Accelerate intelligence at the edge

How to use AI to expedite decision making

Red Hat

Contents

Introduction

Business at the edge is accelerating

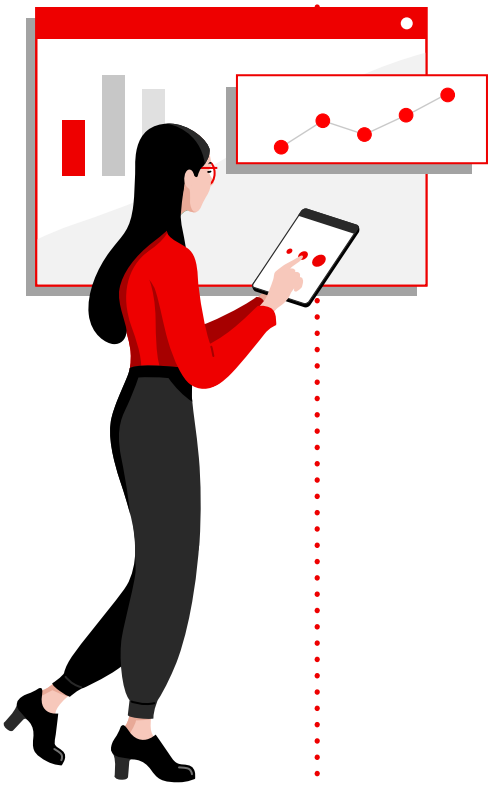
Speed is critical to any organization's success. Faster time to market, faster feature delivery, and faster customer service can make all the difference. Many organizations are also looking to accelerate at the edge of the network, especially when it comes to decision making.

The need for speed is a result of many organizations looking for ways to innovate and deploy new business models at the edge more efficiently. Accelerating decision making at the edge allows organizations to respond to changes in the market, make strategic shifts in their offerings, and gain an advantage over the competition—all of which can strengthen their bottom line.

Consider how a retail store might enhance their customer experience by offering store-specific information via a mobile application. The application might allow customers to see where products are located, check inventory information, or access timely sales and discounts. All of these capabilities require instantaneous data processing–if you show a coupon to someone after they've walked out of the aisle, you're too late.

From a technology perspective, this means processing data locally, near the source of data generation, rather than relying on centralized cloud-based systems. Edge computing makes this shift possible, reducing latency and allowing for real-time decision making, which results in improved efficiency.

However, implementing and running edge environments can pose challenges that are different from those found in traditional on-premise or public cloud environments.



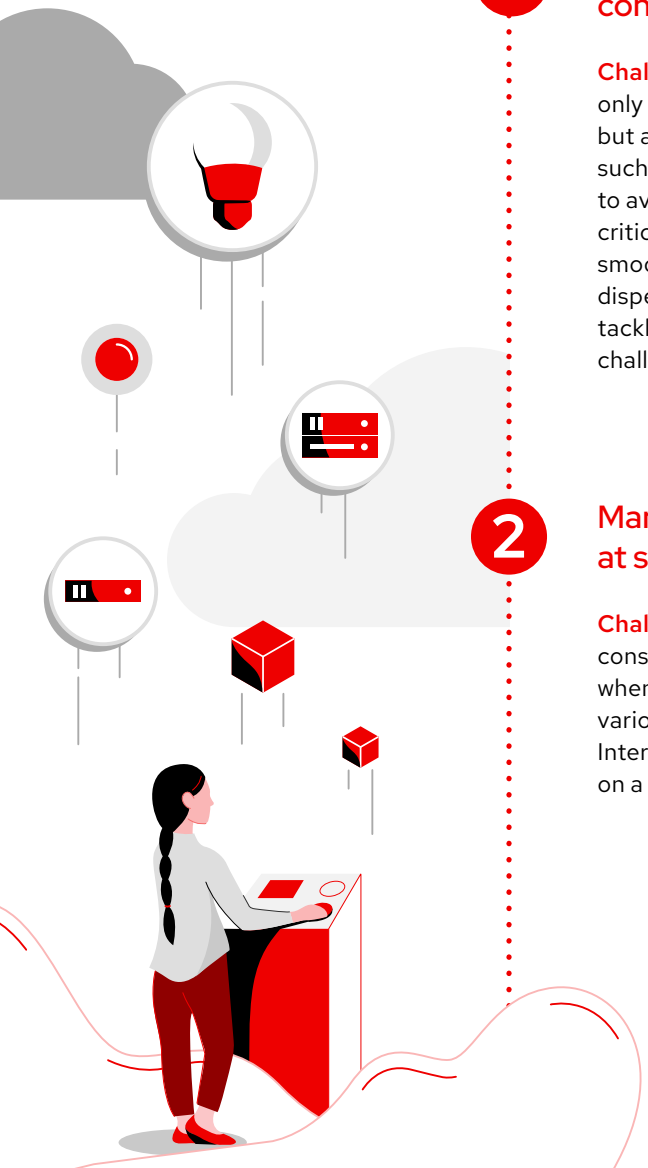
This e-book explores key challenges that can slow decision making, suggests how to overcome these hurdles, and provides insight into how artificial intelligence (AI) and machine learning (ML) can advance your organization's edge capabilities.

