agents. That's the foundation our governance program was missing."

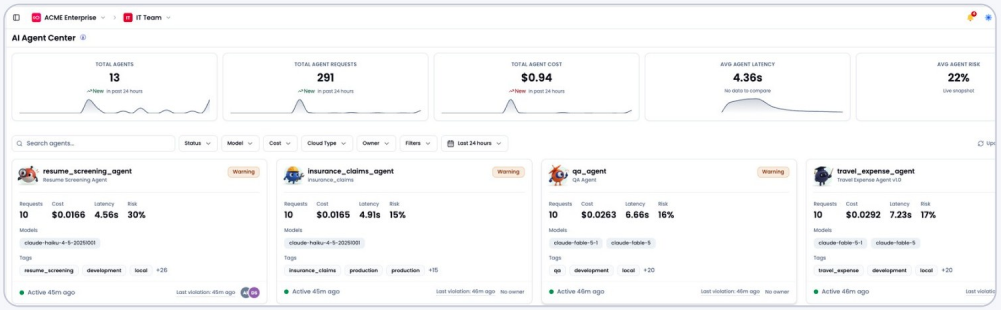
CIO, Fortune 1000 enterprise

The challenge

AI arrived faster than IT could catalogue it: browser and desktop assistants, coding assistants, and agents on cloud AI platforms. Security tools saw network traffic, not AI activity with a person attached. The CIO's office couldn't say which AI was in use, who used it or what it cost, and the board was asking.

The solution

The company made ANTS Platform its system of record for AI. The Endpoint AI Agent and browser extension inventory AI tools; SDK and OpenTelemetry instrumentation captures traces from agents and AI apps. Every event is tied to a named user and department. Discover AI keeps the inventory current, Gauge AI shows where AI is working, and the AI Agent Center gives every agent a cost, a risk score and an owner.



AI Agent Center: every agent with cost, latency, risk, models and owner. Demo tenant.

The results

- 9.5M observations and 6M traces a month from 170+ agents in one record, each linked to an owner, a model and a cost.
- 4 weeks from kickoff to the first complete AI inventory.
- Complete visibility of AI agent ownership: every one of the 170+ agents has a named owner.

CASE STUDY · FORTUNE 1000 PLATFORM ENGINEERING

How a Fortune 1000 team put every Anthropic model call on a budget

Budget AI as the AI gateway for platform engineering's Claude models.

4,500
developers reaching Anthropic models through Budget AI

$5M
annual Anthropic model spend now under budget control

All
of platform engineering's Anthropic traffic runs through one gateway

AT A GLANCE

CUSTOMER
Fortune 1000 enterprise

DEVELOPERS
4,500

INDUSTRY
Software

PRODUCT
Budget AI as the AI gateway

MODELS
Anthropic Claude models

DEPLOYMENT
Self-deployed in the customer's environment

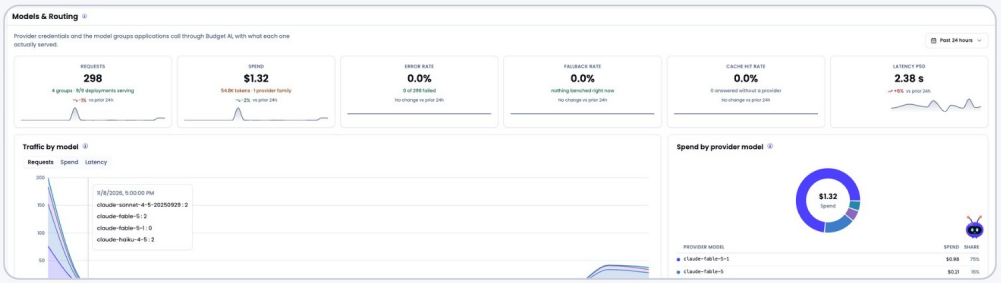
> *"Budget AI gave 4,500 developers one front door to Anthropic. We retired the 10% of keys nobody was using and cut overspend by 20%, without slowing a single team down."*
>
> **VP of IT, Fortune 1000 enterprise**

The challenge

Platform engineering provides Anthropic Claude models to 4,500 developers, a $5M-a-year line item. Provider API keys spread across services and scripts, spend arrived as one invoice with no breakdown by team, and nothing could stop a runaway workload before the bill.

The solution

The team put Budget AI in front of every Anthropic model as the company's AI gateway. Provider credentials are held centrally and never handed out; each app, team or agent gets its own revocable ANTS key with an owner and tags. Claude models sit in model groups, and budgets follow the org chart from project to sub-department: hard budgets refuse requests at the limit, soft budgets alert. Spend and latency are visible by model, key and team.



Budget AI · Models & Routing: requests, spend and latency by Claude model. Demo tenant.

The results

- Every Anthropic model call goes through one gateway, tied to a key and an owner.
- $5M a year in Anthropic spend visible by model, key and department, under budgets that hold.
- Retired the 10% of keys that were unused, cutting overspend by 20%.