NetFoundry

SOLUTION BRIEF

NetFoundry Zero Trust AI Enclaves

Secure and Accelerate Agentic AI Deployments Without Infrastructure Delays

Accelerate AI Innovation Without Increasing Risk

Enterprises are under increasing pressure to deploy agentic AI solutions. Application teams focus on speed—deploying AI agents, integrating large language models (LLMs), and connecting distributed services across cloud, edge, and partner environments. Yet every update becomes a taxing, recurring burden. Every new agent, tool, model, API, or data path can trigger another round of firewall, routing, NAT, VPN, private-link, DNS, approval, and exception work.

Meanwhile, there's a growing disconnect between the urgency of deployment and the realities of securing these new architectures. Security teams must ensure these deployments do not introduce new vulnerabilities, expand the attack surface, or bypass governance controls.
The result is a familiar tension: AI projects are slowed by infrastructure requirements, security reviews, and network changes, while security teams struggle to maintain consistent control over increasingly dynamic environments.

Traditional approaches to connectivity—relying on VPNs, firewall rules, IP allowlists, and exposed APIs—were not designed for the distributed, service-to-service nature of modern AI systems. These approaches introduce operational friction, delay deployments, and create new pathways for attackers. They also rest upon an increasingly risk assumption—that it's safe to allow network reachability before identity and policy determine whether or not a connection should exist.

NetFoundry Zero Trust AI Enclaves solve this by helping organizations deploy AI workloads quickly while maintaining strong, consistent security and governance.

THE AI DEPLOYMENT CHALLENGE

Agentic AI systems are inherently distributed and dynamic. AI agents interact with LLMs, MCP servers, APIs, and data services that may reside across multiple clouds, on-prem environments, or third-party platforms. These interactions evolve continuously as new agents, tools, and integrations are introduced.

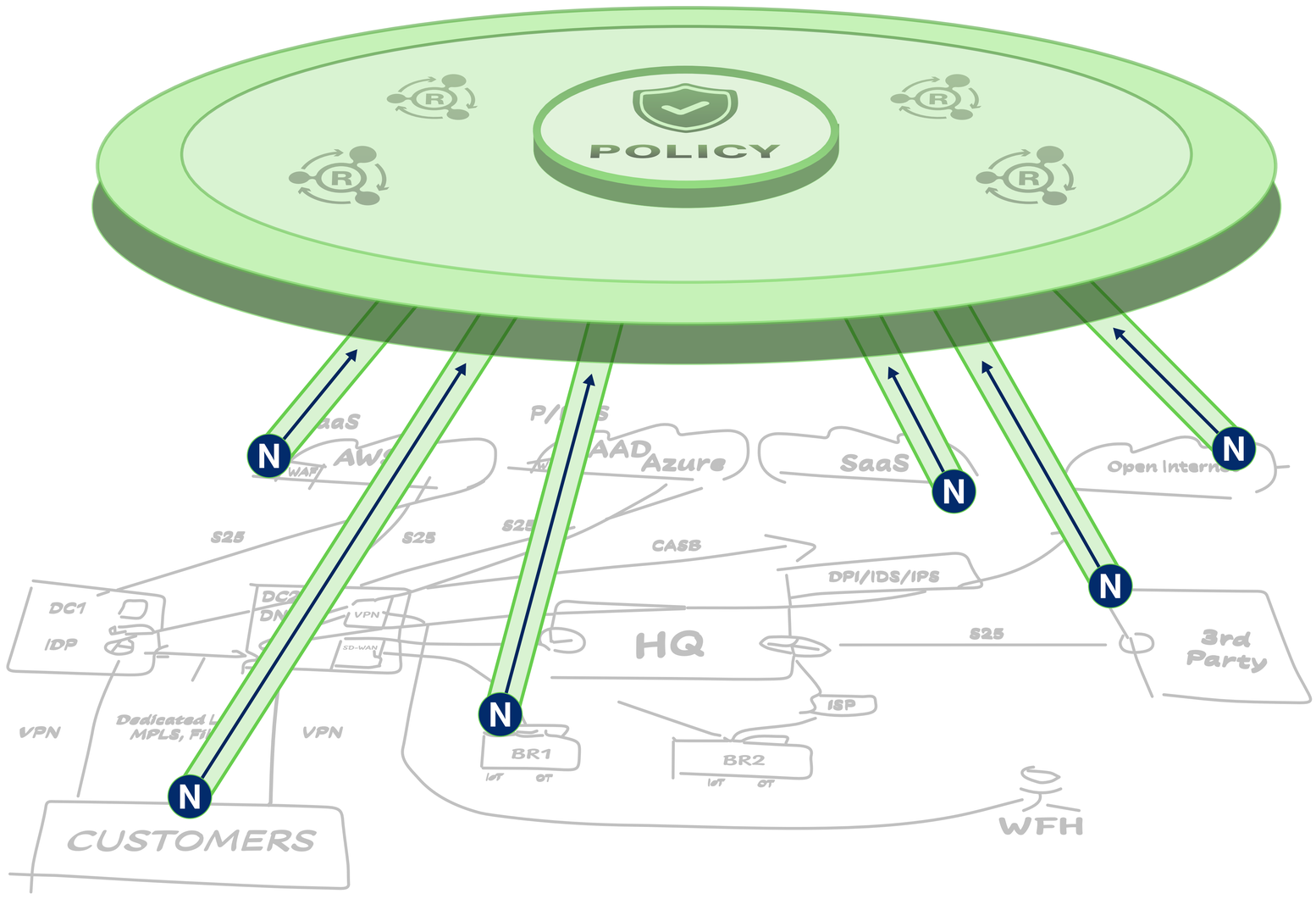
Unfortunately, increased complexity is often addressed via "Shadow AI". Roughly 68% of employees are using unauthorized tools they brought in themselves because the official rollout was too slow.

