

Deploy and Run AI Workloads at the Edge

Government agencies, intelligence organizations, and the military want and need to take advantage of the latest generative artificial intelligence (Gen AI) and other AI models for data analysis but face significant barriers to entry. Most organizations prefer to deploy and run AI on private infrastructure due to the need to maintain strict control of critical data assets and ensure the highest levels of security.

These AI challenges become even greater for edge use cases—especially in mobile or tactical environments where hardware options and on-the-ground technical expertise are limited, network connectivity is intermittent or non-existent, and security is a key consideration. However, Edge AI is essential for running critical workloads where they can make the most difference.

Introducing Icehouse Edge from Flywheel Data

Flywheel Data has partnered with [Spectro Cloud](#), a leading provider of Kubernetes management platforms, to address the needs of organizations that need an easy, flexible, and highly-secure way to deliver AI applications at the edge. Icehouse Edge combines cutting-edge hardware with Spectro Cloud's Palette VerteX to deliver a ruggedized, mobile unit that delivers the GPU acceleration necessary to support AI inferencing.

Icehouse Edge offers simplified management, works in disrupted, disconnected, intermittent, and low-bandwidth environments, and supports easy model tuning and execution with hardened security—all in a unit that requires just 9" x 9" x 17" of space.



An integrated AI software stack enables organizations to start using the device immediately, and easily manage everything—from the edge OS and infrastructure to the models powering innovative AI apps—with less technical expertise and fewer delays and hassles.

Solution Benefits

Icehouse Edge benefits include:

- **Compact, mobile, and rugged.** Dedicated, ruggedized hardware platform satisfies a full range of edge use cases including fixed, tactical, and mobile deployments.
- **Up and running in minutes.** A simple, plug-n-play installation approach enables almost anyone to get up and running quickly without being an expert.
- **Simple to use.** Users get fast results with a minimum of training.
- **Supports any connectivity model.** Can operate online or function in disrupted, disconnected, intermittent, and low-bandwidth (DDIL) environments, based on a unique decentralized architecture.
- **Optimized for Edge AI.** Advanced features simplify deployment and updates of complete AI application stacks, improve AI model accessibility, and more.
- **Ultra-secure.** Protect critical intellectual property and sensitive data with complete security from silicon to app, including validated FIPS 140-2 encryption, immutability, and other advanced security capabilities.
- **Complete lifecycle control.** Manage AI-optimized infrastructure across day 0 and day 2 operations, including full-stack declarative management, zero-downtime upgrades, advanced observability, RBAC, and backup/restore.

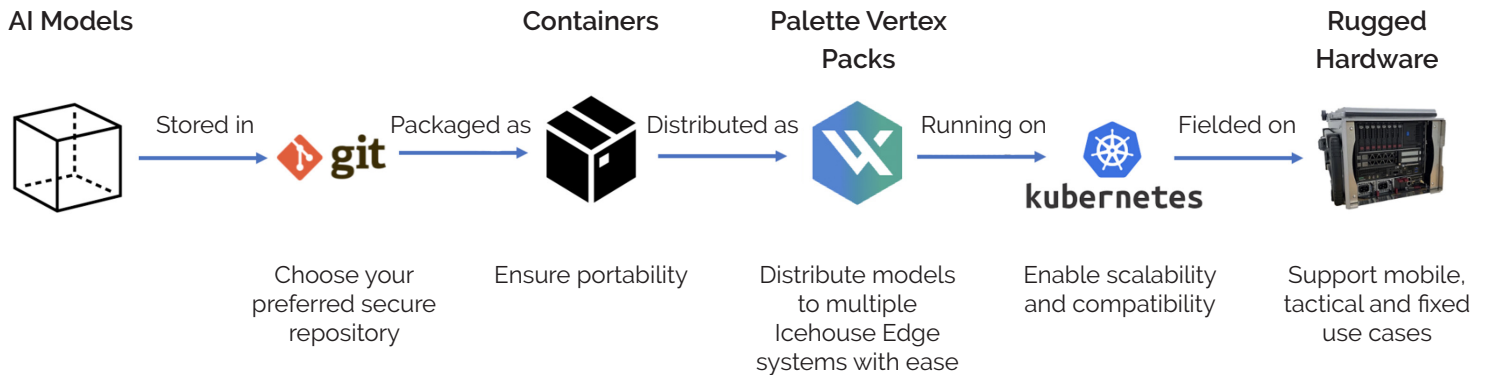
Unleash the Power of Generative AI

Large language models (LLMs) and other AI capabilities are unleashing a range of functionality that can have enormous benefits for government agencies, intelligence organizations, and the military—ranging from document summarization and analysis to translation to image categorization, captioning, and much more. However, due to security and privacy constraints, organizations like these can't simply upload their data to the cloud to take advantage of the latest AI tools.

