

Red Hat/Cisco AI Joint Messaging Guide

Updated June 2025

Clear path. Open road.



Guide overview

How to use this document

Context

Messaging

Additional information

How to use this document

This document presents a unified message about the Red Hat/Cisco partnership and joint AI capabilities.

The information in this messaging guide will help stakeholders at Red Hat, Cisco, and Channel partners understand our target customers' challenges and how to answer the common question, **“Why should I turn to Red Hat and Cisco for AI?”**



Context

Market context

Partnership overview

Joint solution overview

Audience overview

Audience challenges
and considerations

Messaging

Additional information

Market context

With use cases spanning chatbots and virtual assistants to code generation to fraud detection systems, the use of Artificial Intelligence (AI) is gaining momentum across all industries. As a market expected to exceed \$423 billion by 2027 with a 5-year CAGR of 26.9%¹, IDC found that the top business initiatives/objectives for AI included:

- Improving operational efficiency
- Improving customer experience and customer satisfaction
- Improving employee productivity

In this fast-moving market, companies are looking for stable, production-grade AI solutions that they can deploy in their own organizations. In fact, according to the Cisco AI Readiness Report, 97% of global organizations reported an urgency to deploy AI-powered technologies.

But while many want to harness the potential of AI, few are ready to do so.

1. IDC FutureScape Webcast: [Worldwide Artificial Intelligence and Automation 2024 Predictions](#)

2 Gartner: Infrastructure and Operations Summit, November 2023

3 Cisco Global AI Readiness Index, December 2023

Few are ready to harness the potential of AI



of CIOs and tech leaders plan full GenAI adoption within 3 years²



say IT staff/skill recruitment and training is their highest priority to support GenAI²



of IT organizations with infrastructure that are prepared for AI today³

Guide overview

Context

Market context

Partnership overview

Joint solution overview

Audience overview

Audience challenges
and considerations

Messaging

Additional information

With AI adoption spanning many levels of maturity, customers are looking for options that help them get started while also preparing them for full-scale deployments. Organizations are on the lookout for solutions that provide:

- Efficient use of IT resources through monitoring (e.g., resource utilization and model efficacy), resource optimization (e.g., GPU usage), and end-to-end governance;
- Freedom to choose their model approach (build their own or tune foundation models) so that they can use their own data to drive differentiation in both predictive and generative AI use cases;
- Gen AI capabilities that help their teams increase productivity while addressing potential skills gaps; and
- Flexibility to choose where data is stored when training and running models while ensuring that security, regulatory, and compliance requirements are met.



Context

- Market context
- Partnership overview**
- Joint solution overview
- Audience overview
- Audience challenges and considerations

Messaging

Additional information

Partnership overview

For more than a decade, Red Hat and Cisco have partnered to solve the most challenging, real-world problems that developers, cloud administrators, and IT operations teams face in the data center and at the edge. Both companies offer proven, transformative solutions that meet the business and technical needs of enterprise-grade organizations worldwide.

- With more than 20 years of innovation in open hybrid cloud, Red Hat is the world’s leading provider of enterprise open-source solutions, including market-leading Linux, cloud, automation, container, and Kubernetes technologies.
- Cisco is revolutionizing the way infrastructure and data connect and protect organizations in the AI era – helping customers build AI-ready data centers that power all AI workloads anywhere, future-proofed workplaces, and digital resilience.

Cisco and Red Hat build, test, and validate solutions to transform the customer experience, and the entire solution stack can be found on the **Cisco Global Price List**, indicating the deep collaboration and trust between the two companies. And, with more than 20 full-stack [Cisco Validated Designs with Red Hat technology](#), enterprise-grade organizations can securely and seamlessly accelerate their IT modernization journey, enable development and deployment, manage cloud-native applications at scale, and see faster time-to-market and -value with reduced risk – all backed by Cisco Solution Support.

4. Source: IDC white paper, sponsored by Red Hat: ["The Business Value of Red Hat OpenShift"](#), March 2021

Red Hat⁴



Cisco



Joint solution overview

Organizations seeking AI solutions frequently know **that** they should incorporate AI somehow, but they may not know **what** they need or **where** and **how** to get started. The challenges of public cloud, including security of data and cost, make on-premises AI a viable option that offers control, compliance, and cost-effectiveness.

With a joint solution that brings together powerful AI capabilities, networking, security, observability, collaboration, and automation, organizations can streamline deployment and management of AI, freeing their teams to innovate while also preparing for future growth and scale.

1

Red Hat OpenShift

Red Hat OpenShift, the leading hybrid cloud application platform, is trusted, comprehensive, and provides a consistent experience on-prem, in the cloud, and at the edge. With OpenShift as the security-first application development and delivery platform, organizations and IT teams can seamlessly integrate the AI solutions that align with their business goals.