Chapter 1

Identify and overcome challenges at the edge

Edge devices can have many variables that can be as unique as your organization. Whether deployed across hundreds of retail stores, in motion on a commercial airplane, or on a factory floor, every organization must overcome obstacles at the edge. From complexity to automation to a lack of standardization to security, planning for inevitable challenges can help your organization get the most from your edge deployments.

No matter what your industry, as your organization designs and deploys edge computing architecture, it's important to address a number of variables. Here are 4 factors to consider when planning your edge deployments.



1 Maintain operational consistency despite limited connectivity and bandwidth

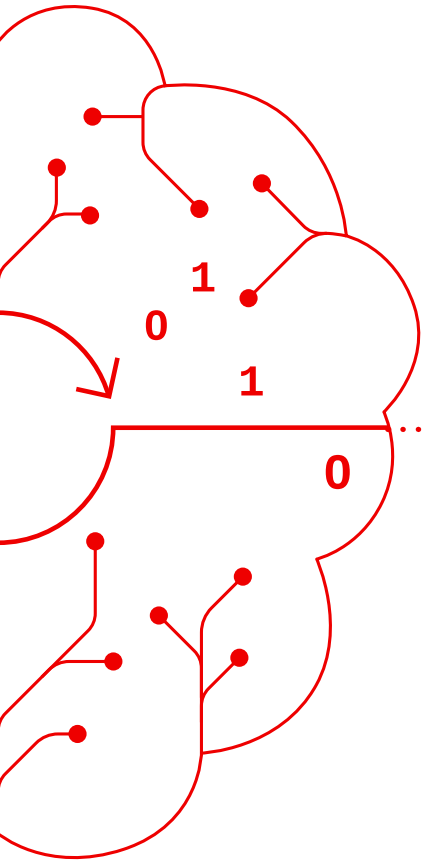
Challenge: Unreliable connections not only degrade customer experiences but also pose significant security risks, such as slow or lost data, and can lead to avoidable costs. To make sure your critical and remote operations run smoothly, especially in geographically dispersed locations, it's essential to tackle connectivity and bandwidth challenges head-on.

Solution approach: By processing data locally, you can minimize the impacts of latency and connectivity disruptions. Focus on edge solutions that help you safeguard your operations and optimize performance, even in the most challenging environments.

2 Manage applications and accelerate deployments at scale anywhere

Challenge: It can be difficult to operate consistently and scale successfully when managing deployments across various topologies and infrastructure. Interoperability for edge deployments on a variety of hardware is essential.

Solution approach: To manage applications, reduce errors, and scale more quickly at remote locations, aim to eliminate manual processes where possible by embracing automation. This approach helps to promote consistent responses that can boost quality control, create positive consumer experiences, and prevent downtime in critical operations. Enterprise automation paves the way for modernization and digital transformation.



3 Realize more efficient decision making and resource utilization

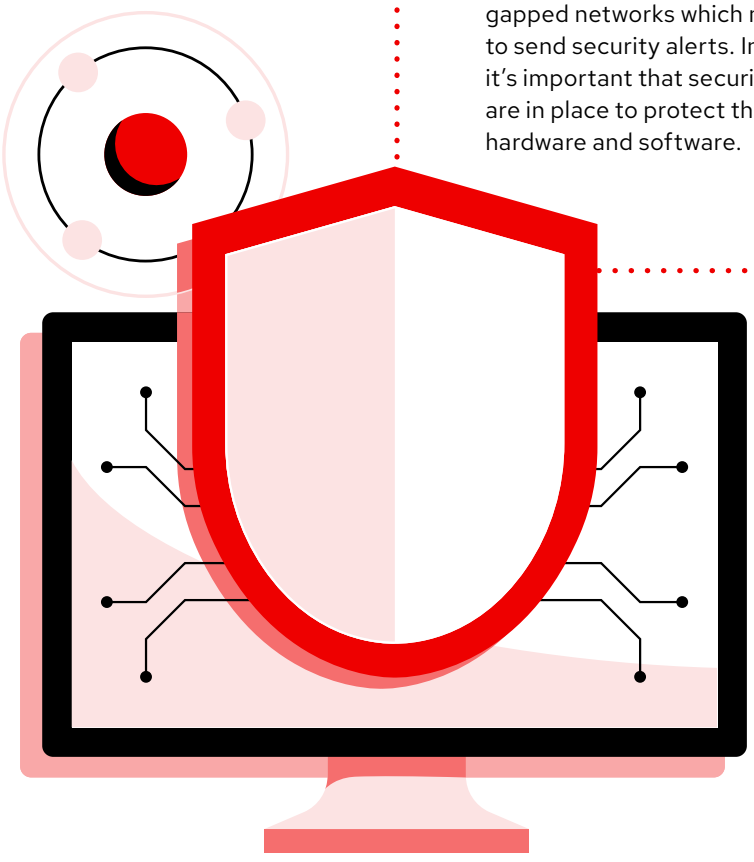
Challenge: Data can be a differentiator for many organizations. However, organizations are burdened with the slow speed at which their data is being sent and received, delaying the resolution of critical business issues and resulting in slower time to market and increased operational costs.

