Technical white paper

GreenLake Intelligence

The agentic AI architecture for autonomous hybrid cloud operations

Contents

Executive summary

Enterprise IT has reached a tipping point. As organizations scale across hybrid, multicloud, and edge environments, the complexity of managing infrastructure has outpaced human capacity. Traditional operation models—built around static dashboards, disconnected tools, and reactive processes—can no longer keep up with the demands of modern digital business.

GreenLake Intelligence introduces a new paradigm: agentic AI architecture purpose-built for autonomous enterprise-scale operations. It's not just AI-ready, it's AI-operational. GreenLake Intelligence enables a mesh of intelligent agents that collaborate, learn, reason, and act in real time across your hybrid estate. These agents form a distributed system that continuously drives optimization, enforces policy, troubleshoots proactively, and augments human teams. It combines:

– A mesh of **expert agents** operating on real-world enterprise telemetry, configs, and support data
– **Unique reasoning capabilities** from our multivendor, multicloud full-stack context
– **Human-in-the-loop with integrated copilots** for natural language interaction, orchestration, and control

This white paper provides details on GreenLake Intelligence: its architecture, how it works, what it enables, and why it represents a foundational shift in how enterprises operate their infrastructure—not just with automation, but with intelligence.

1. The need for agentic AI in enterprise IT

Most enterprises operate a patchwork of environments, tools, and workflows across data centers, clouds, edge sites, and SaaS platforms. While digital transformation has accelerated, operations have lagged behind and are still largely dependent on:

– Manual processes for provisioning, troubleshooting, and compliance
– Disconnected monitoring and observability tools
– Reactive alerts and siloed dashboards
– Human-driven coordination across IT teams

This fragmented state creates four persistent challenges:

1. **Slow, manual workflows** that bottleneck innovation and increase risk
2. **Delayed troubleshooting** that leads to downtime and business disruption
3. **Wasted resources and missed optimization** across the hybrid estate
4. **Overstretched IT teams** lacking the scale or expertise to manage modern environments

At the same time, new business demands—from real-time analytics to AI model hosting—require **operations that are faster, more autonomous, and more reliable than ever**.

2. Why traditional AIOps and automation fall short

Agentic AIOps represents the next evolution of IT operations by introducing autonomous, goal-driven agents that act with greater intelligence, adaptability, and decision-making capabilities compared with traditional AIOps. While traditional AIOps focuses on monitoring, analyzing, and automating IT operations through machine learning and pattern recognition, agentic AIOps takes it a step further by incorporating **autonomy, contextual awareness, and proactive decision-making**.

