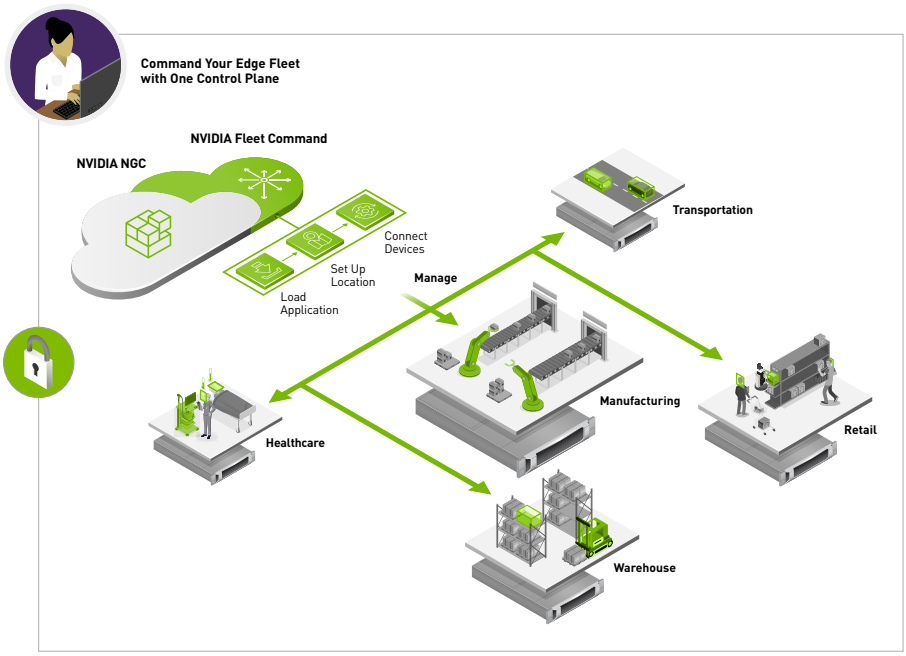
SECURELY DEPLOY AND MANAGE AI AT THE EDGE

WITH NVIDIA FLEET COMMAND

AI at the edge delivers powerful, real-time insights to organizations looking to drive actionable results at lower costs. With lower latency than cloud platforms, it's quickly becoming the go-to AI infrastructure for cutting-edge organizations. But securely deploying, managing, and scaling AI across the edge is complex and challenging. NVIDIA Fleet Command™ combines the security and real-time processing capabilities of edge computing with the remote management and ease of software as a service.

Organizations are deploying edge AI to address the limitations of cloud-only platforms like latency and cost of scaling. However, edge computing presents its own set of challenges. For enterprises gathering insights from many separate locations, installing hardware, deploying software, and maintaining upgrades at each individual location is both time consuming and costly. Additionally, the skilled IT personnel needed to execute these tasks, along with on-call support, can be difficult to arrange for numerous remote or geographically dispersed locations.

NVIDIA Fleet Command is a hybrid-cloud platform for managing and scaling AI deployments across dozens or up to millions of servers and edge devices. Fleet Command allows IT departments to securely and remotely manage a large-scale fleet of deployed systems. Instead of spending weeks planning and executing deployment plans, administrators can bring AI to networks of retail stores, warehouses, hospitals, or city street intersections within minutes. Administrators can add or delete applications, update system software over the air, and monitor the health of devices spread across vast distances from a single control plane.



NVIDIA Fleet Command is a centralized control plane that operates in the cloud. Using one-touch provisioning, Fleet Command pairs with NVIDIA EGX™ servers with ease. Once paired, Fleet Command securely deploys, manages, and scales AI applications across the entire infrastructure in minutes. A resilient software stack means servers self-heal when applications are disrupted, and detailed monitoring dashboards provide continuous insight into application and hardware health.

Effortless Management

Fleet Command's streamlined provisioning, detailed monitoring dashboards, and extensive automated processes allow IT, DevOps, and system administrators of any skill level to manage the deployment and maintenance of AI at the edge.

Centralized, remote management of AI deployments enables over-the-air software updates, remote debugging, and system monitoring, as well as other features such as self-healing systems. These remote management features make maintenance and upkeep easier and AI more accessible and practical for locations that are difficult to access or far from headquarters—which delivers faster, more comprehensive insights that can drive real-time decisions.

Secure Deployment

Attention to security at the edge is essential for ensuring that customer intellectual property and data are protected. However, manual edge deployment and command line interface (CLI)-based edge management tools come with little or no built-in security measures.

NVIDIA Fleet Command has built-in, end-to-end security to ensure intellectual property and application and sensor data are always protected.

It starts with the application. Applications are scanned for vulnerabilities and malware before they're loaded. In addition, signed containers ensure that only authenticated software is deployed to the edge.

At the location, all processed data is encrypted at rest, and the AI runtime is protected from tampering with secure and measured boot. And because systems are on premises, organizations maintain control of where sensor data is stored.

Resilient Software

With many devices at many edge locations, organizations need resilient software to streamline management. Without resiliency, organizations face the incredible challenge of having to manually repair remote systems when applications fail, leading to long periods of downtime when an application isn't running and insights aren't being generated.

Fleet Command features a resilient software stack that allows all systems to self-heal when applications are disrupted. But resiliency means more than infrastructure that can restart applications. It also means that workloads can be migrated to reduce application downtime. If a system is disrupted at one location, Fleet Command can migrate the workload of that system to another on the same network, ensuring real-time insights are never lost.

Power Change with Edge AI

NVIDIA Fleet Command makes it easy to manage, scale, and keep your entire AI deployment running smoothly. From dozens of edge devices to millions, IT admins can take secure, remote control of the fleet, simplifying deployment and powering resilient AI across a network in mere minutes.

To learn more about NVIDIA Fleet Command, visit **www.nvidia.com/fleetcommand**

NVIDIA.