Solution approach: Adopt an automation approach to help accelerate data-driven operations in real time. Analyze data, reduce latency and bandwidth limitations, and deploy edge applications to allow teams to make decisions more quickly and maximize resource allocation. A unified hybrid cloud platform can help create consistency across operations and application development and lifecycle management. This includes the ability to perform on-site data processing on edge devices often optimized for AI/ML modeling, which leads to instantaneous decisions.

4 Enhance security and comply with data sovereignty and regulatory requirements

Challenge: Organizations are rapidly implementing edge computing, often in locations without IT staff or the physical security of a datacenter. This means that edge devices can be located in publicly accessible locations or in air-gapped networks which may be unable to send security alerts. In these cases, it's important that security practices are in place to protect the device hardware and software.

Solution approach: Security teams need sufficient visibility to predict, detect, and proactively address risks. Organizations can improve security for edge devices and environments by setting up controls and policies to maintain proper security posture, governance, and compliance and prevent intermittent connectivity from disrupting ongoing operations.



Addressing these considerations are just the beginning. For organizations looking to improve their agility and speed at the edge, the right technology solution can not only help overcome these challenges, but help teams take the next steps to make more informed decisions at the edge.

Chapter 2

The advantage of AI at the edge

AI and edge computing each offer benefits to an organization, from improved efficiency to greater access to data and insights to enhanced security. However, the advantage of extending these capabilities to distributed computing environments and edge sites can be enormous for industries such as manufacturing, retail, transportation, logistics, smart cities, and more.

For example, in a factory that uses heavy machinery, where downtime can result in significant financial losses, edge computing allows for predictive maintenance that can reduce unplanned downtime. Traditionally, the data collected from machinery would be sent to a central server for analysis, which could delay the identification of potential issues. With edge computing, data from sensors that monitor vibrations, temperature, and other critical parameters can be processed locally in real time.

Across industries and markets, customers are using AI at the edge to be more efficient. Customers such as OMRON began testing Red Hat® OpenShift® in manufacturing to allow real-time data to more easily flow throughout the organization, and in the defense industry, Lockheed uses Red Hat Device Edge to allow models on drones to be changed or replaced to respond to mission needs.

Bringing AI to the edge

The introduction of AI allows organizations to rapidly analyze vast amounts of data generated by sensors on various equipment to identify patterns and anomalies that may lead to failures.

By processing data at the edge, where the machinery operates, AI can provide real-time insights and predictive analytics, allowing for immediate action to prevent downtime and reduce maintenance costs. In the event that deviations from normal operating conditions are detected, such as an unusual vibration pattern that might indicate a bearing failure, it can instantly alert maintenance teams.

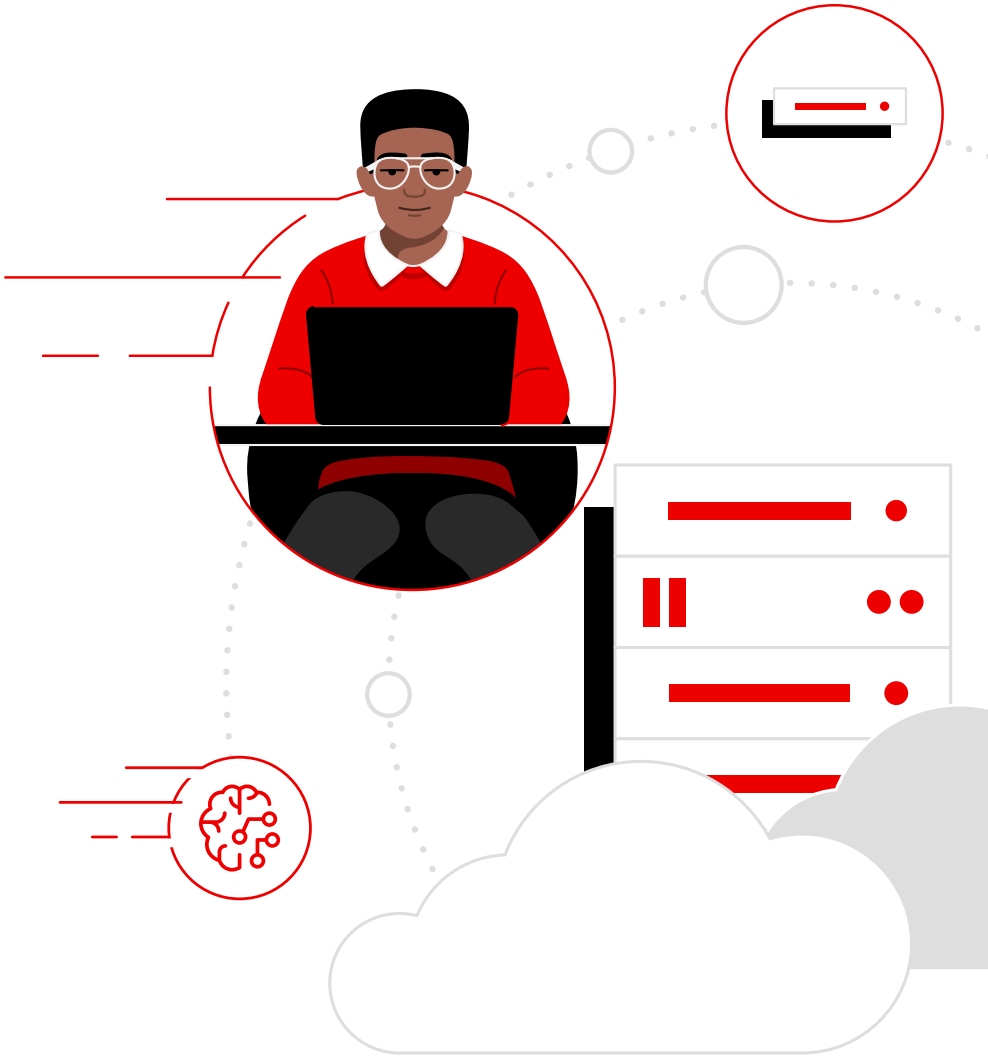
Local intelligence can also prioritize alerts based on the severity of the issue so that the most critical problems are addressed immediately, the right teams are contacted, and processes are initiated without delay.