- **Build AI applications where and how you want.** Consistently build, deploy, and manage applications running on containers, virtual machines, and serverless architectures. Modernize existing infrastructure investments at scale, with the flexibility to run applications across data centers, public clouds, and edge environments.
- **Increase productivity with a comprehensive app platform for cloud-native AI innovation.** Improve developer productivity, automate CI/CD pipelines, and shift security left. Address modernization and deployment hurdles and boost operational efficiency with extensive tooling and services.
- **Rest easy with comprehensive security built in.** Control the trusted sources of content, testing, and deployments. Defend applications from attacks and vulnerabilities in all layers of the platform. Enable and extend secure services through standard interfaces and APIs.

Guide overview

Context

Market context

Partnership overview

Joint solution overview

Audience overview

Audience challenges
and considerations

Messaging

Additional information

2

Red Hat AI

Red Hat AI brings together the power of open source, the flexibility of the open hybrid cloud, a robust partner ecosystem, and experience working with our customers to modernize applications using AI. With Red Hat solutions like **OpenShift AI** – an MLOps platform built on our proven open hybrid cloud platform OpenShift – and third-party AI compatibility, we empower enterprises to increase productivity, create efficiencies, and provide unique customer experiences.

- **Increase efficiency with fast, flexible, and efficient inferencing.** See faster and more cost-effective model deployments with optimized model inference across the hybrid cloud. A vLLM runtime maximizes throughput and minimizes latency, an optimized model repository accelerates model serving, and an LLM compressor helps reduce compute utilization while preserving accuracy.
- **Enhance customization with simplified and consistent experience for connecting models to data.** Improve the accuracy and relevance of model response with a simplified and consistent experience that connects models to data and enables efficient model customization to deliver on domain-specific use cases.
- **Maximize flexibility and consistency when scaling AI across the hybrid cloud.** Choose where to train, tune, deploy, and run models and generative AI applications to facilitate compliance with data constraints, privacy, and security while controlling costs (GPUs, cloud, etc.). Experience operational consistency and stability; deploy, run, and operate models anywhere.
- **Accelerate Agentic AI innovation.** Leverage an agile, stable platform that features a unified API experience and out-of-the-box components that accelerate development and deployment of AI agentic workflows.



Guide overview

Context

Market context

Partnership overview

Joint solution overview

Audience overview

Audience challenges
and considerations

Messaging

Additional information

3

Red Hat Ansible Automation Platform

80% of surveyed business executives say that adopting IT automation is “extremely important” to the future success of their organization.⁵ Red Hat Ansible Automation Platform establishes a solid foundation for AI implementation by simplifying the deployment, management, configuration, and lifecycle of AI models and infrastructure components.

- **Reduce configuration errors or discrepancies.** Use infrastructure as code as a consistent and repeatable method for deploying AI components like servers, models, and data.
- **Prevent disruptions to AI operations.** Connect seamlessly with monitoring and alerting tools, allowing your IT operations teams to proactively identify and address potential issues.
- **Enhance data security and accessibility.** Move data from servers to storage in region, while also making sure that data is accessible to the correct users for training models.



5. Harvard Business Review Analytic Services, sponsored by Red Hat. [“Taking the lead on IT automation: IT leaders as evangelists for their automation strategies,”](#) January 2022.

Guide overview

Context

Market context

Partnership overview

Joint solution overview

Audience overview

Audience challenges
and considerations

Messaging

Additional information

4

Cisco AI PODs

Cisco AI PODs are a full-stack AI-ready system offering a suite of foundational networking, compute power, and storage; network fabrics; accelerated GPUs and CPUs; infrastructure and automation; AI tooling; security; observability capabilities; and Red Hat OpenShift – the industry’s leading container orchestration platform powered by Kubernetes.

- **Deploy with confidence.** Take advantage of AI-ready infrastructure with pre-designed, validated, full-stack architecture bundles for targeted AI use cases.
- **Get started quickly.** Leverage automation frameworks for rapid deployment and adoption of infrastructure.
- **Get the support you need.** Operate with a best-in-class, single-support model for your AI deployment architecture, including enterprise support for Operations Support System tools and libraries.
- **Currently available solutions.** These include Cisco FlexPod AI with NetApp and FlashStack for AI with Pure Storage, with more powerful partnerships coming soon.

Cisco and Red Hat provide fully tested and validated blueprints to build AI solutions that make sense for each company’s unique situation. With [Cisco Validated Designs \(CVDs\)](#) supporting the joint AI solutions, organizations can take the risk out of AI adoption, boosting security, easing implementation, reducing deployment time, and improving total cost of ownership (TCO). It’s what savvy companies need in today’s hyper-competitive landscape: proven partners who will work with them to build their own AI-ready data centers, future-proofed workplaces, and digital resilience.