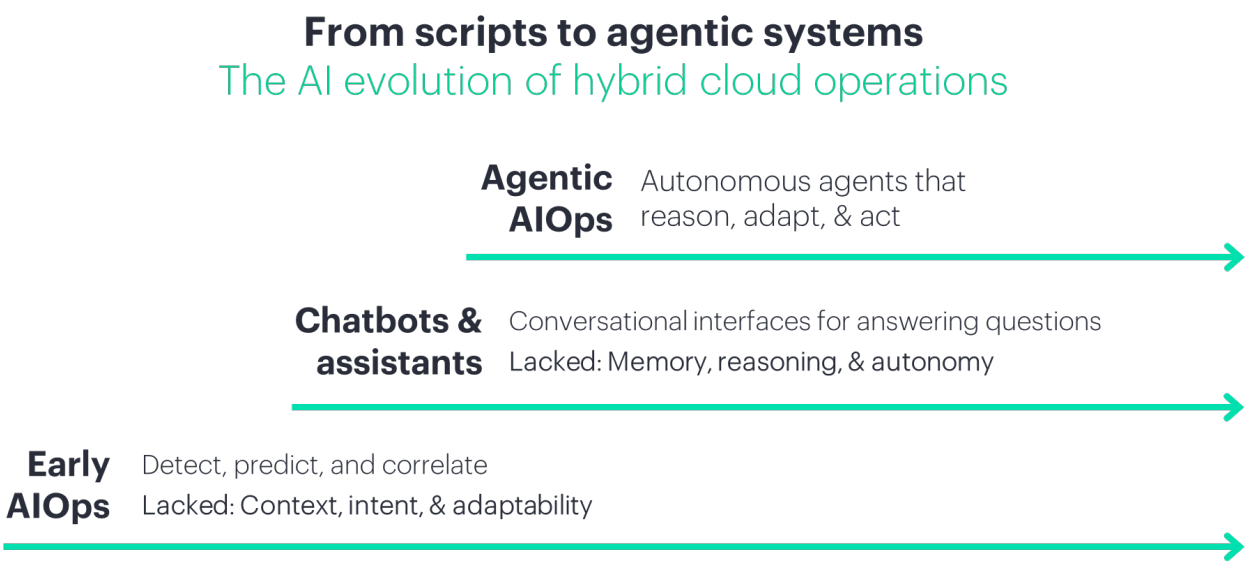


Figure 1. Evolution of AIOps approaches

Traditional AIOps typically relies on predefined workflows, rules, and historical data to identify and respond to anomalies. In contrast, agentic AIOps leverages **intelligent agents** that can dynamically assess the IT environment, adapt to real-time changes, and align actions with organizational goals. These agents are capable of **self-learning**, prioritizing tasks, and collaborating across systems to optimize performance and resilience.

3. GreenLake Intelligence design points

By combining the hybrid cloud flexibility of GreenLake cloud with the **adaptive intelligence** of agentic AI, organizations can operationalize collective intelligence at scale to build human AI systems—without sacrificing trust, transparency, or control.

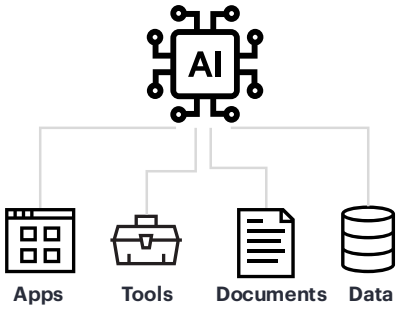
Here's how it comes together:

Human-in-the-loop
Human input is not an afterthought—it's a **structured component of context**. Analysts, engineers, and domain experts can define high-level goals, constraints, and escalation protocols that are interpreted by AI agents during decision cycles.

Interagent communication
GreenLake Intelligence uses open industry protocols to standardize the language of collaboration between distributed agents, whether Hewlett Packard Enterprise or third party, such as ServiceNow, Salesforce, PagerDuty, and such. For example, a model detecting anomalous behavior in financial data can request enhanced resources from another agent responsible for dynamic workload orchestration—creating an emergent system of shared intelligence.

Agentic AI with contextual autonomy
GreenLake Intelligence is designed to give agents the situational awareness they need to make intelligent decisions based on real-time infrastructure metrics and business logic. Interagent communications, based on MCP and other protocols, is critical to the operation and scalability of the system. Whether running on an

What is Model Context Protocol?
Model Context Protocol (MCP) is a framework that ensures AI models operate effectively by providing task-specific context, such as input parameters, user intent, or environmental data. It helps models adapt dynamically, align outputs with objectives, and maintain relevance, improving accuracy, transparency, and reliability in AI-driven workflows.



HPE Alletra Storage MP array, an engineered private cloud system in a data center, or at the edge, agents operate with functional awareness of their environments.

Infrastructure-aware models
With HPE and third-party infrastructure exposed to GreenLake Intelligence, agents can proactively request compute, storage, or network resources or adjust fidelity based on resource availability. This reduces cost, optimizes performance, and helps ensure SLA alignment.

Adaptable goal-driven architecture
Agentic systems are driven by goals, but any system needs to be designed with flexibility to adapt to known goals today, and future goals that may not yet be known. Different goals may require the creation of different expert agents.

Why this matters
As AI becomes more central to mission-critical operations, enterprises need more than model performance—they need model responsibility, adaptability, and explainability. By leveraging a standards-based approach designed to scale as the industry evolves, GreenLake Intelligence helps ensure that models don't operate in isolation but as part of a **human-aligned, infrastructure-aware ecosystem**. By leveraging GreenLake, organizations gain:

- **Real-time collaboration** between AI and humans
- **Elastic, composable infrastructure** aware of workload needs
- **Scalable orchestration of distributed AI agents**
- **Built-in transparency and oversight** to meet regulatory and ethical requirements

4. Inside the GreenLake Intelligence agentic AI architecture

GreenLake Intelligence brings together a number of architectural components that are critical to the application of agentic AI to hybrid cloud operations:

