



CONTROL
PLANE

CASE STUDY

A Virtual Cloud Built for Secure Inference

“

“Building things is so much easier now. But then you hit architecture, security, infrastructure. Control Plane helps us clear that hurdle fast.”

ROSS HENDRICKSON, CTO, INSPECTIV

 **Inspectiv**

KEY DETAILS



Custom AI Inference Layer on Control Plane-Sourced GPUs



Guardrails for Shipping Secure Code at AI Speed



~1/3 the Cost of DIY Infrastructure

COMPANY

Inspectiv

LOCATION

Culver City, CA, USA

INDUSTRY

Cybersecurity /
Application Security

COMPANY SNAPSHOT

Inspectiv is an AI-first security company that has spent about ten years securing some of the largest brands in the business, running large-scale bug bounty programs, scanning environments, and performing penetration testing. Its new AI-native product turns researcher findings into code-level fixes, running on Inspectiv's own AI inference layer.

PROBLEM

Infrastructure Built for Yesterday's Problems

Over the last two years, AI has multiplied the volume of work on both sides of Inspectiv's platform: customers are shipping more code, and researchers are submitting more findings. "With the advent of AI usage across the board, we are getting more code than we've ever had, with more vulnerabilities in it than we've ever had," says CTO Ross Hendrickson.

When Ross joined about four years ago, Inspectiv ran its infrastructure the way most companies do: build by hand and then babysit. They ran a couple of Kubernetes clusters on bare metal in a single region, no infrastructure as code, and a full-time ops person keeping it alive.

That setup could have kept running. It just couldn't change easily. But Inspectiv's requirements were changing. First SOC 2 Type II showed up in enterprise deals. Then customers barred foundation-model APIs, which put Inspectiv in the business of running its own inference layer on GPUs. Then engineers started generating code faster than the deployment process could keep up.

ROSS HENDRICKSON
CTO, INSPECTIV



“

Databases would fall over because we were running the databases inside of Kubernetes. Services would fall over. Products would stop. People couldn't log in. It was extremely chaotic.”

THE SOLUTION

A Cloud Shaped to Their Workloads

Ross evaluated emerging platforms and open-source options against clear criteria: "It needed to be out of the box, almost zero configuration on our side, and Kubernetes-native, so we're not doing something completely custom," with support for Inspectiv's compliance requirements and uptime guarantees. Control Plane's model of combining clouds into one virtual cloud shaped to your workloads met the criteria. Today, Inspectiv's gateways route traffic through Control Plane to dozens of microservices, with Postgres underneath, NATS for event-driven messaging, AWS retained for the storage layer, and GPUs on the same fabric for inference.

Security Out of the Box

"We have secure workloads, we have least-privileged access to those secure workloads. A lot of things you used to have to piece together from different open source or different vendors, you just get out of the box with Control Plane."

"Control Plane makes it super simple for me to say yes to some of the things that used to be tremendously hard."

Ross Hendrickson, CTO,
Inspectiv



**CONTROL
PLANE**

THE SOLUTION

An Inference Layer on Their Own Terms

Inspectiv's new AI-native product runs a find-and-fix loop: vulnerabilities come in from any source (pen tests, scans, the researcher community), and customers get code-level remediation guidance: "you basically get a PR to close out whatever the researcher community finds with automatic remediation validation." Two years ago, while the first version was being built, Inspectiv's customers set a hard constraint: no data to the foundation-model providers. The answer was an inference layer of Inspectiv's own. "We have really sophisticated AI, and our customers' data is not going outside of our zone of control. It's us and Control Plane, and that's it."

Running custom inference on GPUs accessed through Control Plane lets Inspectiv deliver ready-to-apply fixes in a way that fits both its security model and its pricing model. "It gives us a bunch of advantages versus just straight up running inference tokens to a foundation model. We have cost controls. We have scalability. And we have security." It also insulates them from foundation-model pricing swings: "It's really hard when a foundation model changes their pricing structure and suddenly it's 10x. With Control Plane, we follow our curve of usage. Our cost variance is never going to be more than our usage." Also, sourcing the GPUs for their inference layer didn't require a new vendor: "I don't have to vet another vendor or go through compliance and security just to get access to specific GPUs."

"I had assumed there would be some level of support. But when things went wrong, we had five or six people in Slack threads, huddles going, Zoom calls, helping diagnose and fix. It's like getting an NBA team to come play your high school tournament for you."

**Ross Hendrickson, CTO,
Inspectiv**

**ROSS HENDRICKSON
CTO, INSPECTIV**



“

Our customers said: do not send my data to OpenAI. Do not send my data to Anthropic. We solved it by using GPUs in our own inference layer.”

Guardrails: Secure Code at Velocity

Inspectiv built its delivery process around strict controls. “What we’ve leveraged Control Plane for is streamlining the guardrails, the process to get code that’s coming from AI secure and reviewed.” Engineers point AI at templates to stand up a new service on a subdomain, open PRs against infrastructure as code, and spin up ephemeral whole-stack environments for a couple of hours at a time. The internal bar is building a feature and releasing it in a day. “A lot of them have commented that this is the easiest platform to work on: everything out of the box, all in one platform. They can just go do their work.”

More Infrastructure at the Same Cost

Inspectiv’s spend on Control Plane today is about what its infrastructure cost when Ross started, even as the platform has grown from a handful of microservices to dozens of systems plus the GPU inference layer, all scaling with usage. “When workloads are running, I’m billed. When they’re not, I’m not.”

Summary

Inspectiv’s old infrastructure solved the problems it was built for, and none of the ones that came after: a custom inference layer, secure code shipped at AI pace, enterprise compliance. Control Plane addressed all three requirements in a single platform. The same adaptability carries forward: new models, new security threats, new regulations, or the next shift in compute can be absorbed the way the last round was: by reshaping the virtual cloud, not rebuilding the infrastructure.

“If I was to go build what we have and run it in a more traditional way, it would probably cost us three to four times as much. Plus I’d have to hire two or three experienced people just to run infrastructure.”

Ross Hendrickson, CTO,
Inspectiv

About Control Plane

Control Plane is a cloud virtualization platform enabling you to combine any cloud infrastructure into one AI-native virtual cloud.

www.controlplane.com



**CONTROL
PLANE**



CONTROL
PLANE