



A High AI-Q[™]
Company



Edge AI Platform Deployment with Secure, Scalable Kubernetes and GitOps

Re-architecting cloud-native AI applications for rugged, disconnected environments using a resilient pull-based deployment model and Kubernetes-native management.

Overview

We designed and implemented a secure, GitOps-driven platform using ArgoCD and Kubernetes to deploy AI applications across a fleet of geographically distributed edge units.

- Achieved a 95% deployment success rate in intermittently connected environments by transitioning to a resilient pull-based model.
- Reduced operational overhead by 70% through automated dependency management using Kubernetes operators and dynamic Helm configuration.



Client Profile

This US-based startup operates at the frontier of edge computing, uniting compute and real-world AI where data is generated. Their platform is purpose-built for logistics and defense industries, functioning in remote environments where cloud connectivity is often unavailable or highly restricted.

Cloud Dependencies and Connectivity Challenges

Transitioning AI workloads from the cloud to the rugged edge introduced significant technical friction:

- **Service Dependencies:** Heavy reliance on cloud-managed storage and databases required re-engineering for edge-native, self-hosted alternatives.
- **Pipeline Limitations:** Traditional push-based CI/CD pipelines failed due to restricted inbound access and unstable network connections at edge sites.
- **Scalability Barriers:** The need to support a diverse, heterogeneous fleet of hardware while maintaining a unified and consistent operational model.

- **Security & Observability:** Restricted inbound connectivity complicated centralized troubleshooting and necessitated zero-trust outbound telemetry.

QBurst Solution: GitOps-Driven Edge Orchestration

We implemented a secure, GitOps-based architecture that fundamentally shifted the deployment model from "push" to "pull." Key cloud services were re-architected into an edge-native model, with dependencies managed directly through Kubernetes operators.

The solution features a decoupled, resilient architecture:

- **Pull-Based Deployment:** We utilized ArgoCD to automatically pull versioned Helm charts to edge clusters, ensuring that updates are applied whenever connectivity is available, with built-in retry logic.
- **Edge-Native Dependencies:** Core services like Ceph for storage and various databases were deployed via operators to ensure reliable real-time operations at the source.
- **Resource Optimization:** Large AI images are cached in local edge registries to accelerate the rollout of resource-intensive GPU workloads and minimize bandwidth consumption.
- **Localized Intelligence:** The platform performs data filtering and aggregation locally, transmitting only essential metadata to the cloud to reduce storage costs.

Technical Highlights

- **Automated Configuration Engine:** Dynamically generates Helm values based on specific edge compute profiles, reducing manual YAML prep time by 80%.
- **Multi-Tenant Enablement:** Leverages Kubernetes namespaces and ArgoCD to roll out multiple application versions within the same environment safely.
- **Secure Secret Injection:** A custom secrets operator integrates with key vaults to inject sensitive data at runtime, adhering to zero-trust principles.

- **Centralized Visibility Portal:** stakeholders can monitor deployment status and telemetry from a central dashboard without requiring direct inbound device access.

Impact

- **Real-Time AI Processing:** Localized data handling drastically reduced latency, enabling mission-critical decision-making in defense and logistics.
- **95% Operational Reliability:** The pull-based GitOps model ensured continuous application availability even during prolonged network outages.
- **70% Lower Manual Effort:** Automated Kubernetes-native dependency management eliminated complex manual setups and lowered operational overhead.
- **Bandwidth & Cost Efficiency:** Drastic reductions in cloud storage and bandwidth costs achieved through intelligent local data aggregation and filtering.
- **Future-Ready Foundation:** The modular design now supports over 15 geographically distributed edge devices, providing a scalable blueprint for diverse industrial IoT use cases.