AVIATRIX

SECURITY ARCHITECTURE BRIEF

Contain Azure AI Foundry Agents

Powered by Aviatrix + Microsoft

Threat model, enforcement architecture, and compliance evidence for security architecture review

EXECUTIVE SUMMARY

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 Azure AI Foundry Agents.

Azure AI Foundry's Standard Agent in Bring Your Own Virtual Network (BYOVNet) mode has a security gap: it routes all tool-call egress through a default-allow path with no customer inspection point. Prompt injection, overpermissioned managed identities, and compromised Model Context Protocol (MCP) servers can all exfiltrate data through that gap without touching Azure infrastructure.

The Validated Containment Architecture (VCA) for Azure AI Foundry Agents places every agent behind Aviatrix Distributed Cloud Firewall (DCF) at the Spoke Gateway: default-deny egress, selective Transport Layer Security (TLS) decryption scoped to tool-call traffic, Azure Service Tag-based control-plane bypass, and East-West SmartGroup containment. Controller 9.0 or later required. No code changes to the agent. Reversible in seconds.

Threat Model

Azure AI Foundry BYOVNet mode gives customers network isolation. It does not give them network enforcement. The delegated hosted agent subnet has a no-control egress path to the internet. Three attack vectors are specific to this deployment:

Threat Vector	Architecture Implication
Unrestricted tool-call egress	All outbound traffic from the hosted agent subnet to non-private-endpoint destinations routes through an uncontrolled path. A prompt-injected agent with legitimate managed identity permissions can exfiltrate confidential data to any internet destination.

Control-plane TLS conflict	Applying TLS inspection broadly to the hosted agent subnet path breaks ACA runtime operations – Azure Active Directory (AAD) token acquisition, container image pulls, ACA infrastructure calls depend on unmodified certificate chains. The enforcement solution must selectively decrypt only tool-call traffic (optional) and explicitly bypass decryption for Azure control-plane destinations.
Lateral movement to adjacent spokes	A compromised Foundry agent can attempt to reach databases, internal APIs, and other workload spokes connected via the enterprise transit. Without explicit East-West policy, the agent subnet is not isolated from adjacent resources.

IMPORTANT BACKGROUND

The March 2026 Cascade supply chain operation compromised five major software ecosystems in 12 days using trusted code through trusted pipelines. For one Fortune Global 500 Aviatrix customer, the credential exfiltration attempt was blocked at the network layer by a default-deny spoke-gateway policy before the first packet left the virtual private cloud (VPC). Detection saw the attempt in the logs; the architecture stopped it before the logs were needed. Azure AI Foundry agents face the same structural risk: a prompt injection that passes content-layer controls is still stopped at the network boundary if the exfiltration destination is not in the DCF allowlist.

Attack Scenario and Kill Chain

Scenario: A prompt-injected Azure AI Foundry Hosted Agent is directed to collect confidential SharePoint documents and upload them to an attacker-controlled domain. The agent has legitimate managed identity permissions. The exfiltration traffic is valid HTTPS – indistinguishable from normal tool-call work on the wire.

Kill Chain Stage	What Happens	Control
Prompt injection via MCP tool response	Malicious instruction in an MCP tool response redirects the agent toward data collection and exfiltration. This is indistinguishable from a legitimate tool response at the content layer.	DCF only allows approved MCP FQDNs. Unapproved destinations denied and logged.
Data access	Agent uses legitimate managed identity access to collect documents from SharePoint, Azure Blob Storage, or internal APIs. Valid credentials, authorized channel.	SmartGroup identity-aware policy limits agent subnet access to approved internal service destinations only.

Exfiltration attempt	Agent calls HTTPS POST to attacker-controlled domain. TCP SYN traverses the Aviatrix Spoke Gateway in-path.	Default-deny DCF rule blocks. DENY log in CoPilot FlowIQ with workload identity, destination, rule name, timestamp. Connection never completes.
Lateral movement attempt	Agent attempts to reach adjacent spoke private endpoints – databases, internal APIs – outside the approved tool-call FQDN list.	East-West SmartGroup deny rule blocks traffic from foundry-agents to other-spoke-resources. Structurally prevented before any packet is forwarded.
Audit trail	Incident response team needs evidence for compliance reporting and EU AI Act Article 9 audit.	CoPilot FlowIQ DENY and PERMIT logs – workload, destination, rule citation, timestamp. Continuous evidence from minute one.

