AVIATRIX

TECHNICAL BRIEF

Contain Gemini Enterprise Agents

Powered by Aviatrix + Google Cloud

Aviatrix Distributed Cloud Firewall · Reference Architecture for Gemini Enterprise Agent Platform

Threat Context

Gemini Enterprise Agent Platform run in a Google-managed tenant project. By default, their tool calls, MCP connections, and non-Gemini Enterprise Agent Platform model calls are indistinguishable on the wire from legitimate work, with no inline enforcement point the customer's security stack can observe. Two threat vectors are in scope for this Validated Containment Architecture:

- Egress to attacker infrastructure via prompt injection or compromised dependency (OWASP LLM01, LLM05).
- Egress to unsanctioned model providers, including OpenAI, Anthropic, Mistral, and Perplexity (OWASP LLM08 excessive agency).
- DNS-tunneled exfiltration via UDP/53 to external resolvers.
- Lateral movement from the agent spoke to adjacent workloads in the customer fabric.

LAB-VALIDATED THREAT SCENARIO

The scenario set proves the architecture against OWASP LLM Top Ten and MITRE ATLAS in a live deployment. LLM01 prompt injection driving tool-abuse exfiltration is closed by default-deny on the attacker domain. LLM02 DNS-tunneled exfiltration is closed by the UDP/53 deny. LLM05 supply-chain compromise on a sanctioned host is closed on the GKE shape by selective TLS decryption and a URL-filter deny, proven side-by-side: compromised path blocked, clean path returns HTTP 200. LLM08 excessive agency (shadow model call) is closed by the named vca-vertex-shadow-model-deny rule. Every blocked flow is visible in CoPilot FlowIQ with a human-readable rule name carrying the shape suffix (-a for managed runtime, -b for GKE).

Insertion Pattern

This Validated Containment Architecture supports two deployment shapes. Both use the same Distributed Cloud Firewall policy model and the same CoPilot dashboard.

Managed runtime shape (Shape A)

Gemini Enterprise Agent Platform's managed runtime runs in a Google-managed tenant project the customer cannot see into. The insertion pattern forces all non-Google agent egress out of that tenant project through a Private Service Connect interface (PSC-I) network attachment into a customer VPC. Inside that VPC, the Aviatrix gateway is inserted at Layer 3: the network-attachment subnet routes to the gateway, which performs NAT and acts as a transparent forward proxy.

There is no HTTPS_PROXY and no change to the agent. The agent makes the same outbound calls it always made. The gateway is transparently in path. The gateway is also the RFC 1918 egress next-hop Google requires under VPC Service Controls, satisfied at the route layer, not in the application. RFC 1918 east-west traffic, such as calls to a private MCP server, routes natively over PSC-I and is inspected on the same gateway.

A standalone single-spoke deployment with no transit is supported as an on-ramp and extends to the transit topology later.

GKE shape (Shape B)

For a GKE-hosted custom ADK runtime, enforcement runs through Aviatrix Kubernetes SmartGroups and spoke in-path enforcement on pod egress. The GKE shape adds selective transparent TLS decryption for URL-path enforcement, scoped tightly to a destination FQDN SmartGroup. The one optional container change for this shape is adding the Aviatrix CA to the agent container image's trust store.

What does not change in either shape

Neither shape requires an HTTPS_PROXY setting, an SDK change, a redeploy of the agent, or any modification to the application. The enforcement point is at the network layer. The agent does not know it is there.

Rule name suffix convention

DCF rules carry a shape suffix to distinguish them in FlowIQ logs: -a for the managed runtime shape and -b for the GKE shape. Every blocked flow is visible in CoPilot FlowIQ with a human-readable rule name.

Prerequisites

Before configuring Distributed Cloud Firewall enforcement, verify the following are in place:

- Aviatrix Controller 8.1+ for SNI and domain baseline; Controller 9.0+ for TLS decryption, URL-path filtering, and egress IDS/IPS.
- GCP project with VPC Flow Logs enabled.
- IAM role for the Aviatrix controller with permissions to create Network Attachments and manage Private Service Connect endpoints.
- Gemini Enterprise Agent Platform agent deployed with a PSC-I network attachment configured (managed shape) or ADK runtime on GKE with the agents namespace (GKE shape).

- Private Service Connect endpoints for Gemini Enterprise Agent Platform (aiplatform.googleapis.com) and googleapis.com, with Cloud DNS private zone and peering configured.
- CIDR range defined for the network-attachment subnet (minimum /28, two IPs per max_instances).
- For GKE shape: Aviatrix CA distributed to the agent container image trust store.

DEPLOYMENT TIME

Under 45 minutes on a fresh GCP project with one terraform apply per layer. The managed-runtime shape alone comes up in about 25 minutes. GKE provisioning is the long pole. Destroy is one command and leaves zero orphans.