What does this mean for AI-enabled applications at the edge? In order to successfully run AI at the edge, many organizations may need to modernize their edge IT resources and processes to support the demands of AI applications. Automation can help address these needs with greater speed and consistently.

Move away from your monolith to make the most of AI at the edge

Bringing AI to edge computing environments removes the cost and latency of moving data back and forth from edge devices to centralized cloud computing. To get to this point, it's critical that organizations break free from system architectures that may be decades old or that aren't interconnected.

Some organizations may find that they are restricted by their aging systems and are unlikely to have the ability to take advantage of technology such as AI/ML. In this case, container technology can help develop and deploy these new applications at the edge. Containers are lightweight and portable, and run efficiently with a focus on security across a variety of devices and platforms. Their modular nature provides the flexibility developers require to quickly and efficiently iterate on applications.



Add flexibility to make decisions more quickly

Kubernetes container orchestration is an important tool for the successful deployment of AI at the edge. Both single node Red Hat OpenShift and compact clusters (3-node OpenShift) can help organizations make rapid decisions at the edge by providing flexible and scalable deployment options tailored for edge environments. In addition to flexibility, Red Hat OpenShift brings much-needed consistency through a single application platform that allows developers to write an application once, for deployment anywhere, and allows operations teams to manage a single environment, from core to cloud to edge.

Compact clusters, or 3-node OpenShift deployments, provide a balanced approach by offering high availability and redundancy in a compact configuration. In a 3-node setup, Red Hat OpenShift runs on 3 nodes that can manage workloads collectively, providing resilience and continuous operation even if 1 node fails. This configuration is particularly beneficial for edge locations that require a robust and reliable infrastructure to support critical applications and data processing.

Single node OpenShift is a standard configuration of a Red Hat OpenShift cluster that consists of a single control plane and worker node that brings all the features, benefits, and consistency of Red Hat OpenShift to an even smaller footprint. It is designed for environments with limited space and resources, allowing organizations to deploy Red Hat OpenShift on a single physical or virtual node.

This compact deployment is ideal for remote locations where deploying a full-scale cluster isn't feasible. By running Red Hat OpenShift on a single node, organizations can bring Kubernetes-based container orchestration and management capabilities to the edge. This configuration allows for the local processing of data and execution of AI/ML workloads, minimizing latency and supporting real-time decision making.

