
edge
edge
edge
Red Hat
Unlock business value at the edge
Rediscover the Red Hat Edge portfolio

Table of contents

Introduction

Unlock opportunity at the edge

As industries adopt more digital tools and intelligent systems, the edge is becoming a strategic opportunity to boost business value.

In manufacturing, it's where vision systems inspect products in near real time. In retail, it powers in-store analysis and transaction processing. In energy and utilities, it monitors remote infrastructure and maintains uptime in unforgiving conditions.

These use cases rely on shifting processing power closer to users and data sources—whether in a factory, a point-of-sale (POS) machine, or a backroom server. The edge of the network is where insights emerge, where decisions are made, and increasingly, where AI is operationalized. It's helpful to think of the edge as more than extending the datacenter. Consider how your organization can transform operations to bring compute power and vital intelligence right to where data is generated to unlock real business value.

Edge is accelerating and evolving quickly

Today's organizations need to process, analyze, and act on growing volumes of data in real time. Most of this data is being generated outside of traditional datacenters and cloud environments, in locations not designed to process and distribute information to multiple sources. From internet of things (IoT) sensors and robotics to embedded devices and localized compute power, the edge is where critical business intelligence begins.

But as edge expands, success will require managing compute resources strategically. This means treating the edge not as a fragmented collection of point solutions, but as an integrated part of your IT infrastructure. This will require the ability to standardize, scale, and automate how edge systems are built, monitored, and maintained–especially across hundreds or thousands of endpoints.

For many organizations, edge computing is supporting:

- **Real-time decision making.** AI and machine learning (ML) models can infer and act on data without needing to call home to the cloud.
- **Operational resilience.** Applications can continue running even when connectivity is unreliable or unavailable.
- **Data privacy and sovereignty.** Sensitive data can remain on site, streamlining compliance and mitigating risk.
- **Cost savings.** Roundtrips to the cloud can be eliminated, reducing bandwidth costs and latency bottlenecks.

This convergence of operational needs and technological possibility is supporting momentum in edge adoption, particularly in sectors such as telecommunication (telco), manufacturing, energy, logistics, and retail.

Edge management is difficult, yet essential

Despite these benefits, managing edge environments remains complex. Organizations must provision, configure, monitor, and update thousands of distributed devices with limited visibility, often over unreliable networks, and without on-site staff.

Organizations often struggle with basic operations such as device onboarding, operating system (OS) updates, and visibility across their edge environments. Without centralized control and automation, these tasks quickly become unscalable and prone to error.

Common challenges include:

- Manual onboarding and configuration.
- Drifting OS versions and inconsistent application states.
- Inability to manage AI model lifecycles in disconnected locations.
- Security risks in air-gapped or aging environments.

To scale effectively, organizations need:

- Declarative, policy-based management.
- Remote observability and orchestration.
- Transactional update capabilities (all or nothing).
- Repeatable deployment pipelines (for applications and AI models).
- Tools that abstract complexity but preserve control.

Empower your edge operations

To succeed at the edge, enterprises need to overcome these management obstacles and turn the edge into a true extension of their hybrid infrastructure.

This is where Red Hat's platform-based approach delivers unique value. With consistent, open infrastructure that spans the complete edge spectrum, supported by familiar tools and automation, a platform approach reduces complexity and improves outcomes. But, like most IT solutions, edge requires a tailored approach to be effective.

This e-book explores the opportunities for enterprises at the edge. It also provides a practical guide to how the Red Hat® portfolio can help operationalize AI at the edge, streamline OS deployment and update processes, scale Kubernetes architectures across topologies, and automate the full edge device lifecycle–from deployment to retirement.

Chapter 1

Discover Red Hat Edge

Red Hat has long been a leader in delivering hybrid cloud infrastructure, open source innovation, and automation solutions. These same principles are now being applied at the edge, to bring together the technologies, tools, and automation needed to build and manage a consistent, security-focused, and scalable edge infrastructure.

Built for complexity at the edge

Red Hat provides a comprehensive edge AI architecture that builds on the strength of open source, hybrid infrastructure, and DevOps methodologies.

Backed by decades of open source leadership and hybrid cloud expertise, Red Hat is helping organizations do more at the edge by providing:

- **Greater flexibility and choice.** Support for a wide range of hardware and on-premise environments.
- **Consistency of operations across hybrid environments.** A unified platform and toolset that span datacenter, cloud, and edge environments.
- **Open source DNA.** Standardized application program interfaces (APIs) and community-powered innovation at every layer of the stack.
- **IT and OT convergence.** Solutions designed to work across operational and technical teams to bridge skills gaps.

Why Red Hat for edge environments?

Organizations adopting edge computing need more than individual products—they need a unified strategy. Red Hat's edge portfolio gives organizations the flexibility, consistency, and control needed to extend their IT environments from the datacenter and cloud all the way to the farthest edge.

Red Hat's edge portfolio is designed to help organizations in 5 key areas:

1 Build and operate a common platform across core, cloud, and edge

Red Hat's approach to edge allows for a consistent infrastructure and application experience across environments. Unlike rigid, vendor-prescribed stacks that require significant rewrites, Red Hat's open platform allows organizations to extend their existing investments to the edge without rearchitecting.

2 Automate and manage edge infrastructure anywhere, even when disconnected

Using automation and edge-native lifecycle tooling, organizations can provision and maintain devices across thousands of sites with limited bandwidth or on-site IT.

3 Add a security focus to every layer of the stack, from hardware to application

Red Hat solutions include built-in layered security measures developed through a rigorous software supply chain security process. Organizations benefit from end-to-end visibility and continuous updates to keep distributed infrastructure resilient and protected.

4 Operate in heterogeneous environments through open innovation

Edge deployments rarely fit a universal mold. Red Hat's platform and vast partner ecosystem support custom hardware, niche software, and specialized workloads, helping organizations deploy and innovate on their own terms.

