AVIATRIX

SOLUTION BRIEF

Contain AI Agent Harnesses — OpenClaw / NemoClaw

Powered by Aviatrix Cloud Native Security Fabric

Aviatrix Validated Containment Architectures are lab-tested containment deployment blueprints for the AI platforms enterprises are actually running. Ship-ready, policy-included, validated before they arrive. This Validated Containment Architecture covers the OpenClaw / NemoClaw Agents.

FOR PLATFORM AND AI WORKFLOW ENGINEERS

Deploy powerful agentic loops and workflows in private subnets, allow only approved egress destinations, and produce independent security evidence of allow and deny decisions.

The Threat

Agent Harness technologies like OpenClaw, Hermes, and NemoClaw are valuable because they keep working after the first prompt: research, coding, ticket triage, SaaS automation, browser tasks, and internal API calls. The same capabilities make them privileged non-human operators. If the runtime is prompt-injected, extended with a risky skill, or supplied with a compromised package, the question is not only what the agent intended to do. The question is what the network actually let it reach.

- **Prompt injection or tool abuse** redirects sensitive data to an attacker-controlled host.
- **Supply-chain compromise** – a malicious package, plugin, or skill phones home from the terminal app.
- **Shadow model calls** – the agent sends prompts or retrieved data to an unapproved model provider.
- **DNS exfiltration** – sensitive data is encoded in DNS queries sent to an attacker-controlled domain over UDP or TCP ports.
- **Lateral movement** – a compromised agent VPC pivots to adjacent workloads in the cloud fabric.

The OpenClaw Agent Runtime Containment Validated Containment Architecture is a lab-reviewed, policy-included AWS blueprint that places an Aviatrix Spoke Gateway in the egress path of the agent host and enforces a default-deny destination policy with Aviatrix Distributed Cloud Firewall (DCF). It treats the agent VM as a potentially compromised operator and moves the control point to the cloud network – outside the runtime the agent can influence.

TRUST BOUNDARY

The enforcement point is the Aviatrix Spoke Gateway and DCF policy, not the agent runtime. Sandboxing, model guardrails, IAM, and MCP authorization remain in place; this Validated Containment Architecture adds the independent network control the agent cannot rewrite from inside the harness.

The Architecture

The baseline deployment places the agent runtime in a private subnet with no public IP and no native NAT or Internet Gateway route. The subnet default route points to the Aviatrix Spoke Gateway. DCF applies ordered policy using SmartGroups for source identity and WebGroups for destination allow-lists. CoPilot FlowIQ records DCF allow and deny decisions; VPC Flow Logs provide complementary AWS network-flow telemetry. Administration is through AWS SSM Session Manager, so the host needs no inbound rule and no public SSH.