Audience overview

Our audiences are broken into two categories:



Primary personas

These are the IT decision makers who are responsible for making or influencing final purchases. **As high-level executives, they are primarily influenced by messaging about the business impact of a solution – the *why* behind the *what*.**



Secondary personas

These are the IT influencers who provide strategic guidance and/or recommendations to the decision makers. **They are primarily hands-on-keyboard engineers and leaders who need to understand how a solution works and how it will affect their day-to-day work.**

Context

- Market context
- Partnership overview
- Joint solution overview
- Audience overview
- Audience challenges and considerations

Messaging

Additional information



Primary (IT decision makers)

Title	AOR	Purchase Responsibility	Involvement
Chief Executive Officer (CEO)	Leads overall company strategy & approves major expenditures	Sets strategic direction, approves major budgets	Final Approver (potentially Key Influencer)
Chief Artificial Intelligence Officer (CAIO)	Leads company AI strategy, development, & implementation	Identifies gaps & opportunities, provides technical expertise & leadership	Key Influencer & Approver
Chief Information Officer (CIO)	Leads IT strategy, aligns with business goals, & drives innovation	Provides strategic guidance	Key Influencer & Approver
Chief Technology Officer (CTO)	Oversees technology strategy & development	Provides technical expertise & leadership	Key Influencer & Approver



Secondary (IT influencers)

Title	AOR	Purchase Responsibility	Involvement
Chief Financial Officer (CFO)	Ensures financial stability and profitability	Evaluates financial impact & provides input	Key Influencer
Chief Operating Officer (COO)	Approves major expenditures & oversees operations	Ensures cost-effectiveness, operational alignment	Key Influencer & Approver
VP Infrastructure & Ops	Establishes & maintains IT infrastructure strategy	Manages IT systems & networks	Key Influencer & (potentially) Approver
VP/Director of IT	Oversees day-to-day operations & management of IT systems	Manages procurement process	Key Influencer & (potentially) Approver
Software Engineer/DevOps Engineer	Writes and maintains code, deploys & manages applications	N/A	Key Influencer
Data Scientist	Analyzes data, builds & trains models, extracts insights	N/A	Key Influencer

Audience challenges and considerations

Our target audiences have multiple challenges to navigate when evaluating their future with AI.

Specific persona-based challenges and considerations include:

Title	Focus	Challenges/Considerations
Chief Executive Officer (CEO)	Overall business impact, financial implications, alignment with long-term goals	- Long-term return on investment (ROI) - Aligning solution purchases with company vision and strategy - Balancing financial risks with potential benefits
Chief Artificial Intelligence Officer (CAIO)	AI strategy, talent acquisition, ethical AI, vendor selection	- Securing budget - Building AI talent - Managing ethical risks - Integrating AI into business operations
Chief Information Officer (CIO)	IT strategy, business alignment, cost optimization, security, compliance	- Securing data & ensuring compliance with regulations - Managing & optimizing costs - Selecting vendors that align with strategic goals
Chief Technology Officer (CTO)	Solution performance, technology alignment, business strategy	- Selecting the most suitable technology solution - Ensuring solution performance & security - Aligning upgrade with long-term technology roadmap

Context

- Market context
- Partnership overview
- Joint solution overview
- Audience overview

Audience challenges and considerations

Messaging

Additional information

Specific persona-based challenges and considerations include:

Title	Focus	Challenges/Considerations	
VP Infrastructure & Ops	Infrastructure compatibility, performance, security	- Incorporating new tech while protecting existing investments - Ensuring security while driving upgrades - Simplicity and efficiency	
VP/Director of IT	Technical viability, cost-effectiveness, vendor reliability	- Navigating budget constraints - Assessing vendors - Ensuring seamless integration of new technology with existing systems	
Chief Financial Officer (CFO)	Ensuring that funds are used responsibly	Proof of ROI/TCO	
Chief Operating Officer (COO)	Developer productivity, automation, code quality, CI/CD pipelines, infrastructure automation, deployment pipelines, monitoring, security	- Tool integration & compatibility with existing systems - Learning curve & potential job displacement	- Ethical implications - Data privacy & security risks - Operational overhead
Software Engineer/DevOps Engineer	Ensuring that funds are used responsibly	Financial, security, reputational risks	
Data Scientist	Data quality, model accuracy, algorithm selection, tool capabilities	- Data availability - Model performance	- Ethical implications - Explainability