And beyond complexity, a more fundamental issue is limiting organizations' ability to scale AI securely: the way machine identities are managed today.

THE NETFOUNDRY SOLUTION

Zero Trust AI Enclaves

The NetFoundry solution avoids the limitations of firewalls and VPNs by implementing virtual connectivity at layer 7. It implements an AI enclave perfectly suited for AI connectivity. This enclave is independent of the underly network and invisible to anyone who's not authenticated. Unlike the traditional model, the enclave performs authentication and authorization before connection. Until then, no routable path exists.



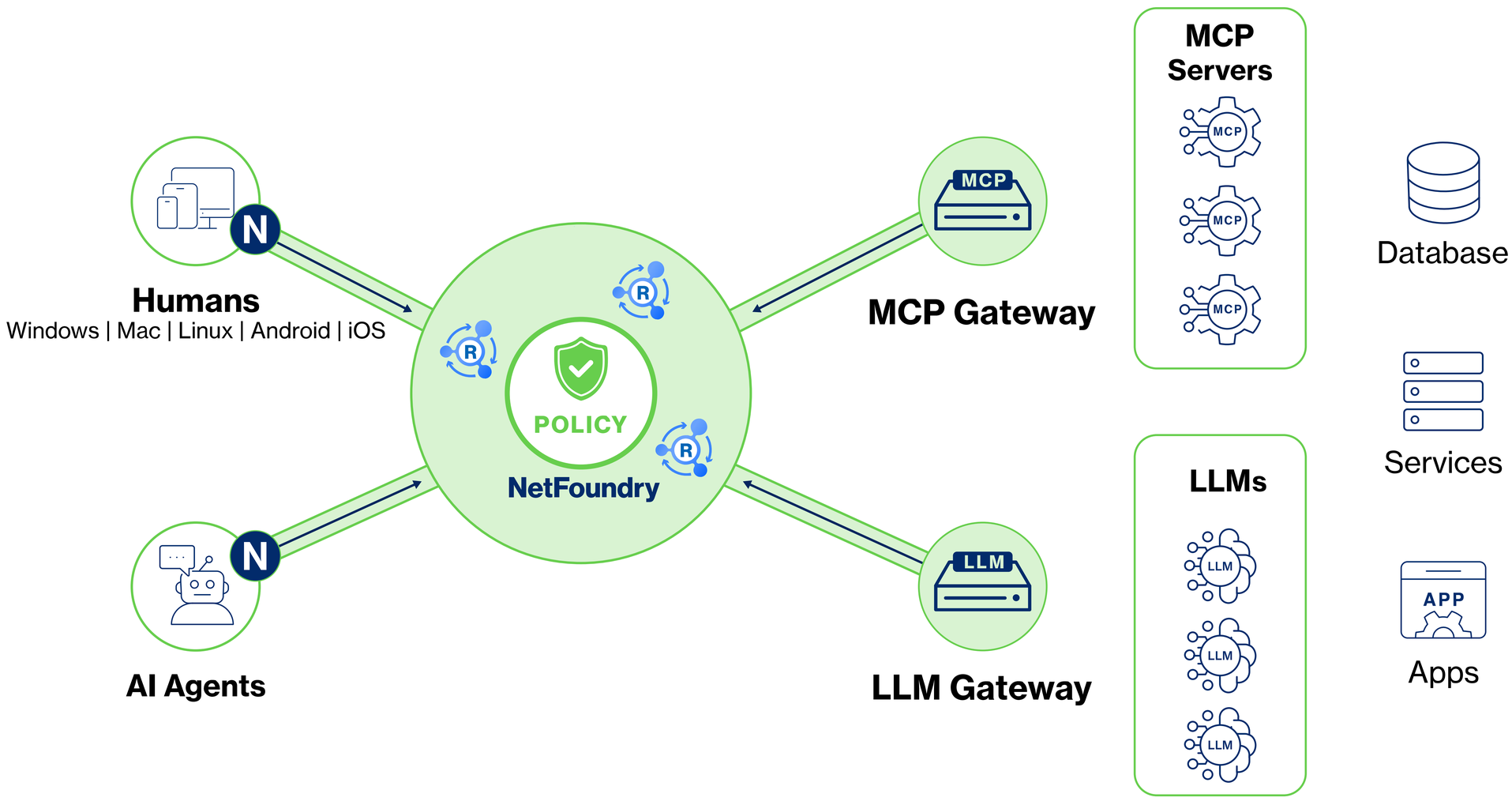
Uniquely Designed to Accelerate and Secure AI Deployments