POINT OF INTERVENTION

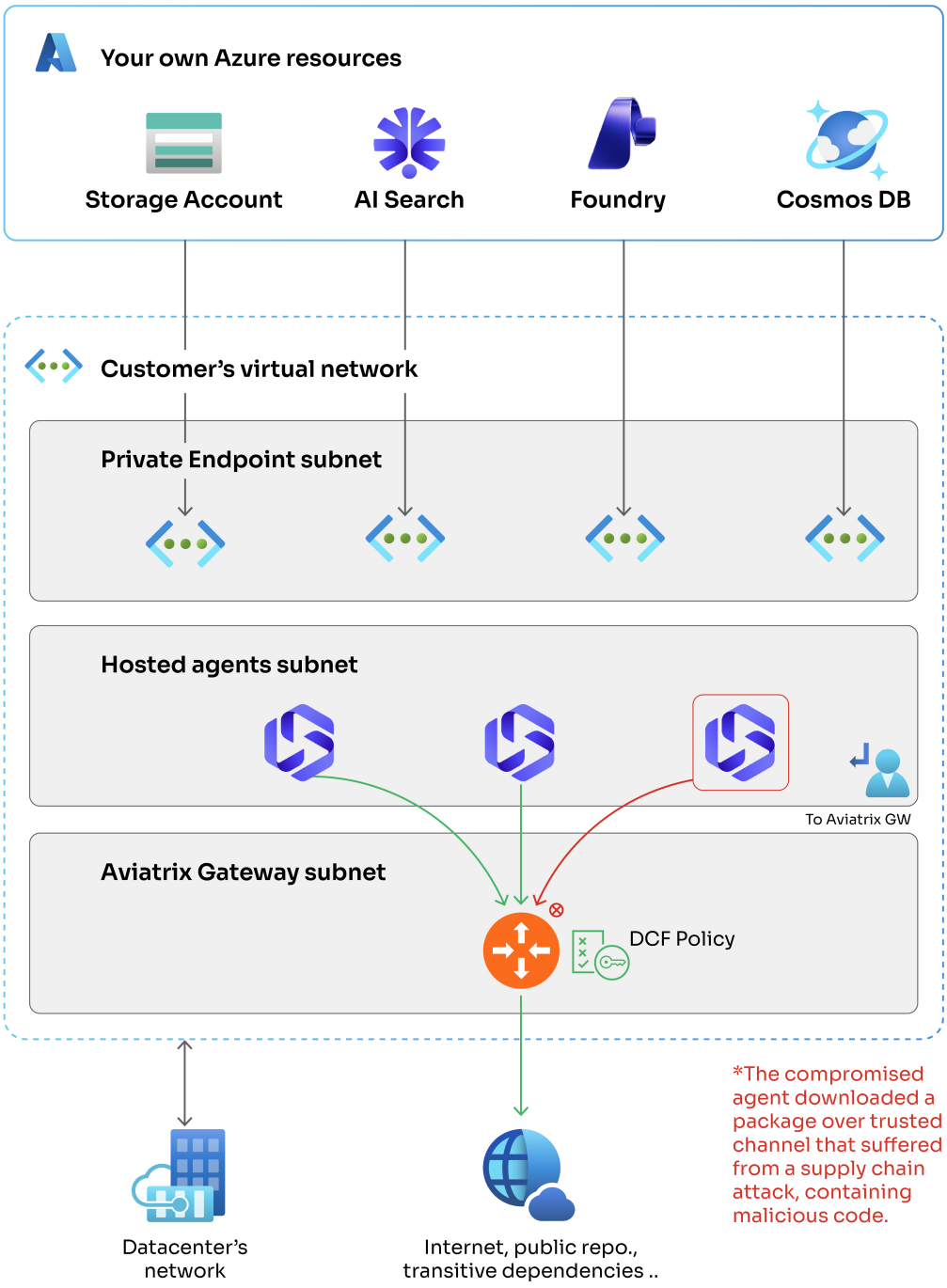
Aviatrix Distributed Cloud Firewall enforces a default-deny policy pack at the Spoke Gateway – in-path for every egress flow from the hosted agent subnet. Tool-call traffic is selectively decrypted and inspected. Azure control-plane traffic is explicitly bypassed via DECRYPT_NOT_ALLOWED rules scoped to Azure Service Tags and ACA runtime FQDNs. East-West SmartGroup policy denies lateral movement. No code changes to the agent container. No sidecar. No agent on the runtime.

Enforcement Architecture

Enforcement runs at three layers. All are transparent to the agent container and require no application changes.

Enforcement Layer	Technical Detail
Payload inspection – prompt content, tool arguments, model responses	Aviatrix Spoke Gateway deployed in the Foundry spoke VNet. A User-Defined Route (UDR) on the hosted agent subnet routes all non-private-endpoint egress through the Spoke Gateway. DCF allows and decrypts (optional) tool-call traffic (external MCP servers, Code Interpreter, REST API calls) via a specific tool-call rule. Azure control-plane traffic decryption is explicitly bypassed: two DCF rules – hosted-agent-requirements- fqdn (WebGroup) and hosted-agent-requirements- svctag (SmartGroups on Azure Service Tags: AzureActiveDirectory, AzureContainerAppsService, MicrosoftContainerRegistry, AzureFrontDoorFirstParty, AzureContainerRegistry) – permit hosted agent platform traffic before the default-deny rule evaluates. The hosted agent runtime never sees a certificate it does not trust.