5 Accelerate outcomes with validated solutions and ecosystem support

Red Hat collaborates with partners to build and share pretested deployment patterns and blueprints for common edge scenarios. These resources are open and downloadable, helping teams more efficiently with greater confidence.

Red Hat's edge portfolio helps organizations to modernize infrastructure, deploy applications and AI models with confidence, and scale resilient, security-focused systems across thousands of locations.

Chapter 2

The AI advantage at the edge

To make the most of the data that is generated and captured across distributed environments, many organizations are deploying AI capabilities closer to where data is created. IDC projects overall spending on AI will reach US$632 billion by 2028, with a compound annual growth rate (CAGR) of 29.0% from 2023 to 2028.[1] A major area of this investment is in AI at the edge, known as edge AI, which combines the power of ML with the responsiveness and resilience of edge computing.

Rather than sending data back to a central datacenter or public cloud for analysis, AI models are increasingly being trained in non-resource-constrained environments and deployed to the edge for real-time inference. This allows organizations to analyze, decide, and act on local data within milliseconds—all without relying on continuous connectivity.

What is real-time inference?

This refers to the process where an AI model makes predictions immediately as new data arrives. It prioritizes low latency and quick responses, making it suitable for applications requiring instant feedback or action, based on live data streams.

1 IDC. "IDC FutureScape: Worldwide Artificial Intelligence and Automation 2025 Predictions." Document # US51666724, Oct. 2024. (Registration required)

From increased efficiency to enhanced security posture, edge AI is delivering benefits in real-world implementations, such as:

- **Predictive maintenance:** Manufacturing facilities can use ML models at the edge to detect warning signs of malfunction. Identifying when a machine or system is likely to fail allows for proactive repairs, which helps optimize maintenance costs and minimize unplanned downtime.
- **Loss prevention:** Instead of relying on the manual efforts of security teams, retail stores can use ML models to monitor multiple security camera feeds in real time. Models can flag suspicious activity, identify theft, and minimize losses.
- **Energy efficiency:** Edge AI can be applied to optimize energy consumption in industries where efficiency is essential, such as manufacturing, transportation, and agriculture. Sensors can collect real-time data about energy usage patterns and needs, helping organizations strategically manage energy resources.

As AI-powered applications proliferate, edge AI helps reduce costs, boost productivity, increase uptime, and improve customer experiences.

Why edge AI matters now

Edge environments such as factory floors, retail backrooms, and remote energy installations are all sources of vital data, but that data often resides where conditional challenges such as latency, bandwidth, or data sovereignty limit the practicality of cloud-based processing.

Edge AI addresses these constraints by allowing organizations to:

- **Make faster decisions.** Inference happens at the source, delivering insights in real time.
- **Improve reliability.** Local processing continues even when connections drop.
- **Reduce costs.** Eliminate the need to send low-value, high-volume data to the cloud.
- **Boost data security.** Sensitive data stays local, supporting compliance and privacy.

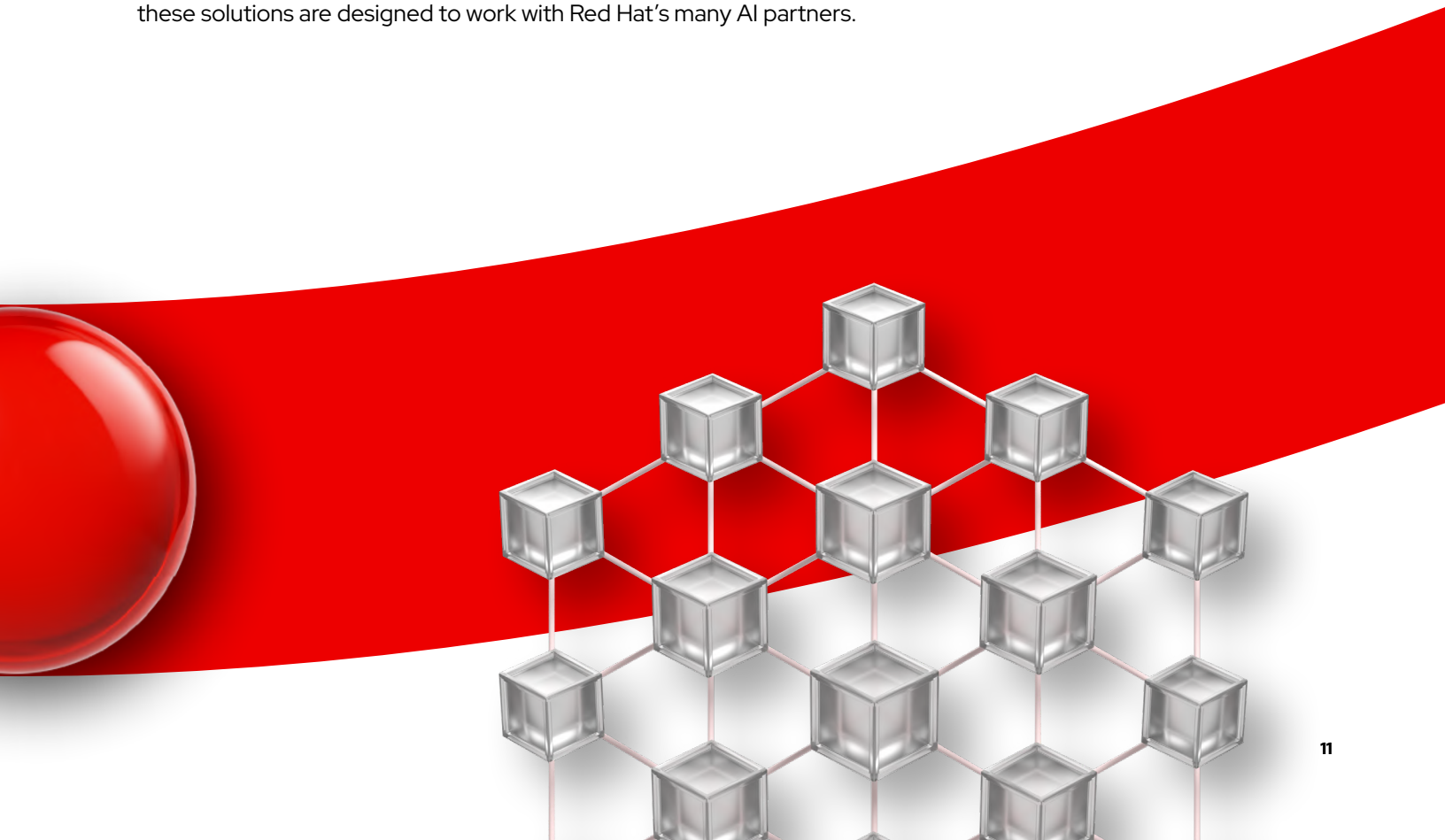
As AI adoption grows across industries, organizations require a scalable, security-focused, and consistent way to manage the full AI lifecycle across the hybrid cloud, from model development in the datacenter to inference at the far edge.