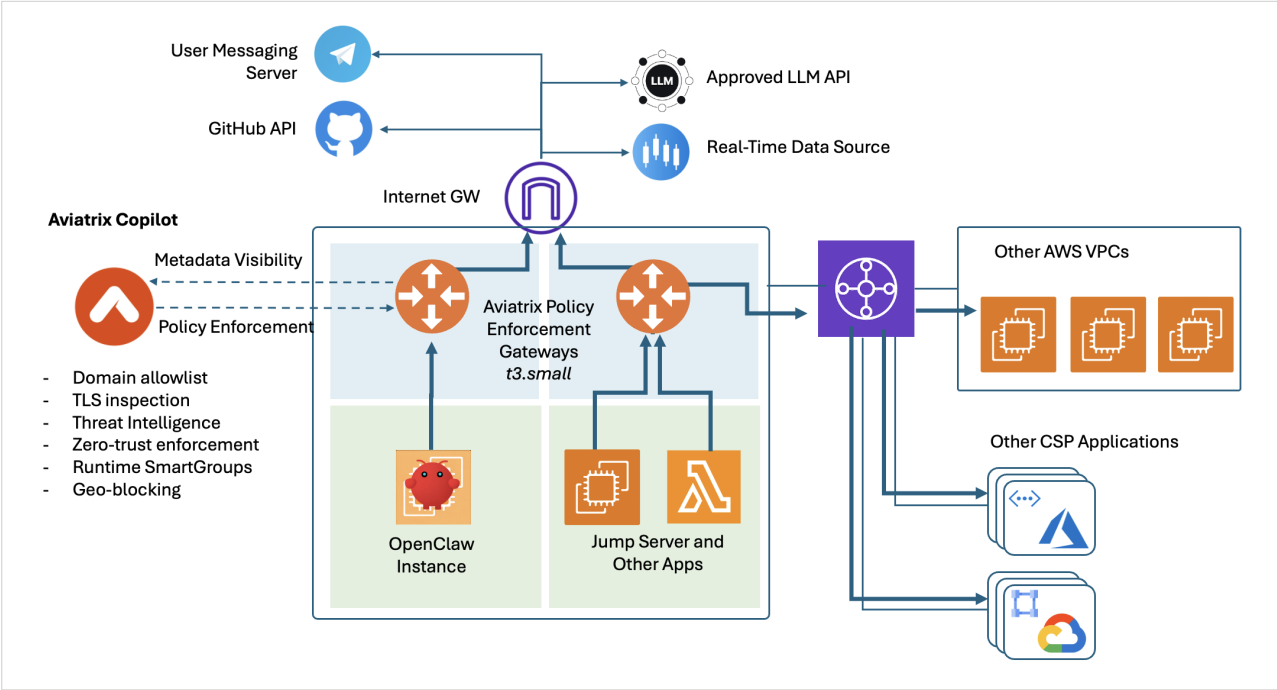

Figure 1. Reference insertion pattern for AWS agent runtime egress containment.

Component	Purpose for platform engineers
Private agent subnet	Runs the VM or container host with no public IP. Access is through SSM, not inbound SSH from the internet.
Aviatrix Spoke Gateway	In-path egress enforcement outside the agent runtime. The subnet default route sends outbound traffic to the gateway.
Distributed Cloud Firewall	Applies ordered deny and permit rules. The default action denies any routed destination that is not explicitly permitted.
SmartGroups	Identify the source: agent subnet, agent class, environment, team, or Kubernetes label selector.
WebGroups	Define approved web destinations and FQDNs per agent class. Use DCF network rules for approved non-web protocols and services.
CoPilot FlowIQ	Shows allowed and denied flows with source identity, destination, rule name, action, and timestamp.

What's Included

Deliverable	What ships
AWS Terraform blueprint	VPC, private subnet, OpenClaw/NemoClaw VM, SSM-only management, VPC Flow Logs, Aviatrix Spoke Gateway, and DCF policy objects.
Agent-class presets	Locked-down, coding, research, support, and demo tfvars profiles to copy instead of starting from a blank policy.
Validation tests	Verify approved destinations are reachable; unapproved destinations, external DNS resolvers, and unapproved model APIs are blocked; and each result is visible in FlowIQ.
Operations docs	README, AGENTS.md, security model, domain tiers, preflight checks, route checker, and rollback guidance.

Compliance Evidence

For SOC 2, HIPAA, PCI-DSS, FedRAMP, DORA, ISO 27001, NIST AI RMF, and EU AI Act programs, CoPilot FlowIQ provides runtime-independent evidence of what the agent was allowed to reach, what it attempted to reach, and which policy made the decision. Terraform history adds peer-reviewed, versioned proof of every destination change tied to an owner.

Get Started

This Validated Containment Architecture is available now. Use the blueprint as the first secure agent landing zone: one AWS account, one VPC, one agent class, monitor first. Once the class is stable, enforce it and make it a reusable path for other teams. The result is a deployable architecture engineers can understand, security teams can audit, and platform teams can scale.

SCOPE NOTE

This Validated Containment Architecture focuses on egress containment. The same Aviatrix Spoke Gateways and DCF SmartGroups can later add east-west microsegmentation between agent spokes and adjacent workloads as a phase-two layer – it is not required for the first AWS egress deployment.

Request a 30-minute architecture review.

We walk through the policy model in your environment, identify the OpenClaw, NemoClaw, or Hermes agents you have running today, and show you the egress paths you currently cannot see.

About Aviatrix

Aviatrix® is pioneering the Cloud Native Security Fabric – the architecture the Containment Era requires. The Cloud Native Security Fabric governs every workload communication path across every cloud, every VPC, every Kubernetes cluster, and every serverless function, from a single policy engine. Universal propagation. Enforced at the workload during runtime. Trusted by more than 500 of the world's leading enterprises. For more information, visit aviatrix.ai.