SmartGroup and WebGroup Design

Object	Type / Scope / Purpose
sg-agent-psc-subnet	CIDR SmartGroup over the PSC-I network-attachment subnet. Sized at two IPs per max_instances, minimum /28. Source identity for the managed-runtime shape.
sg-agent-gke	Kubernetes SmartGroup matching k8s_namespace=agents. Source identity for the GKE shape.
sg-vertex-psc	CIDR SmartGroup over the Gemini Enterprise Agent Platform Private Service Connect endpoint. The only sanctioned model-class destination.
wg-sanctioned-tools	WebGroup of approved tool FQDNs defined by the platform team. Updated via pull request against the Terraform repo.
wg-shadow-models	WebGroup covering api.openai.com, *.anthropic.com, api.mistral.ai, *.perplexity.ai, and any other unsanctioned model providers named by governance. Used in vca-vertex-shadow-model-deny.
wg-supply-chain-ioc (GKE shape)	FQDN SmartGroup scoping selective TLS decryption. Narrow-scoped to destinations where URL-path filtering is needed (e.g., raw.githubusercontent.com). Everywhere else decrypt_policy=DECRYPT_NOT_ALLOWED.

CRD-defined WebGroups (per MCP server) are separate from the above and live in the cluster, not on the controller.

Distributed Cloud Firewall Policy Pack

Rules are ordered. The shadow-model deny is placed ahead of the tool allow-list. Rule names carry a shape suffix: -a for the managed runtime, -b for GKE.

Rule Name	Action	Source	Destination	Notes
vca-vertex-shadow-model-deny-a/b	Deny + Log	sg-agent-psc-subnet / sg-agent-gke	wg-shadow-models	Placed ahead of tool allow. Logged by name in FlowIQ.
vca-vertex-allow-vertex-a/b	Permit	sg-agent-psc-subnet / sg-agent-gke	sg-vertex-psc	First-party model traffic. Never decrypted.
vca-vertex-allow-tools-a/b	Permit	sg-agent-psc-subnet / sg-agent-gke	wg-sanctioned-tools	Approved tool FQDNs. Updated via PR.
vca-vertex-deny-dns-a/b	Deny	sg-agent-psc-subnet / sg-agent-gke	Any, UDP/53	Blocks DNS-tunneled exfiltration.
vca-vertex-deny-eastwest-a/b	Deny	sg-agent-psc-subnet / sg-agent-gke	All other spokes	East-West isolation. Blast Radius stops at agent VPC.
vca-vertex-url-filter-b (GKE shape only)	Deny	sg-agent-gke	wg-supply-chain-ioc (URL path match)	Requires Controller 9.0+. Selective TLS decryption on wg-supply-chain-ioc scope only.
vca-vertex-default-deny-a/b	Deny + Log	sg-agent-psc-subnet / sg-agent-gke	Any	Default-deny catch-all. Must be last.

TLS Decryption Policy

Destination Scope	decrypt_policy	Result
wg-supply-chain-ioc (GKE shape only)	DECRYPT_ALLOWED	URL-path filtering active. Aviatrix CA must be in agent image trust store.
All other destinations	DECRYPT_NOT_ALLOWED (explicit)	No MITM. First-party Google traffic (Gemini, Google APIs) is never decrypted.

Domain syntax: exact hostnames or leading wildcards (*.example.com). Bare * is rejected by the controller. Do not widen the decryption scope without re-validating certificate pinning behavior on the new destinations.

Known Constraints

Constraint	Workaround / Status
Managed runtime TLS decryption not available in v1	Google manages the container image; no verified CA injection path today. Managed shape runs SNI and domain policy only.
A2A multi-agent fan-out not covered in v1	Not yet documented whether inter-agent hops traverse PSC-I or Google backbone. Agent Gateway is Google's governance point today.
GKE CRD-based DCF policy not yet end-to-end tested	GKE shape leads with SmartGroup-based policy. CRD-based policy deferred.
Certificate pinning on decrypted destinations	Decryption is scoped tightly by design. Re-validate cert behavior before adding destinations to wg-supply-chain-ioc.

Appendix: GCP-Specific Prerequisites

The following GCP resources must be in place before running terraform apply:

- PSC-I network attachment in the agent VPC, with the network-attachment subnet CIDR assigned to sg-agent-psc-subnet.
- Private Service Connect endpoints for aiplatform.googleapis.com and googleapis.com in the workload VPC.
- Cloud DNS private zone peered to the workload VPC, resolving *.googleapis.com and *.aiplatform.googleapis.com to the PSC endpoint IPs.
- VPC Flow Logs enabled on the agent VPC and the network-attachment subnet.
- Aviatrix gateway deployed in the workload spoke VPC, with the network-attachment subnet routing default via the gateway IP (satisfies Google RFC 1918 egress next-hop requirement under VPC Service Controls).
- For GKE shape: node pool with the agents namespace label applied; Aviatrix CA certificate added to the agent container image trust store.

The full Terraform is at **aviatrix-blueprints/blueprints/gemini-enterprise-agents**. The DCF policy lives in the same repository as the agent config. Adding a tool endpoint is a pull request, not a change ticket to another team.

Aviatrix Validated Containment Architecture for Gemini Enterprise Agent Platform removes the unknown from agentic deployment by enforcing policy at the network layer.

Ask your Aviatrix account team for a guided deployment.

Explore Validated Containment Architectures for other AI platforms.

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 plane. One rule. Universal propagation. Enforced at the workload, not at a chokepoint. Trusted by more than 500 of the world's leading enterprises. For more information, visit aviatrix.ai.