A closer look at Red Hat's approach to edge AI

Red Hat provides an extensive edge AI architecture that builds on the strength of open source, hybrid infrastructure, and DevOps methodologies.

- Red Hat OpenShift® AI offers a flexible MLOps platform for building, training, and deploying AI models.
- Red Hat Device Edge, built on MicroShift and Red Hat Enterprise Linux®, brings the full AI runtime to resource-constrained locations.
- Red Hat Enterprise Linux AI, for large language models (LLMs) and generative AI, supports foundation model development and alignment in security-focused environments.
- Red Hat Ansible® Automation Platform orchestrates infrastructure from core to edge, automating provisioning, scaling, and updates, and provides an automated foundation to support AI implementations at scale.
- Red Hat OpenShift with Red Hat OpenShift Pipelines supports building, testing, and deploying both applications and containerized versions of AI models, streamlining end-to-end MLOps workflows.
- Red Hat Advanced Cluster Management for Kubernetes helps deploy and manage AI models across hybrid and edge environments, whether using Red Hat OpenShift or MicroShift.
- Red Hat Quay provides a security-focused and scalable registry for distributing and hosting AI models alongside containerized applications.

Together, these technologies allow customers to operationalize AI workloads wherever they need them—from in-vehicle inference to predictive maintenance on the production line. In addition, these solutions are designed to work with Red Hat's many AI partners.



Chapter 3

Lightweight innovation for edge environments

Modern edge environments require innovative solutions that are powerful and purpose-built, while also being lightweight and portable.

Whether on the shop floor, in a wind turbine, or inside a POS system, today's edge deployments require systems that are compact yet capable of doing more than ever before.

As organizations expand their edge footprint, they need solutions that provide performance under industry-specific constraints without adding management complexity or introducing fragility.

Red Hat is supporting this shift with a consistent approach to how operating systems and applications are built, deployed, and maintained at the edge. With new tools and features introduced through Red Hat Device Edge—a flexible platform that supports different workloads across small, resource-constrained devices at the edge—organizations can now apply container-native workflows to the OS itself. This brings new levels of consistency, automation, and flexibility to the edge software lifecycle.

Discover OS deployment using image mode in Red Hat Device Edge

Red Hat Device Edge introduces image mode, an innovative solution that allows organizations to build, package, and distribute an OS just like containers.

Image mode represents a significant evolution from the OSTree-based approach used in earlier Red Hat Enterprise Linux for edge deployments. While OSTree manages OS updates via layered filesystem snapshots, image mode packages the entire OS as a bootable container (bootc) image, allowing for simpler, more streamlined deployments.

This container-based OS image is managed using the bootc tool, which makes sure the image is bootable by embedding necessary bootloader and kernel metadata.

This shift aligns edge OS management with cloud-native DevOps workflows, making it easier for teams familiar with container tooling to build, test, and deploy edge operating systems. Teams can take advantage of the same tooling that is used for application containers (e.g., CI/CD tools), but for the OS.

Using image mode allows teams to:

- Build, test, and deploy edge OSes as open container initiative (OCI)-compliant container images.
- Converge application and OS pipelines into a single GitOps-enabled continuous integration and continuous delivery (CI/CD) workflow, using the same processes and tools for application containers and the OS or base platform.
- Gain reproducibility and portability across hybrid and multicloud environments.
- Streamline updates with better control, security focus, and rollback options.

This modern, declarative approach eliminates the need for manual configuration or master images. Now, edge operating systems can be treated like versioned, traceable, and auditable artifacts, delivering stronger governance and accelerated time to value.

Image mode is the preferred method for edge OS deployment starting with Red Hat Enterprise Linux 10. OSTree-based Red Hat Enterprise Linux for edge remains supported in Red Hat Enterprise Linux 9 for existing workloads but will be superseded by image mode in upcoming releases.

Security is built in

Image mode benefits from cryptographic validation, security-focused provenance, and toolchain consolidation. This means that organizations can confidently deploy updates to remote, disconnected devices knowing the system is verifiable, rollback-capable, and supported by Red Hat's trusted supply chain process.

The bootc tool also makes sure that security contexts, such as SELinux labels, are preserved within the bootc image, maintaining compliance with strict security policies in regulated environments.

New capabilities at the edge for real-time and AI-enhanced innovation

Since its initial release in 2023, Red Hat Device Edge has evolved to support more advanced and specialized use cases across industrial, retail, and many other industries.

The platform now includes a host of features that give organizations greater performance, precision, and flexibility at the edge.

Real-time performance for latency-sensitive applications

To meet the demands of latency-sensitive edge environments, Red Hat Device Edge introduces a suite of performance-tuning capabilities. These features are designed to deliver deterministic, real-time behavior where milliseconds—or even microseconds—matter most, such as in industrial control systems, robotics, and other safety-critical operations.