Layer 2 – East-West lateral movement containment	If attached to a transit, SmartGroup East-West deny rule prevents traffic from the foundry-agents SmartGroup to other-spoke-resources. A compromised agent cannot reach databases, internal APIs, or other workloads in connected VNet spokes – even if those resources are reachable by other enterprise transit spokes.
Layer 3 – Policy as code	The DCF WebGroup allowlist, SmartGroup definitions, and UDR configuration are Terraform-native. They live in the same repository as the agent configuration and update in the same pipeline run. No out-of-band network team coordination. No pressure to over-permit because the change process is slow.



Architecture: Azure AI Foundry BYOVNet with Aviatrix Spoke Gateway in the egress path.

WHY THIS IS DIFFERENT FROM AZURE FIREWALL

Azure Firewall on a stick solves egress for agent on one stack. Enterprises end up running Foundry today, AWS Bedrock AgentCore next quarter, and self-hosted LangGraph on Kubernetes alongside both. Each platform has its own insertion model, firewall requirements, and policy syntax. Aviatrix DCF SmartGroups and WebGroups apply consistently across all three with one control plane, one policy model, and one audit log. A security team can answer "what can any of our agents reach, across all platforms?" from a single console rather than maintaining separate firewall rulesets per agent type. And unlike Azure Firewall, Aviatrix handles multiple agent VNets with overlapping RFC 1918 IP address spaces natively – a common pattern when teams deploy isolated agent environments.

Architectural Boundaries

The VCA governs network reachability for Azure AI Foundry agent workloads. The following are explicitly out of scope:

Out of Scope	What Governs It Instead
Inference traffic (agent to Azure OpenAI private endpoint)	Private endpoint only – RFC 1918, never leaves the VNet, never traverses the Spoke Gateway. Not in scope by design.
Internal resource traffic (Cosmos DB, AI Search, Storage private endpoints)	Private endpoint traffic – stays inside the VNet. Not in scope by design.
Payload inspection (prompt content, tool arguments, model responses)	Azure AI Foundry Guardrails for content-layer enforcement. Aviatrix governs network reachability; Guardrails govern content. Both are required for a complete posture.
Shadow AI discovery (ungoverned Foundry agent workloads)	AgentGuard Shadow AI Discovery – discovers every AI workload via cloud telemetry with no gateway insertion required. Feeds directly into VCA policy targeting.

Compliance Evidence

For HIPAA, PCI-DSS 4.0, SOC 2, EU AI Act, DORA, and CISA Zero Trust Maturity Model environments, auditors require architectural proof of enforcement – not a policy document.

Evidence Artifact	What It Proves and Framework Mapping
CoPilot Distributed Cloud Firewall DENY and PERMIT logs	Continuous, timestamped record of every blocked and permitted egress attempt as needed. Workload identity (SmartGroup), destination FQDN or IP, rule name, timestamp. Maps to PCI-DSS 4.0 Requirement 10.2, SOC 2 CC7.2, and HIPAA §164.312(b). Continuous record of every permitted egress decision. Proves least-privilege intent is architecturally enforced. Required for EU AI Act Article 9 governance audit trails and NIST 800-53 AU-2.
East-West deny logs	Evidence that lateral movement from the Foundry agent subnet to adjacent spoke resources is structurally prevented. Maps to NIST 800-53 SC-7 and CISA Zero Trust Maturity Model Network pillar tier 3.
SmartGroup model with hosted agent identity	Identity-based policy attribution – policy follows workload identity, not IP address. Addresses EU AI Act high-risk system governance requirements and FedRAMP AC-3 access enforcement.
Baseline DCF policy pack (Terraform, version-controlled)	Infrastructure-as-code evidence of network-layer least-privilege – auditable, reproducible, and diffable. Maps to PCI-DSS 4.0 Requirement 1.3, SOC 2 CC6.6, and ISO 27001 A.13.1.3.

The proof of enforcement is the architecture itself – not a statement about the architecture.

Deployment Paths

Two deployment paths accommodate both existing Aviatrix customers and net-new deployments:

- **Existing Aviatrix Enterprise Transit customers:** Attach the Foundry spoke to the existing transit, apply the SmartGroups and WebGroups from the blueprint, and activate the policy pack. Same control plane, same console, same policy model as every other spoke in the fabric.
- **New customers or standalone Foundry deployments:** Deploy a standalone Aviatrix Spoke Gateway inside the Foundry VNet. One UDR on the Hosted agent subnet, no hub peering, no on-premises connectivity required. The standalone gateway connects to Enterprise Transit later if the network footprint grows.

Next Steps

The Validated Containment Architecture for Azure AI Foundry Agents is available today. The Terraform blueprint, scenario user interface, and deployment guide ship together. Aviatrix Controller 9.0 or later is required for selective TLS decryption.

Request a 30-minute architecture review.

Walk through the enforcement model in your environment, map your current Foundry agent inventory, and identify 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 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.