- **Expert agents:** Software systems based on LLMs fed with enterprise infrastructure context (telemetry, logs, configurations). Agents are continuously active and autonomous.
- **Reasoning and agent orchestration:** The system that orchestrates the **chain-of-thought** interaction across HPE and third-party agents. Agent action may be event-initiated or human-initiated. Shared context, memory, and intent across agents are enabled.
- **Adaptive copilots:** Human-in-the-loop is enabled through a copilot interface system in GreenLake, allowing for interaction in human-friendly formats. Human operators define the level of autonomy granted to the agentic system. Copilot interactions may be optimized with the benefit of situational context for precision to the task at hand.
- **Governance and guardrails:** Identity- and context-aware authentication of agent objectives and interactions.

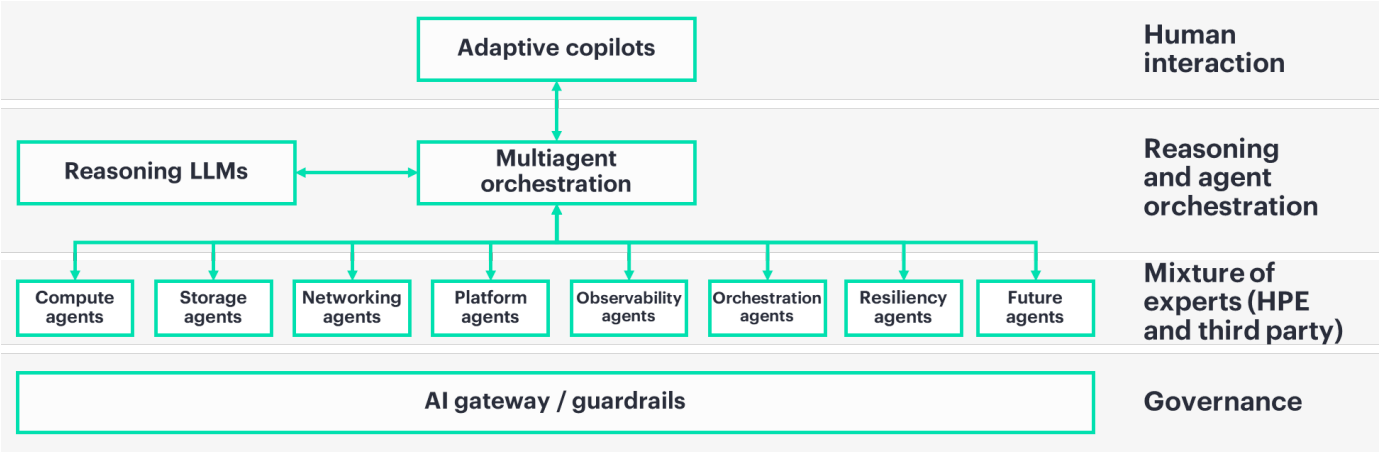


Figure 2. GreenLake Intelligence architecture

5. Mixture of experts for goal flexibility

Key to the architecture of GreenLake Intelligence is the orchestration of expert agents within and across domains. Expert agents are based on LLMs suited to a particular task category and fed with data specific to the domain.

Because GreenLake is a hybrid cloud platform, domains cover areas of infrastructure, runtimes, applications/workloads, operations, and more. Expert agents operate on goals, and different goals may require the creation of different expert agents that work within the orchestrated system.

By implementing a system of orchestrated expert agents, GreenLake Intelligence can deliver value based on deep expertise; coverage across day-0, day-1, and day-2 scenarios; and adaptability to future use cases.

The breadth of GreenLake, coupled with the flexibility of the multidomain agentic orchestration in GreenLake Intelligence, allows for an extensive set of practical applications. Initial priority use cases include:

- **Streamlined service delivery:** Autonomous provisioning and balancing of workloads based on complex multivariate parameters of cost, performance, and security
- **Workload and platform optimization:** AI-driven workload placement, sustainability tuning, and capacity forecasting
- **Proactive troubleshooting:** Agents coordinating real-time root-cause analysis with embedded observability