- **Low-latency kernel support** for ultra-fast compute scenarios.
- **CPU core isolation** to help teams maintain consistent performance for priority workloads.
- **Workload partitioning** to prevent jitter and resource contention.
- **Maximum latency targets** of lower than 50 microseconds to allow for demanding workloads like software-defined industrial control systems, robotics, and safety-focused operations.

Advanced networking support

As edge deployments grow more complex, networking flexibility becomes essential. Red Hat Device Edge provides advanced networking capabilities that support multinetwork connectivity, modern addressing standards, and fine-grained traffic control. This allows organizations to tailor networking to meet the needs of their applications, locations, and compliance requirements.

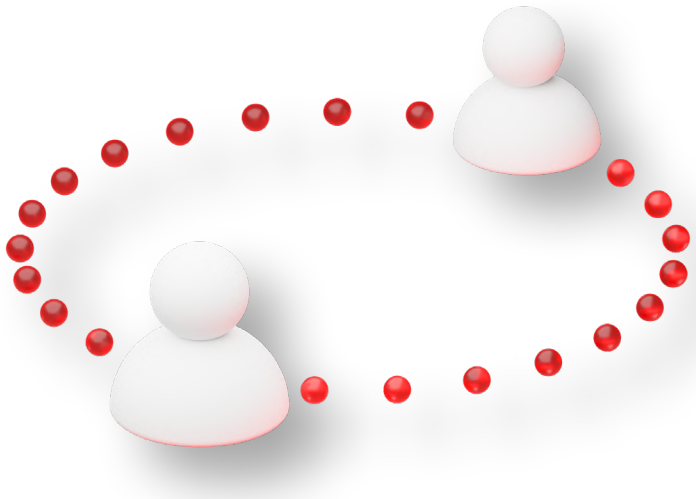
- **Multus integration.** This open source project allows Kubernetes pods to attach to multiple network interfaces per pod, providing separation of control, data, or sensor networks.
- **IPv6 support.** Including single-stack and dual-stack configurations, IPv6 allows organizations to address connectivity in modern, high-scale environments.
- **Support for use cases.** This support includes telcos, industrial automation, and retail environments, where traditional IPv4 addressing may fall short.

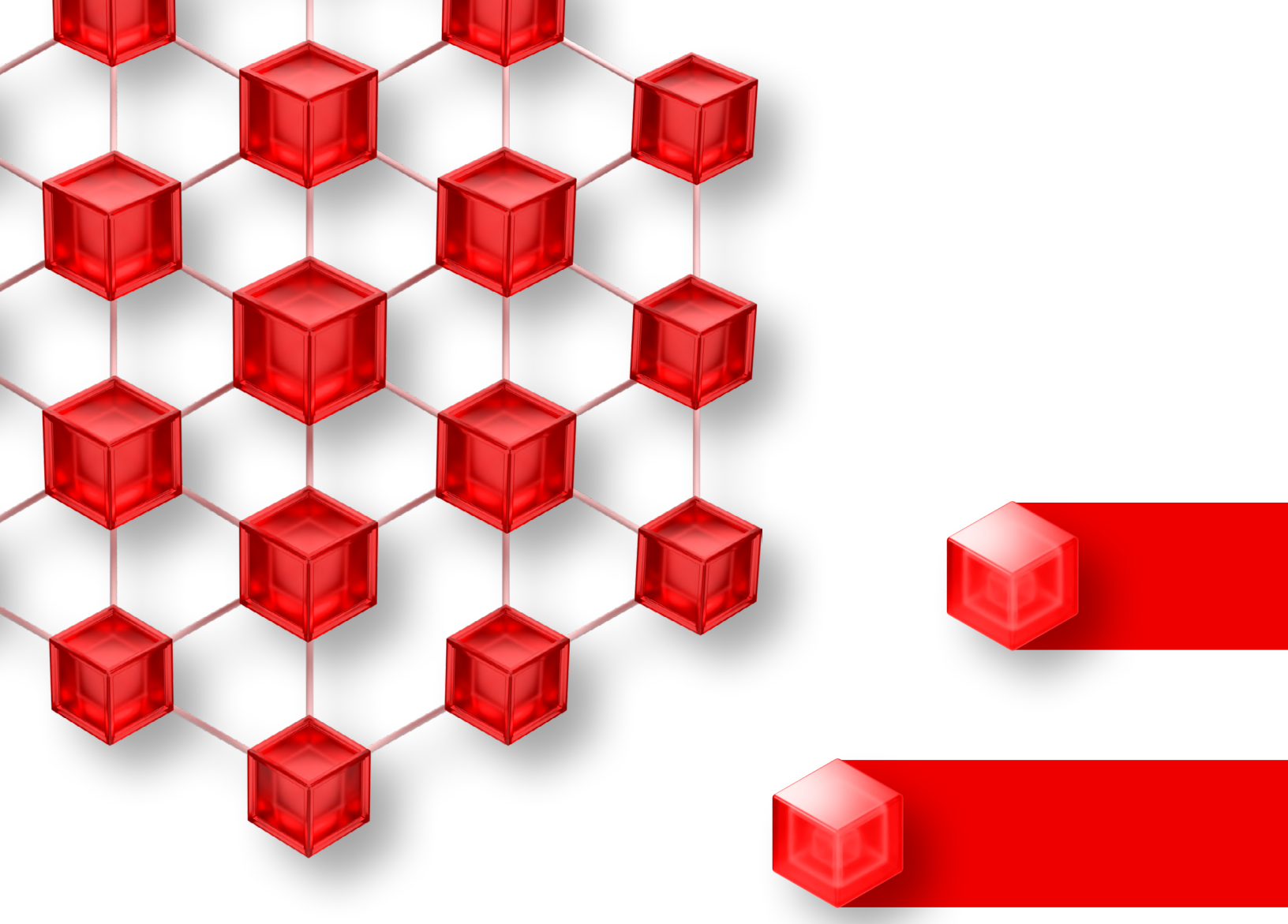
FIPS-ready compliance and security

In highly regulated industries and government environments, security and compliance are foundational. Red Hat Device Edge supports stringent security requirements with built-in features that meet cryptographic compliance at the OS level.