Icehouse Edge makes it fast and easy for teams to take advantage of the latest AI models to process critical data without compromising security or causing unnecessary delay.

How Icehouse Edge Works

Icehouse Edge is designed to deliver all the hardware and software necessary to enable AI operations in a simple-to-deploy and simple-to-use form factor suitable for deployment anywhere including the [tactical edge](#). The solution can be delivered with your preferred AI models pre-deployed, or you can import and easily deploy your own. Icehouse Edge reduces the time necessary to deploy and run edge AI models and see results. Build your preferred AI application stack and deploy it across any number of edge locations with a few clicks.



Optimized for Cloud Native AI Stacks

Most AI models run in containerized, Kubernetes-based environments, but that introduces a lot of complexity and raises the bar when it comes to infrastructure performance, availability, cost and security at the edge. [Palette Vertex with Edge AI](#)—Spectro Cloud's battle-tested Kubernetes management platform for public sector organizations like USAF and US Navy—solves these challenges by providing full-stack control across OS, Kubernetes, AI engines and models within Icehouse Edge. Palette Vertex is verified for FIPS 140-2 across management plane, distribution, and managed clusters.

EdgeAI enables you to:

- Deploy an AI stack easily, from fast, secure hardware onboarding to repeatable 'blueprints' using Cluster Profiles.
- Choose your preferred OS, Kubernetes distribution, AI engines and models, and any of 50+ software integrations—or bring your own.
- Eliminate configuration drift, with declarative management and one-hour reconciliation loops that ensure your stack will always stay as designed across all your edge locations..
- Deploy additional AI models via your own private repositories and from Hugging Face.
- Update edge models frequently, with easy repository integration, safe, version-controlled zero-downtime updates, and canary deployments with rollbacks.
- Secure edge infrastructure and the software stack with tamper-proof immutability, integrated [SBOM and other security scans](#), persistent data encryption, granular [RBAC](#), and other advanced security controls.
- Meet all your compliance requirements with FIPS, ISO 27001, SOC 2.



Unlimited AI Use Cases

While Icehouse Edge has been tested on use cases such as document summarization and image categorization and captioning, it's designed to provide a general operating framework. The solution can run virtually any AI model that an organization needs to run now or in the future, while keeping operations simple so that users can focus on results. It's a great fit for small teams operating at the edge in mobile or tactical situations where connectivity is a challenge or where there is simply too much data to move and analyze in a timely manner.

The Icehouse Edge solution is not intended for large teams, data center deployments, or situations requiring extensive model building

Getting Started with Icehouse Edge

With ease of deployment, ease of use, and security built in, Icehouse Edge is your quickest, easiest route to Artificial intelligence at the edge. If you are ready to jumpstart your Edge AI journey, Flywheel Data can help you stand up a solution based on Icehouse Edge that is tailored to your mission requirements and have you operational and compliant in the shortest time possible.

If you're ready to learn more, contact Flywheel Data about a customized Icehouse Edge workshop.

Customer references are available upon request.

About Spectro Cloud

Spectro Cloud uniquely enables enterprises and public sector organizations to deploy and manage Kubernetes in production, at scale. Its Palette next-gen Kubernetes management platform gives IT and platform operations teams effortless control of the full Kubernetes lifecycle across multiple clouds, data centers, bare metal and edge environments. [Palette VerteX](#) uniquely addresses the security challenges of government and highly-regulated industries, backed by the Spectro Cloud Government team.

Learn more and get started with Palette at spectrocloud.com

About Flywheel Data

Flywheel Data provides elite solution design, system integration, software development, and product resale for data-driven organizations.

Based on our experiences with the US Government and top commercial companies; Flywheel Data recognizes that data and people are at the center of a successful, data-driven organization.

Our goal is to arm our clients with the right tools, platforms, and culture to accelerate data-driven insights.



3033 Wilson Boulevard, Suite 700
Arlington, VA 22201

Phone: (703) 647-4137

Email: info@flywheeldata.com

www.flywheeldata.com