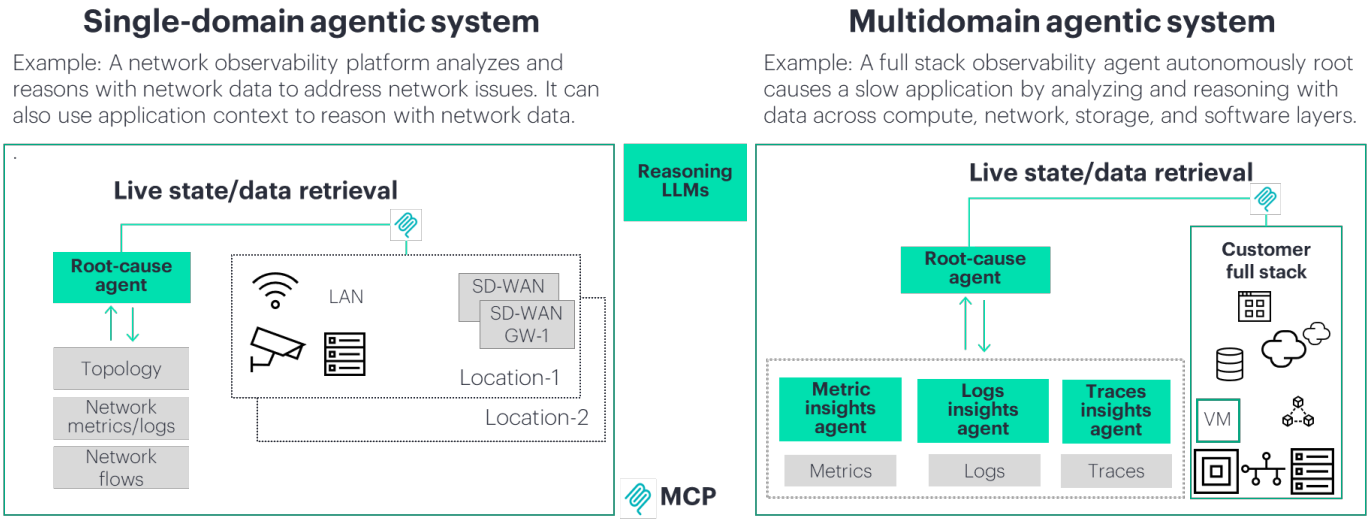


Figure 3. Single-domain and multidomain use case examples

6. Human-in-the-loop through adaptive copilots

As agentic systems take on the heavy lifting of managing complex, multidomain environments, the need for an intuitive and collaborative interface has never been more critical.

A copilot acts as the connective tissue between human operators and autonomous systems, enabling natural-language interactions, contextual insights, and actionable decision-making. At the multidomain level, a copilot provides a unified view across the entire hybrid estate, empowering users to oversee and guide agentic systems with clarity and ease.

But the value of a copilot doesn't stop there—its impact extends deeply into member applications. Within these apps, embedded copilots can address domain-specific challenges, such as network optimization, real-time troubleshooting, or workload orchestration.

By tailoring copilot capabilities to individual applications while integrating seamlessly with the broader system, copilots enable precise control and collaboration. This dual-layered approach ensures that the overall hybrid environment and specific, granular operations both benefit from intelligent, human-aligned interaction, redefining how organizations engage with agentic systems.

7. In-depth review of a use case: Network troubleshooting

AIOps has been in use for networking for many years. The combination of complex, widespread topologies, device number and diversity, speed of change, and the immediacy and consequence of end-user impact has made AIOps both a necessity and a benefit in the networking space.

Networking AIOps has evolved from anomaly detection and predictions to include generative AI to enable optimized search and document summarization. Agentic AI adds autonomous reasoning powered by rich context to enable **rapid-precision, root-cause identification, and remediation**.

HPE Aruba Networking Central—a key set of services on GreenLake—leverages GreenLake Intelligence to provide advanced reasoning for complex network/security conditions and insights.

Operators interact through a multimodal conversational copilot optimized for security-first, AI-powered networking. An array of domain-trained and coordinated agents enables precise root-cause analysis with suggested or automated remediation for many networking conditions.