When installed on Red Hat Enterprise Linux in Federal Information Processing Standards (FIPS) mode, MicroShift core components use Red Hat Enterprise Linux cryptographic libraries that have been submitted for FIPS 140-3 validation.

This maintains compliance with strict U.S. federal and international standards for cryptographic operations, which is a requirement for regulated industries and government deployments.





AI model serving at the edge

As AI adoption expands, organizations need lightweight, flexible ways to deploy inference models in environments with limited compute, power, or connectivity. Red Hat Device Edge supports edge-native AI workloads by allowing model serving directly on resource-constrained hardware. This brings intelligence closer to the data and allows for real-time insights without dependency on centralized infrastructure by providing several benefits.

- Support for NVIDIA Jetson Orin and other low-power AI accelerators.
- Tech preview of Red Hat OpenShift AI model serving on MicroShift using raw deployment mode to:
  - Deploy AI inference models as KServe manifests.
  - Run Red Hat OpenShift AI-supported ServingRuntimes on lightweight Kubernetes.
- Bringing AI inference directly to the edge, supporting real-time prediction and vision processing workloads without reliance on centralized infrastructure.

Use cases unlocked by lightweight edge innovation

Red Hat's lightweight edge platform is ushering in a new generation of innovation across use cases and industries.

Industrial manufacturing

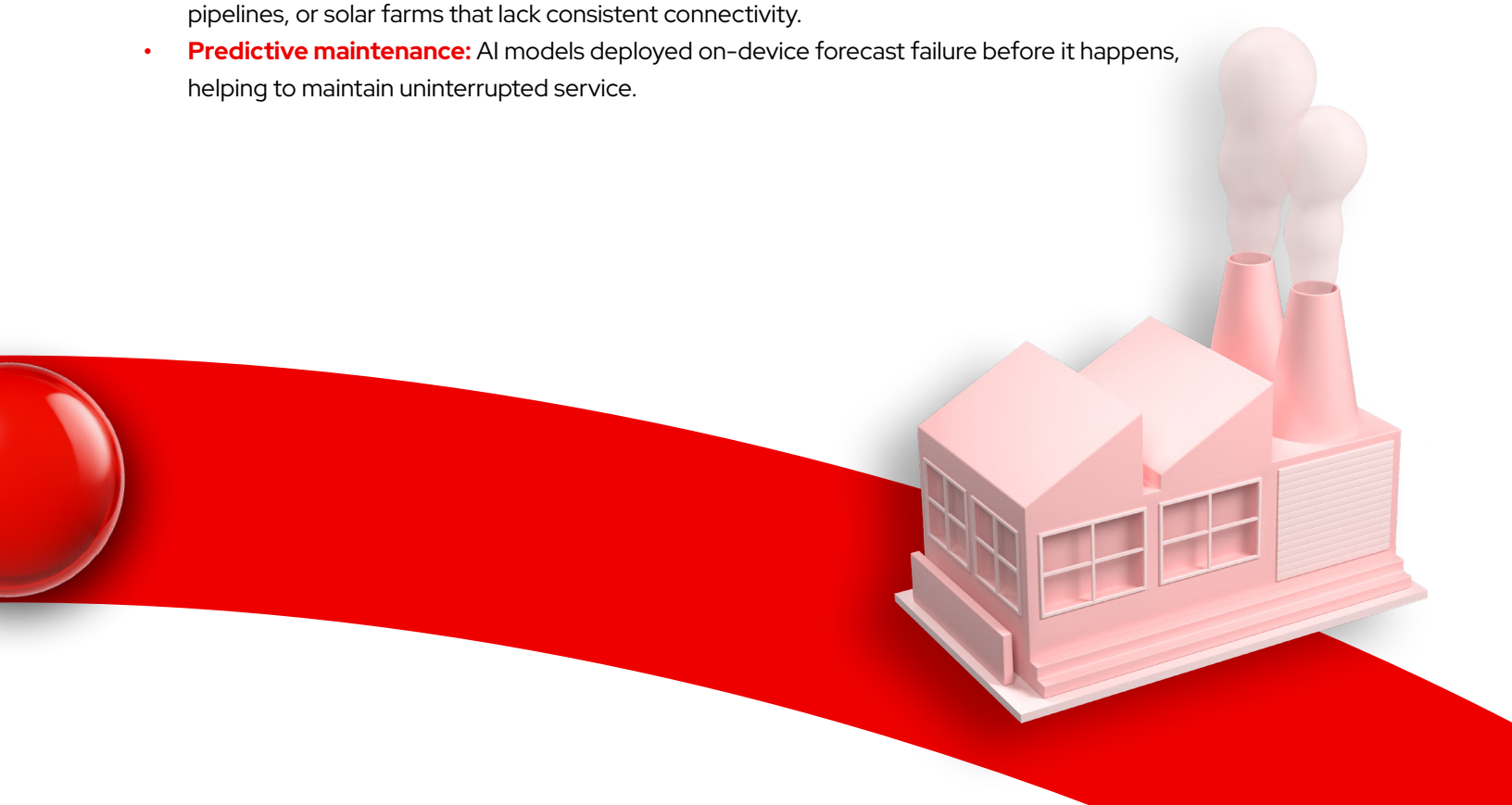
- **Precision control and automation:** Real-time kernel and workload partitioning maintain low-jitter control for robotics and machinery.
- **On-device quality assurance:** Cameras on production lines run vision models on MicroShift with NVIDIA Jetson, detecting defects in real time, boosting quality control and assurance.
- **Process safety:** FIPS-compliant deployments ensure critical operations meet regulatory requirements.