Orchestrate containers at the edge using Kubernetes

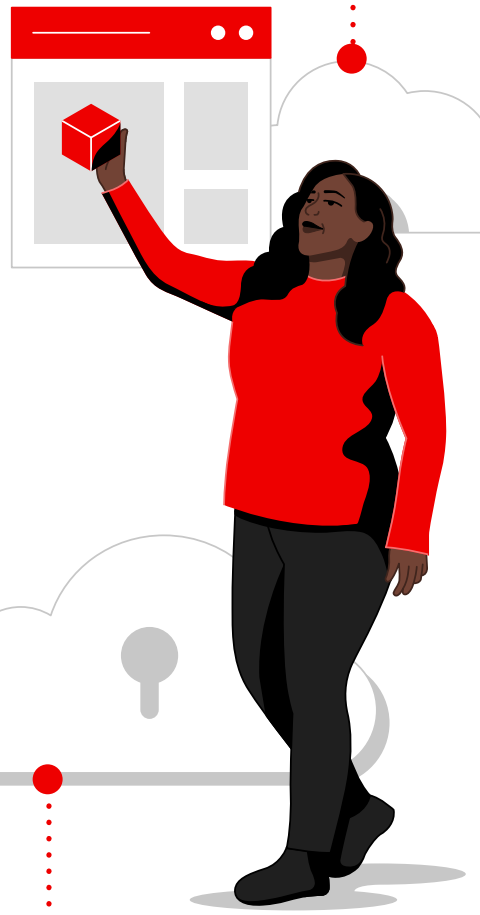
Need to go even smaller? Red Hat Device Edge includes the Red Hat build of MicroShift, a lightweight Kubernetes container orchestration solution derived from Red Hat OpenShift that uses a consistent Kubernetes API to extend operational consistency and scalability for hybrid cloud deployments to the furthest reaches of the edge. MicroShift can be placed in highly space-constrained locations and run on extremely low-power hardware while allowing teams to use familiar tools and processes.

Red Hat Device Edge brings a new level of scalability and consistency to cloud-native environments. It allows applications to run where they are most useful—at the edge. Whether your application needs to be attached to an assembly line, building products, in a retail store, or onboard ships at sea, applications can comfortably operate on small, efficient hardware from a vast ecosystem of partners while handling virtually any conditions.

Mix and match Red Hat OpenShift compact clusters, single nodes, and Red Hat Device Edge to create a bespoke environment that can address any edge situation and benefit from a single, agile, consistent platform from core to cloud to edge.

Chapter 3

Red Hat's approach to accelerating decisions at the edge



Organizations across all industries are overseeing an increasing number of computing locations, some with intermittent connectivity and limited physical access. The explosion of data volume from these edge-based devices makes it challenging to turn the steady stream of information into actionable insights.

Extending AI to the edge can address this challenge by improving visibility and accelerating decision making. By providing a consistent IT foundation, Red Hat helps organizations boost operational consistency, even in instances of limited connectivity, to deliver insights and experiences at the moment they're needed.

Observability, monitoring, and platform engineering tools

Red Hat Enterprise Linux® AI offers a stable and security-focused foundation for running AI workloads at scale. It supports latency-sensitive applications so AI models can operate efficiently on a variety of hardware, from edge devices to central datacenters.

This flexibility allows organizations to deploy AI/ML applications across their entire infrastructure with consistent performance and reliability. This consistency helps organizations use data intelligence more efficiently, which translates to optimized operations, enhanced customer experiences, and the ability to offer better products and services to the end customer.

Red Hat OpenShift AI further enhances decision making by providing a robust platform for developing, training, and deploying AI/ML models. This platform integrates with edge computing environments, allowing AI models to run efficiently on edge devices. By using Red Hat OpenShift AI, organizations can gather and process vast amounts of data at the edge, generate actionable insights, and make data-informed decisions.

Red Hat OpenShift brings the consistency of a unified platform to the edge, speeding up development and allowing teams to scale with their environment. With the power of Red Hat OpenShift, a hybrid cloud can seamlessly extend all the way to the edge, bringing innovation where it is needed most.

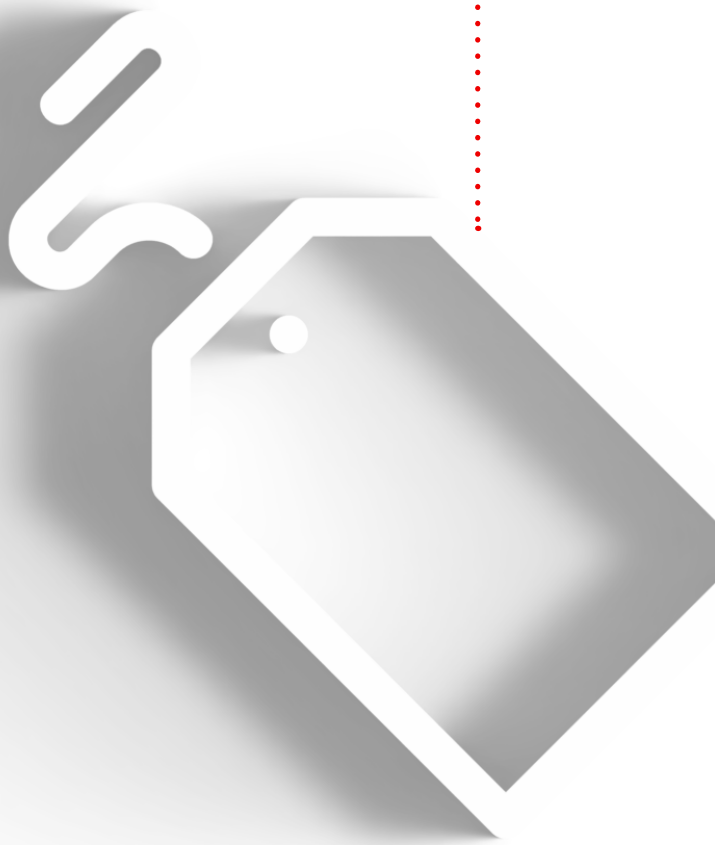
Red Hat Device Edge allows organizations to deploy lightweight and scalable applications on edge devices so that data can be processed close to its source. This local data processing reduces latency and bandwidth usage, which is crucial for real-time analytics and decision making. Organizations using edge computing can help maintain continuous operations even in environments with limited connectivity.