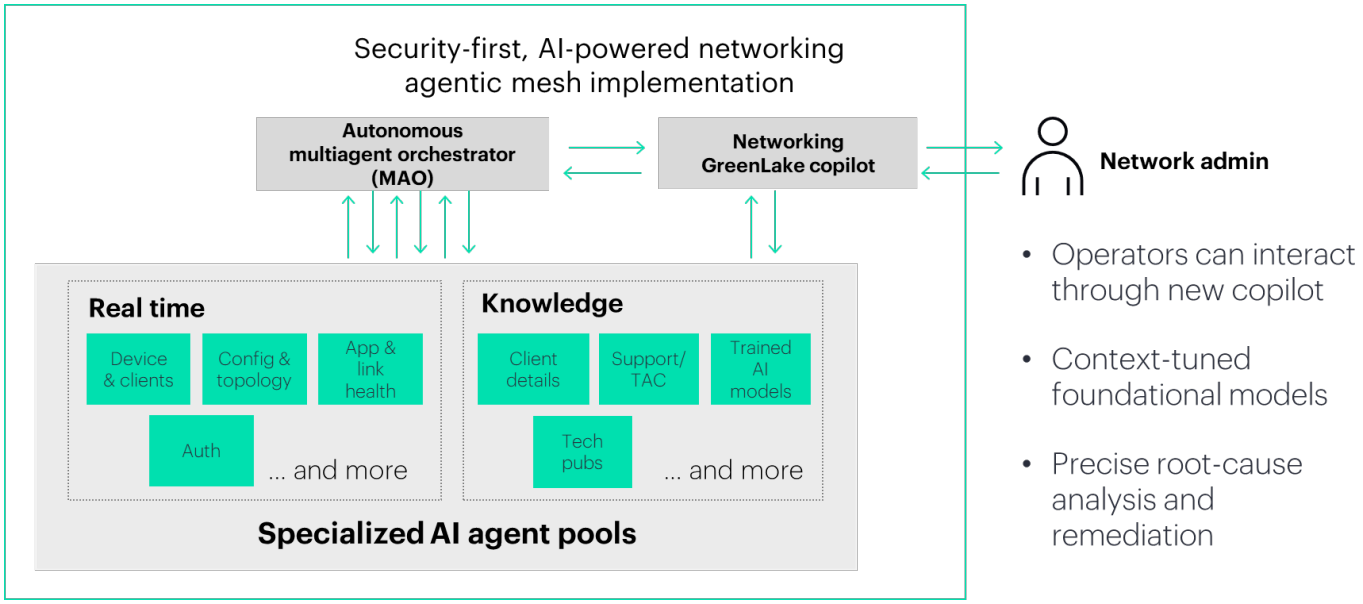


Figure 4. Multiagent flow and user interaction in networking

Let's examine how this flow works in an actual example (Figure 5). In this case, unacceptable latency impacting service quality for Slack users has been detected.