Retail

- **Smart store infrastructure:** Multus networking, an open source project that allows Kubernetes pods to attach to multiple networks, creates separation of in-store systems (including point of purchase, video analytics, security, and among others) while streamlining deployment.
- **Loss prevention:** Red Hat OpenShift AI model serving detects theft in real time using in-store video feeds without cloud roundtrips.

Energy and utilities

- **Remote site control:** Image mode streamlines OS and application updates to wind turbines, pipelines, or solar farms that lack consistent connectivity.
- **Predictive maintenance:** AI models deployed on-device forecast failure before it happens, helping to maintain uninterrupted service.



Chapter 4

Scale Kubernetes from the core to the edge

As edge strategies scale and adapt over time, organizations need the flexibility to run workloads in a wide range of conditions—from high-availability (HA) backroom clusters to low-power devices in disconnected environments.

Kubernetes offers the agility and consistency needed, but only if it can be adapted to these diverse deployment models.

A common platform for the edge

Red Hat OpenShift brings the full power of enterprise Kubernetes to the edge. By offering multiple topologies—3-node, 2-node, and single-node clusters—Red Hat OpenShift allows organizations to scale their cloud-native infrastructure beyond the datacenter without compromising availability, security posture, or operational consistency.

Compact 3-node OpenShift cluster

In many edge locations, space and power are limited, but availability is not negotiable. That is where a 3-node compact cluster is beneficial.

This configuration delivers full Kubernetes functionality and HA across 3 physical servers using minimal rack space. It's an ideal solution for edge deployments that need resilience and local compute without the overhead of a full-scale datacenter.

When paired with Red Hat OpenShift Data Foundation, this topology provides integrated compute, storage, and networking. This creates a complete, cloud-like environment at the edge.

Common use cases include:

- Retail backrooms supporting inventory systems, POS, and AI-powered loss prevention.
- Industrial sites running local automation and analytics workloads.
- Branch offices or regional hubs requiring localized processing with HA.

Organizations can manage these clusters with the same tools used across their enterprise:

- Red Hat Advanced Cluster Management for deployment and lifecycle management.
- Red Hat Advanced Cluster Security for Kubernetes for workload protection.
- Ansible Automation Platform for integrated infrastructure provisioning and orchestration.

2-node OpenShift cluster (tech preview)

To further reduce hardware requirements while maintaining HA, Red Hat offers 2 innovative deployment models tailored for cost-sensitive environments.

2-node with arbiter (TNA)

This approach technically functions as a 3-node cluster, with the 3rd node—a lightweight arbiter—serving only to maintain etcd quorum.

- Ideal for retail, industrial, or telco deployments where a 3rd full node is cost-prohibitive.
- Arbiter can be colocated or remote.
- Supports Red Hat OpenShift Virtualization and partner software-defined storage (SDS) solutions.
- Requires <500ms end-to-end latency to the arbiter node.

2-node cluster with fencing (TNF)

A true 2-node configuration, TNF uses proven Red Hat Enterprise Linux HA technologies, such as corosync or pacemaker, for quorum and fencing.

- Uses base management controller (BMC) with RedFish for automated node fencing in a split brain situation.
- X86, bare metal only—suited for highly cost-constrained locations.

Both variants offer a middle ground between single-node simplicity and full HA—balancing resilience, performance, and cost.

Single-node cluster

For edge sites with the most constrained resources, Red Hat offers a single-node solution, which provides the full Red Hat OpenShift experience on a single machine.

Unlike Red Hat build of MicroShift (see below), which is optimized for minimal Kubernetes on micro devices, Red Hat's single-node approach includes all standard OpenShift components, making it ideal for running full-featured workloads where HA isn't required.

A single-node solution is best suited for:

- Telco open radio access network (ORAN) deployments where compute is deployed at the cell site.
- Remote industrial environments with limited space and compute.
- IoT aggregation points collecting data from sensors, cameras, or devices.
- Disconnected or air-gapped scenarios needing full container orchestration on a single machine.

Despite its minimal footprint, single-node OpenShift is fully manageable with the following tools:

- Red Hat Advanced Cluster Management for centralized management.
- Red Hat Advanced Cluster Security to maintain workload security focus.
- Ansible Automation Platform for supporting network, router, and switch configurations.

By reducing operational sprawl and eliminating the need for custom tooling, single-node OpenShift extends modern cloud-native operations to even the most remote edge locations.

MicroShift

For ultra-small form factor deployments with extreme resource constraints, Red Hat build of MicroShift offers a lightweight Kubernetes distribution derived from Red Hat OpenShift.

Designed specifically for edge use cases–including industrial gateways, ruggedized embedded systems, or small ARM/x86 devices–MicroShift provides just enough Kubernetes to containerize workloads without the full overhead of Red Hat OpenShift.

MicroShift is best suited for:

- Remote or mobile edge environments with intermittent or unreliable connectivity.
- Space- and power-constrained controllers, gateways, or industrial devices.
- Compact retail systems or IoT endpoints.
- Devices requiring extended uptime and local autonomy.

Use cases across topologies

Each OpenShift topology aligns with different business requirements and environmental constraints. Here's how customers are putting them to use:

Retail

- Three-node clusters support store operations and AI-powered surveillance with local resilience.
- Two-node configurations reduce hardware costs across large chains while maintaining HA.
- Single-node OpenShift supports smaller locations or franchise stores with plug-and-play deployment.

Manufacturing

- Single-node OpenShift and 2-node clusters support edge compute on production lines.
- AI/ML inference runs at the edge while models are managed centrally via Red Hat OpenShift AI.

Telco

- Single-node OpenShift and MicroShift allow for lightweight RAN deployments and IoT gateways.
- Three-node clusters provide regional aggregation with scalable services at the near edge.

Energy and utilities

- Single-node or MicroShift deployments serve remote substations or generation sites where minimizing physical infrastructure is key.
- Single-node OpenShift powers control systems in air-gapped environments with limited network access.