Red Hat Ansible® Automation Platform provides an enterprise framework for building and operating IT automation at scale, including deployment and management of AI/ML models and data. Available as part of the platform is Event-Driven Ansible, which can help organizations quickly automate common tasks based on real-time events or triggers. Organizations can predefine desired responses to accelerate incident resolution when certain conditions occur.

Chapter 4

Explore common use cases for AI at the edge

To better understand how Red Hat helps accelerate informed decision making at the edge, let's take a look at a few examples across 3 industries.



Retail

Edge locations such as retail stores are critical points of revenue generation and customer interaction. Simple, customer-centric functionality is essential.

1

Enhance the customer experience

Red Hat Device Edge can power in-store customer experiences such as an augmented reality (AR) mirror that lets shoppers try on virtually anything in the store, without spending extended time in a dressing room.

2

Streamline back office and inventory functions

Catalog and pricing systems powered by Red Hat OpenShift can help run essential business functions, streamlining inventory and ordering to make sure that high-demand items are always available when needed.

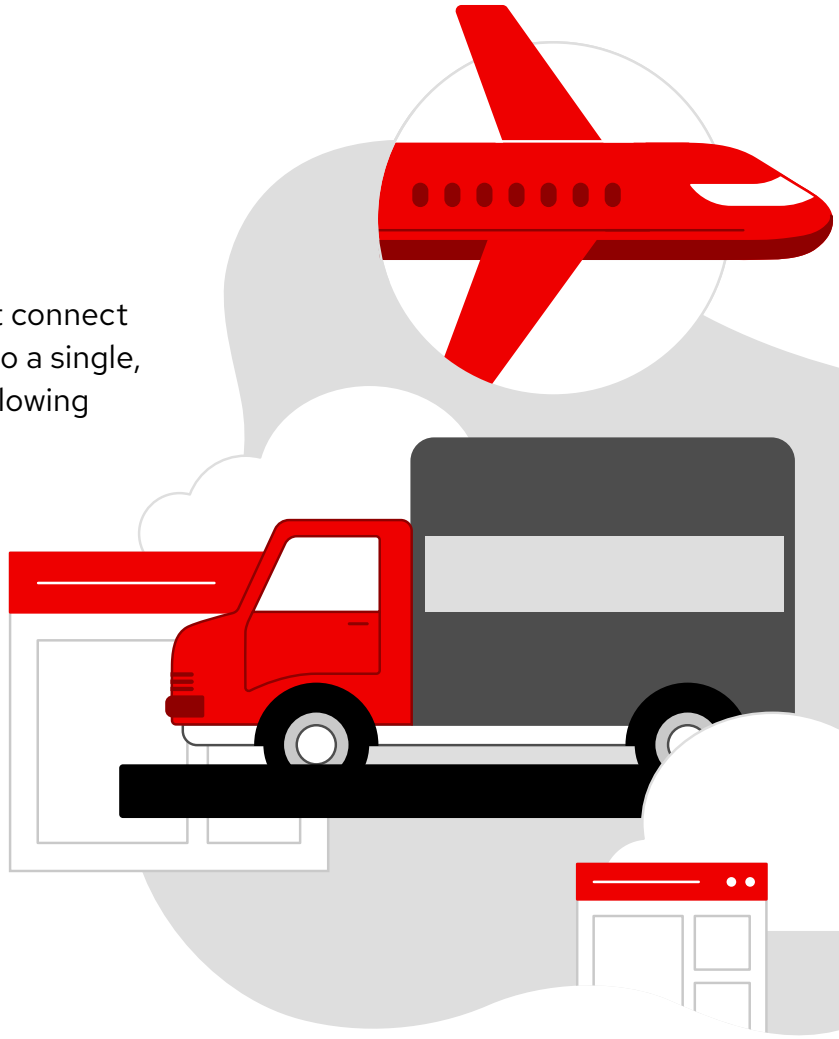
3

Power the point of sale

Everything leads to the point of sale. Red Hat Device Edge allows for rapid transaction processing, improved customer experiences, and reliable system performance without reliance on centralized datacenters. Network connectivity is essential for self-service solutions, point of sale, and payment systems and applications. Ansible Automation Platform can be used to automate networks end-to-end and orchestrate network processes.