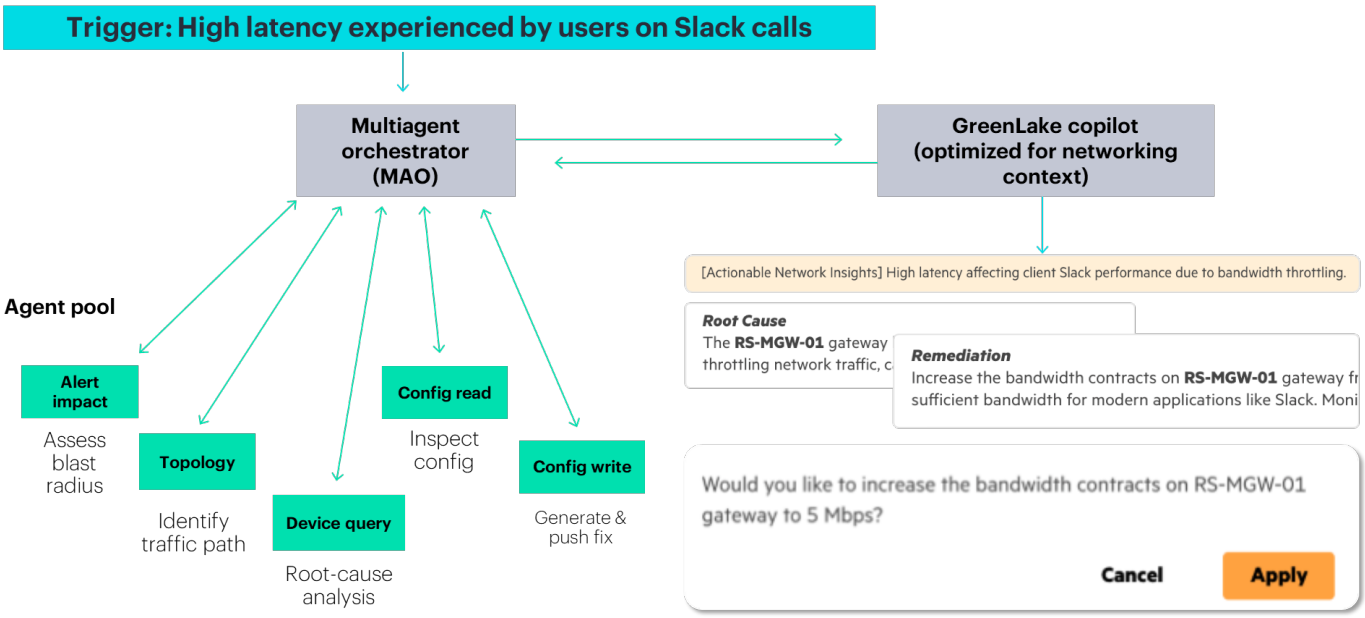


Figure 5. Network latency agent flow and copilot interaction

The orchestrator leverages several expert agents to handle the issue:

1. **Alert impact agent** identified the blast radius—Five clients at RS05 site are experiencing a median latency of 101 ms.
2. **Topology agent** understands the switches and gateway that are in the path when clients access Slack.
3. **Device query agent** dived into root-cause analysis: 38 contract lookup failures, indicating recurring issues with bandwidth contract resolution on the gateway. Verified that the switches in the path (CORE1 and AGG1) show no congestion issues, mitigating them as potential causes.
4. **Config read agent** extracted the gateway settings and gateway configuration and discovered two bandwidth contracts ("aloha-ssiddownstreamper-userui" and "aloha-ssidupstreamper-userui") set at only 512 Kbps each.
5. **Config write agent** applied the fix in real time, increasing the bandwidth contracts to at least 5 Mbps would provide sufficient capacity for Slack and similar applications.

In this environment, the system is set up to ask for operator approval of recommendations through a simple OK in the copilot. Systems can be configured to act with greater or lesser independence, based on enterprise policies.

8. Extensibility, ecosystem, and partner value

Enabled by GreenLake, GreenLake Intelligence benefits from the same security, API frameworks, and validated partner ecosystem that are already a mature part of the GreenLake ecosystem.

GreenLake APIs provide a robust and flexible way to interact programmatically with HPE cloud services and infrastructure-as-a-service (IaaS) offerings. They enable developers and IT teams to automate operations, integrate GreenLake services with existing workflows, and manage hybrid cloud environments efficiently. These APIs are part of GreenLake, designed to support multicloud, edge, and on-premises environments.

GreenLake APIs use RESTful principles, allowing for simple and secure communication between applications through HTTP. Through these APIs, users can provision and manage resources, monitor usage and performance, and access consumption analytics. They offer granular control over workloads, enabling tasks like deploying virtual machines, configuring storage, or scaling resources based on demand. Security is a priority, with API authentication typically using OAuth 2.0 and role-based access control (RBAC) to verify that only authorized actions are performed.

This delivers flexibility to leverage partner capability in GreenLake Intelligence or enable ISVs and developers to build their agents or copilots in areas such as security, data protection, operations, and more.

9. Recap

GreenLake Intelligence represents a transformative shift in enterprise IT operations, enabling a move beyond traditional AIOps. By integrating intelligent, autonomous agents with real-time collaboration, proactive troubleshooting, and infrastructure-aware adaptability, it delivers unparalleled operational efficiency, scalability, and resilience.

With humans-in-the-loop and governance at its core, GreenLake Intelligence offers transparency, trust, and alignment with organizational goals. As enterprises face increasing complexity across hybrid, multicloud, and edge environments, GreenLake Intelligence offers a future-ready solution that combines innovation, adaptability, and intelligence to meet the demands of modern digital business.

Learn more at

HPE.com/us/en/GreenLake/Intelligence.html

Visit HPE.com

Chat now

a50013545ENW

HEWLETT PACKARD ENTERPRISE

HPE.com