Chapter 5

Manage operations across the edge

For edge computing to deliver real business value, it must go beyond standalone deployments and scale across thousands of locations. From provisioning devices to managing ongoing updates, organizations need a way to standardize operations, mitigate risk, and adapt to change—without sending IT teams to every location.

Red Hat brings automation and management to the edge with a unified platform approach. Central to this approach is the introduction of Red Hat Edge Manager.

Red Hat Edge Manager

Red Hat Edge Manager, a product within the Red Hat Device Edge offering, provides a comprehensive, security-focused solution for managing fleets of edge devices that is streamlined, scalable, and built for the edge. Edge Manager provides the following business value to our customers:

Simple

- Intuitive edge operations for operational technology (OT) teams to onboard devices, run updates, and troubleshoot. Streamlined and intuitive processes, from onboarding to provisioning to deployment, support OT personnel with minimal IT staffing at the edge.
- Streamlined management of edge devices and applications through a policy-based approach.
- Flexible management options for your platform, configurations, and workloads from our vast ecosystem of open source partners.

Scalable

- Policy-based deployment maximizes operational efficiency. A desired-state configuration model provides consistent and scalable deployments, reducing manual effort, enforcing configuration consistency, and preventing configuration drift across fleets of devices.
- Resilient agent-based architecture allows for comprehensive device management in intermittent network connectivity or constrained environments.
- Autonomously manage and monitor devices to check for new configurations and report device status to increase productivity and efficiency for thousands of devices.

Built for edge

- Hardened security for device authentication and communication provides security-focused, scalable fleet management from onboarding to decommissioning.
- Monitor fleet device resources and complete device health checks to build insights into operations.

Managing edge devices and software effectively requires tight collaboration across organizations and systems, along with centralized management and orchestration of edge infrastructure and applications. This centralized approach also introduces a plethora of challenges, including connectivity demands, heterogeneous hardware, operational complexity, and more.

If used with Ansible Automation Platform, Red Hat Edge Manager provides security-focused organizations to streamline the deployment, configuration, and lifecycle management of infrastructure and workloads across diverse, distributed, and sometimes disconnected edge environments.

Infrastructure at your fingertips

Automation is the foundation for managing hybrid cloud, and using Ansible Automation Platform, that foundation can extend out to the farthest edge.

Ansible Automation Platform helps team to:

- Orchestrate infrastructure consistently across datacenters, cloud, and near-edge locations.
- Automate complex provisioning workflows, from networking to storage to Kubernetes clusters.
- Scale repeatable configurations and updates across thousands of devices.

With event-driven automation capabilities, Ansible Automation Platform can respond to events, such as device failures or policy violations, without manual intervention. This reduces downtime, boosts efficiency, and allows edge teams to move more quickly with fewer resources.

A unified control loop

When used together, Red Hat Edge Manager and Ansible Automation Platform create a powerful control plane for fleet-wide edge operations.

Edge Manager is built to manage large fleets of edge devices, and is especially useful in environments with intermittent or low-bandwidth connections, remote or air-gapped locations, and strict security and compliance requirements. It provides:

- **Zero-touch onboarding and remote updating** for edge devices.
- **Declarative desired state management** for OS, application, and configurations.
- **Self-healing capabilities** that detect and correct drift.
- **Fleet templating** for allowing consistent, repeatable configuration at scale, regardless of device type or location.

When Edge Manager and Ansible Automation Platform are used together, Red Hat delivers a unified user experience, allowing operations teams to manage infrastructure and workloads in a single place. Ansible Automation Platform prepares the environment by configuring upstream networking, storage, switches, routers, and identity services. Once the site is ready, Edge Manager takes over to handle image deployment, new device enrollment, and ongoing health and configuration of devices. This can include aligning devices with declarative configurations stored in Git, which is optional and based on the organization's chosen workflow. As part of device onboarding, Ansible workflows can also be used to update IT service management (ITSM) systems for improved visibility, tracking, and compliance reporting across the deployment lifecycle.

Centralize edge operations with unified management

Running edge deployments across thousands of locations presents a dual challenge: device lifecycle management and application and cluster operations. For holistic edge operations, organizations need platforms that unify both perspectives.

- **Red Hat Edge Manager:** For device fleet provisioning, configuration, and policy-based deployment and health.
- **Red Hat Advanced Cluster Management:** For centralized policy, governance, lifecycle, and workload management of Kubernetes clusters, including those running on the edge (such as MicroShift or Red Hat OpenShift).

Edge management in action: Use cases

Deploying a new retail store

When a new store location is brought online, Ansible Automation Platform and Edge Manager can automate the full edge environment setup:

- Ansible Automation Platform provisions network infrastructure, including configuring routers, switches, firewalls, wireless access points, and in-store servers.
- Edge Manager deploys the edge nodes, including provisioning containerized applications responsible for point of sale (POS) systems, video surveillance, or inventory control.
- Ongoing operations are managed remotely, including updates, health monitoring, and policy enforcement (handled from a central location).

This approach reduces set-up time from days to hours and eliminates the need for costly on-site technicians.

Resolving device issues

Edge Manager provides built-in alerting capabilities that automatically monitor your edge devices and notify you when issues occur. In the case a device needs to be replaced or decommissioned:

- Ansible Automation Platform removes the device from the network and deconfigures any associated services or policies.
- Edge Manager decommissions the device.
- A new device can be shipped, plugged in, and onboarded, repeating the lifecycle with minimal human touch.

This closed-loop workflow maintains speed and repeatability at scale.

Why automation is essential at the edge

As edge environments expand, automation provides more than efficiency; it offers greater opportunities for operational success. Organizations using Red Hat automation and management solutions are able to:

Increase scalability
Apply consistent configurations and updates across thousands of nodes with minimal overhead.

Boost agility
Adapt infrastructure and services based on dynamic business needs and edge site requirements.

Focus on security and safety
Update patches, security policies, and rollback strategies—mitigating risk and addressing vulnerabilities.