Transportation

From roads and railways to intermodal sites that connect trains and trucks, edge computing contributes to a single, fluid supply chain that reliably keeps inventory flowing from the source to customers.



1

Chart the way for wayside devices

Red Hat Device Edge keeps remote transportation connected back to head office for logistics monitoring, on-board cargo status updates, and more.

2

Coordinate to accelerate transportation

Red Hat OpenShift can power multimodal transportation management to create the most efficient scheduling, planning, and optimization.

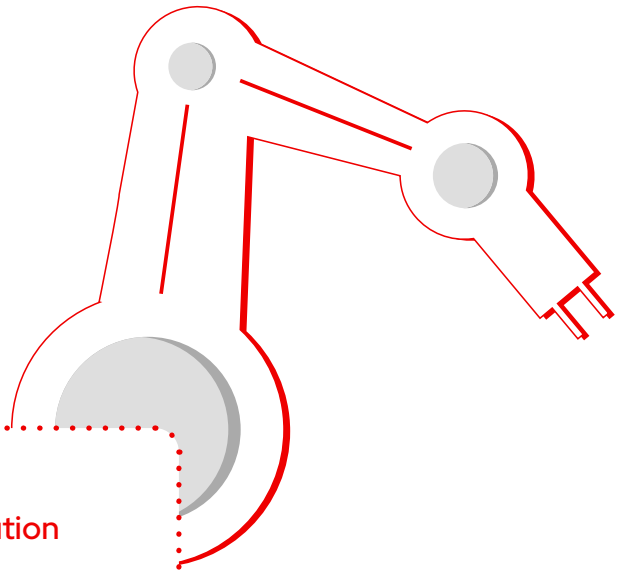
3

Keep safety and security top of mind

Red Hat Device Edge helps to protect people and resources with on-site video AI/ML for fast, private monitoring.

Manufacturing

When downtime is not an option, AI at the edge can help make sure the factory, plant, or assembly line is always running on time.



1

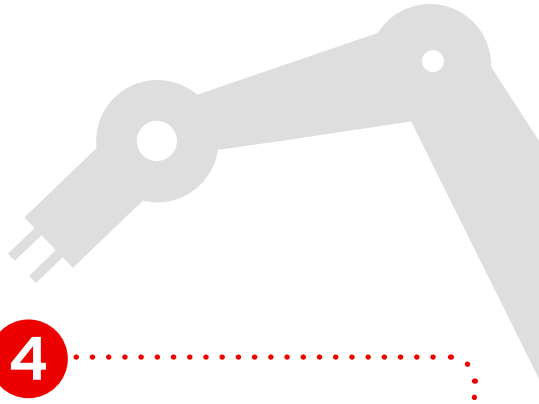
Plan ahead by predicting equipment failure

Predict and prevent equipment failure using Red Hat Device Edge to proactively catch anomalies before they become issues. For example, detection of irregular performance could trigger an event to Event-Driven Ansible, allowing for self-healing.

2

Manufacturing execution systems (MES)

Red Hat OpenShift powers the high-availability systems that keep production on schedule. Ansible Automation Platform helps to manage the configuration of manufacturing execution systems.



3

Level up your product inspection

Put AI/ML directly on the assembly line to immediately detect errors in production.

4

Roll out autonomous guided vehicles (AGV)

Red Hat Device Edge navigates AGVs to keep your inventory supply chain moving continuously and efficiently.

Discover how edge computing is revolutionizing these industries and more; explore Hatville, the miniature city where edge computing comes to life.

Chapter 5

Protect your edge environment with Red Hat's layered security approach

Use of edge computing has accelerated for many organizations across all industries. A critical aspect of any edge and AI adoption strategy is security, including how systems and their vast stores of business-generated data can be kept safe from threats. Securing edge deployments can become especially complex as business data is processed in much higher volumes in more remote places, far from the traditional security perimeters.

Red Hat provides trusted open source software that helps organizations implement a consistent, layered security approach across their hybrid cloud infrastructure, application stacks, and product and business process lifecycles.