- Every element has a **unique identity**, as you would with a human being, with each machine identity receiving a certificate. The certificate is bound to the particular app or server, so it can't simply be stolen and reused by an attacker.
- Scoping can be established by **site, machine, or individual application component**. In fact, everything needed to connect with the overlay network can be compiled directly into a container or an application component via an API.
- Connections are established virtually, with **full end-to-end encryption**, and a separate layer 7 routing infrastructure.
- From the underlying network's perspective, **all connections are outbound**—so no changes to firewalls are required. All inbound ports can be closed with a "default deny" approach, resulting in far better security and no operational updates.
- Connectivity is **authorized by policy with least privileged access**, and it can be enabled instantly and programmatically.
- Management and visibility is **by identity, not IP address**, for better control, faster problem resolution, and deep insight. Access can be instantly granted or revoked, solely based upon identity. There's no dependency on infrastructure, networks, or integrations.

AI Connectivity Challenges	Benefits of Zero Trust AI Enclaves
Deployments slowed by network change requirements	Accelerates deployments by using only outbound connections. No changes are required to VPNs, firewalls, or other network setup.
Intrusion risk from expanded attack surface	The AI Enclave is completely private and invisible from the outside. Each connection is authenticated before it's established, and all connectivity is controlled via policy. If identity and policy do not authorize the interaction, no routable service path is created.
Poor visibility into connectivity across multiple networks	All traffic is associated with an identity rather than an IP address, allowing for meaningful understanding of connectivity and traffic.

Secure and Govern AI Interactions at Scale

A key advantage of the Zero Trust AI Enclave is the ability to centralize control and governance across all AI communications while eliminating the risks associated with secrets-based access.



The LLM Gateway provides a secure control point for managing interactions with external and internal LLMs. Organizations can route requests across multiple models, enforce usage policies, and implement load balancing or failover strategies—all without exposing endpoints or distributing API keys.

The MCP Gateway similarly enables secure, standardized access to MCP servers and services using identity-based authentication. Agents can only discover and invoke MCP servers and tools that are inside their policy scope. This approach removes the need for AI agents to have access to secrets, reducing operational overhead and eliminating a major source of risk.

Because all access is identity-driven, organizations gain full visibility into AI activity. Token usage can be tracked for cost management and governance, and policies can be enforced consistently across environments.

Key AI Gateway Functionality

Semantic routing

To determine if a query requires a public model (e.g., GPT-4o) or a self-hosted private model, optimizing for cost and privacy.

MCP protocol inspection

To parse MCP traffic and enforce rate limits on tool usage (e.g., “Max 5 SQL queries per minute”).

AI Firewall

To sanitize inputs for PII before they enter the secure network, acting as a specialized semantic filter rather than a generic packet inspector.

Cost tracking

Budget, limit, and track the dollar cost usage of expensive AI assets by team and project.

Operational Simplicity and Agility

By eliminating reliance on secrets and network-based controls, NetFoundry dramatically simplifies the ongoing operation of AI systems. Policies are defined centrally using identity-based rules and can be updated programmatically. Changes take effect instantly, enabling organizations to respond quickly to evolving requirements. Access can be granted on a just-in-time basis or revoked immediately, providing precise control over AI agent behavior.

Management of the AI Enclave is based upon identity, not IP address. This is true as well for *visibility*, which greatly simplifies tasks such as troubleshooting.

Furthermore, NetFoundry offers several deployment options to meet the needs of your AI solution.

OpenZiti

- Community support
- Self deployed and managed (connectors, routers, controller), self orchestrated

Self-Hosted

- Enterprise grade support (24x7)
- Self deployed and managed (connectors, routers, controller), self orchestrated
- Guidance for resilient, reliable, scalable production architecture & deployment
- Operations, logging, assurance and platform LCM tools
- Contracted relationship (indemnification)
- Use case and/or implementation documentation

Cloud

- Enterprise grade support (24x7)
- Fully managed by NetFoundry with 99.95% uptime SLA
- Guidance for resilient, reliable, scalable production architecture & deployment
- Fully automated LCM of hosted and LCM support for self-hosted components
- Contracted relationship (indemnification)
- FIPS compliant
- Use case and/or implementation documentation
- SOC 2 Type II compliant

NetFoundry – The Leader in Zero Trust AI Enclaves

NetFoundry was founded by the inventors and maintainers of OpenZiti—the world’s most-used open source solution for zero trust connectivity. NetFoundry helps organizations secure connectivity between components of widely distributed workloads across any environment. NetFoundry’s Zero Trust AI Enclaves enable organizations to deploy AI workloads quickly while maintaining strong, consistent security and governance.

Learn more at www.netfoundry.io

NetFoundry