Reduce downtime
Streamline recovery processes, manage device health, and avoid costly outages.

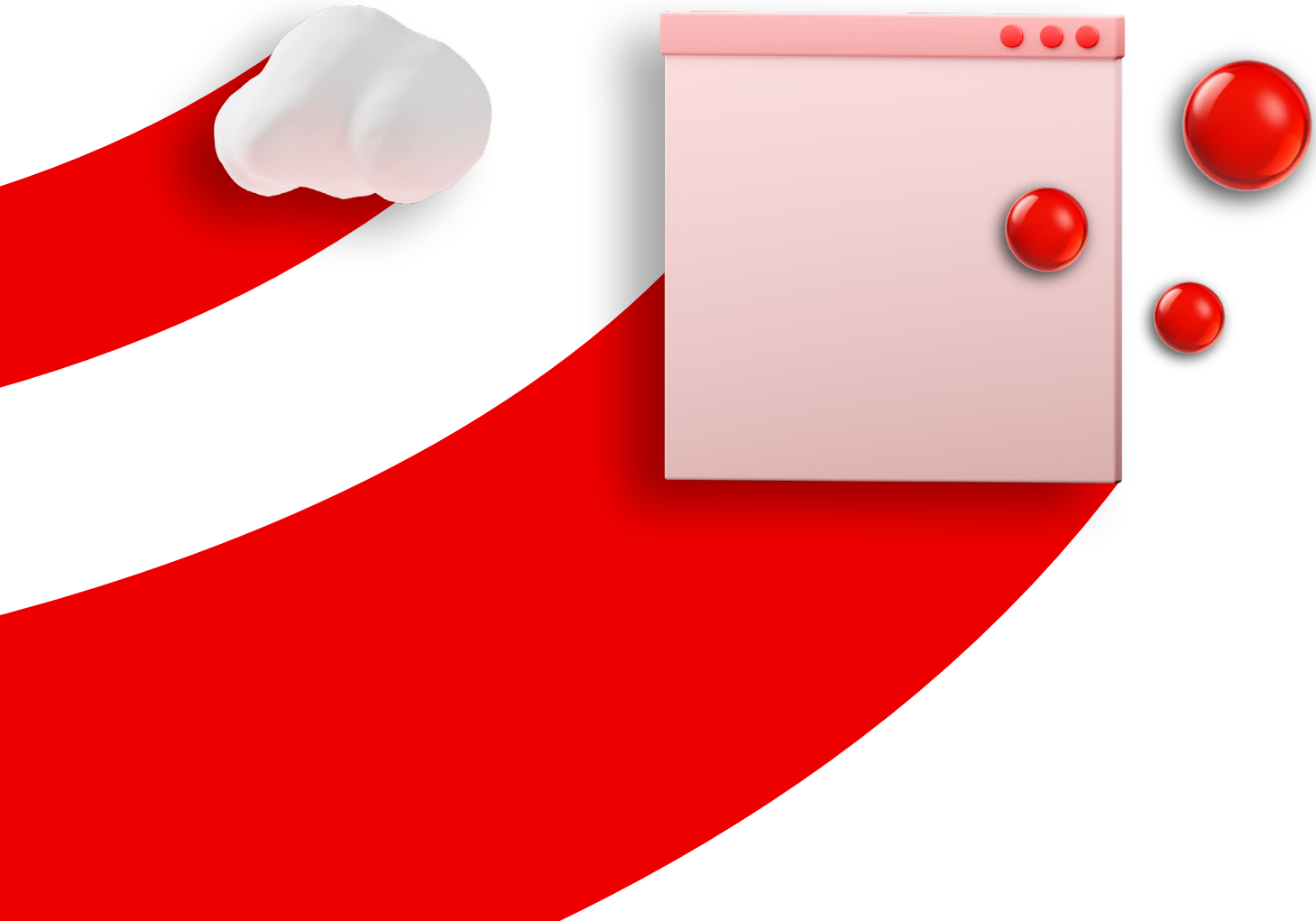
Improve efficiency
Minimize human error, streamline manual processes, and reallocate IT resources to strategic work.

Red Hat Advanced Cluster Management for Kubernetes

For organizations deploying containerized workloads at the edge, Red Hat Advanced Cluster Management for Kubernetes helps to consistently manage, observe, and operate hybrid, multicloud, and edge clusters from a single control plane. Integrating Edge Manager with Red Hat Advanced Cluster Management provides:

- Streamlined, security-focused onboarding of edge devices and clusters, including automated registration and image-mode management of Red Hat Enterprise Linux and MicroShift.
- Unified, policy-based lifecycle management using Red Hat Advanced Cluster Management clusters and applications (OpenShift and MicroShift), and Edge Manager for OS images and containerized workloads running directly on edge devices.
- Consistent, near-real-time state monitoring and remediation, correlating device health (via Edge Manager) and workload health (via Red Hat Advanced Cluster Management).

This creates a truly holistic, automated, and scalable edge management platform that spans devices, clusters, and workloads, everywhere your business operates.



Learn more

Bring your edge strategy to life

Unlocking the full value at the edge requires more than point solutions. It demands a consistent platform, automation, and open flexibility across every environment.

As edge environments grow more distributed and data-intensive, organizations must adopt a smarter, more unified approach to managing operations. That means moving beyond tactical fixes to a strategic vision that allows for repeatable, intelligent operations—from Day 0 deployment to ongoing updates and AI-driven insights. Automation, resilience, and centralized control are no longer optional; they are essential for scaling edge success.

Red Hat helps organizations meet these needs with a comprehensive, modular edge portfolio. Whether you are building lightweight infrastructure with Red Hat Device Edge, scaling Kubernetes with Red Hat OpenShift, automating deployments with Ansible Automation Platform, or operationalizing AI at the edge, Red Hat delivers the consistency, security focus, and agility to accelerate your edge strategy.

Connect with a Red Hat expert
to learn how Red Hat can help your organization architect a modern edge environment.