Adopt a zero trust approach

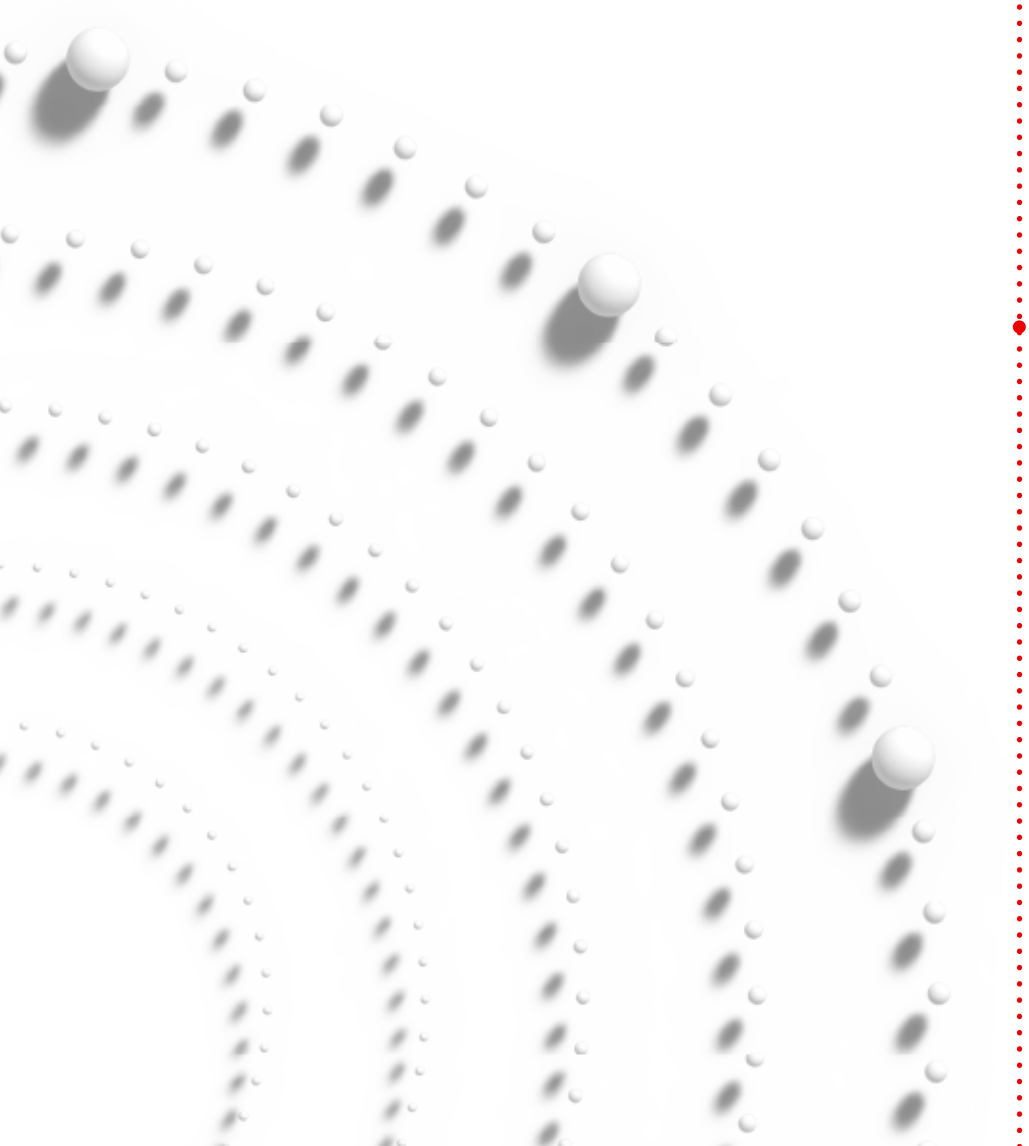
Red Hat employs a zero trust security approach, which means that every interaction within the system is based on the principle that no entity, whether inside or outside the network, is inherently trusted. By applying this model across the entire suite of platforms, tools, and managed cloud services, Red Hat enhances security for both proprietary and third-party applications deployed by customers and partners. This approach includes rigorous application isolation, authentication, and encryption processes to meet Federal Information Processing Standards (FIPS), closing security gaps and preventing unauthorized access.

Deploy automation and intelligence at the edge

Red Hat helps organizations mitigate security risks and reduce the associated costs through automated security assessment and immediate remediation of issues. Red Hat's platforms support the journey toward DevSecOps, integrating security into every phase of the development lifecycle.

Red Hat Device Edge and Red Hat OpenShift AI allow local data processing to reduce latency and promote immediate analysis and response to potential threats. Red Hat also provides continuous security updates and patches to address vulnerabilities promptly, minimizing vulnerabilities.

Ansible Automation Platform provides a proactive approach to identifying and remediating security vulnerabilities, reducing downtime and operational costs. This approach to security management is critical in maintaining a security-focused and resilient edge computing environment.



Protect the software supply chain

Red Hat's security focus applies to the entire software supply chain, and includes rigorous provenance checks, image and vulnerability assessments, extensive quality assurance, and regression testing.

The sum of these parts provides a comprehensive approach to protecting edge deployments against threats and helps organizations move toward a true end-to-end DevSecOps approach. This comprehensive security framework supports organizations in maintaining governance and compliance, even in highly distributed edge computing environments.

Learn more

Discover how Red Hat can help your organization gain greater visibility, boost security, and accelerate decisions at the edge.

> **Read** Red Hat: Transform your organization's edge operations with AI here.

North America
1 888 REDHAT1
www.redhat.com

Europe, Middle East, and Africa
00800 7334 2835
europe@redhat.com

Asia Pacific
+65 6490 4200
apac@redhat.com

Latin America
+54 11 4329 7300
info-latam@redhat.com

facebook.com/redhatinc
@RedHat
linkedin.com/company/red-hat

redhat.com