Guide overview

Context

Messaging

Joint value proposition

Value propositions

Conversation starter

Copy blocks

Additional information

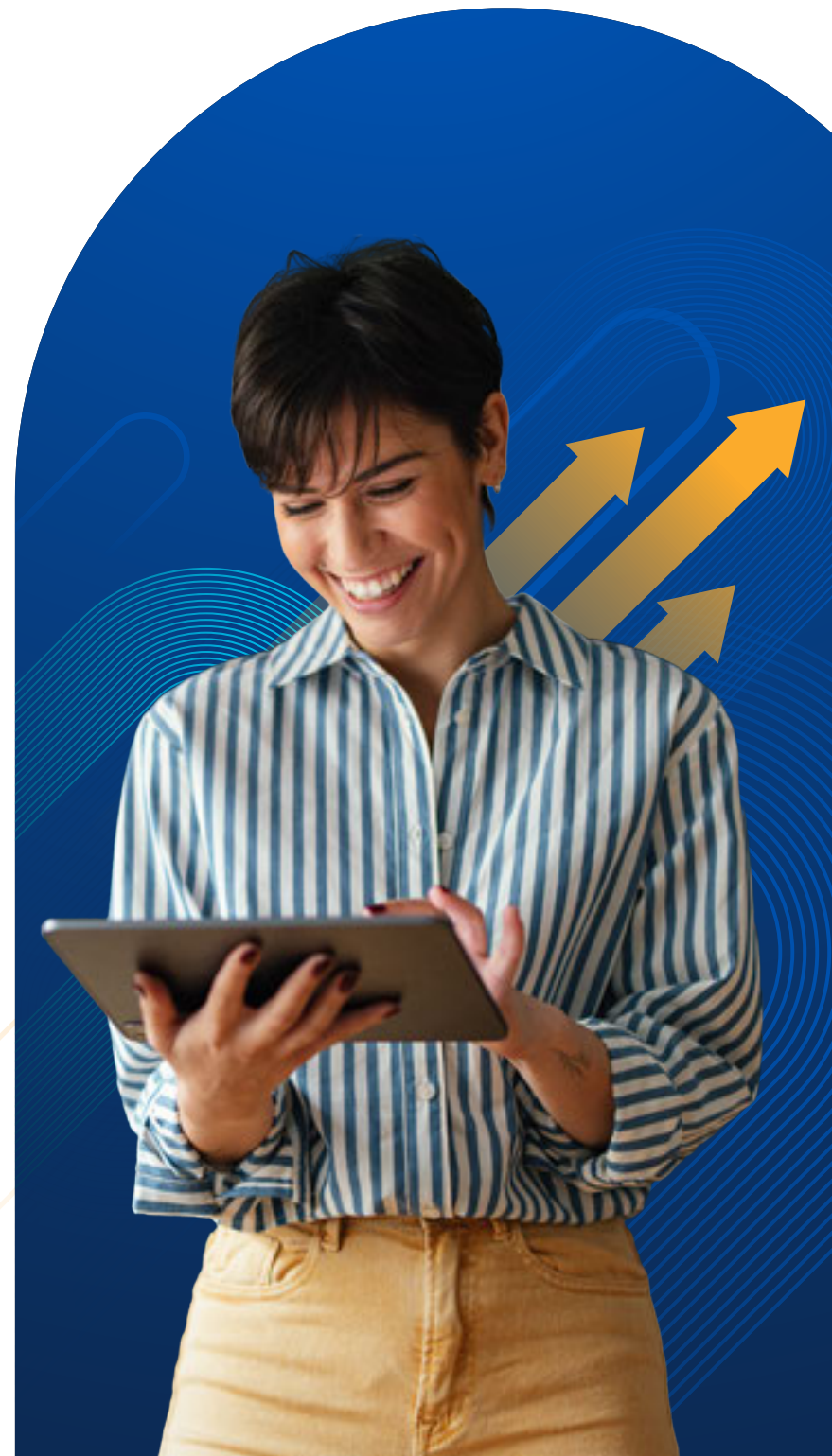
Partnership joint value proposition

Clear path. Open road.

For more than a decade, Red Hat and Cisco have partnered to build highly secure, powerful solutions that accelerate digital transformation and IT modernization. It's what you need in today's hyper-competitive landscape: proven partners who will work with you to build your own AI-ready data centers, future-proofed workplaces, and digital resilience.

Flexible, simplified management of hybrid cloud infrastructure and cloud-native applications. Virtualization that doesn't break the bank... and positions you for future growth. Automation that frees your team to innovate. And AI capabilities that can give them wings.

We've cleared your secure path to digital transformation. The open road awaits.



Key messages and value propositions



Enjoy a simplified purchasing experience.

Cisco AI PODs are easy to purchase, with pre-configured bundles based on validated designs that are tailored to specific use cases.



Streamline, simplify, and accelerate time to value.

Simplify AI infrastructure through tightly coupled resources and one-call support, operationalize deployments with validated designs and automation, and quickly adapt and deploy AI-enabled applications, moving your team from hypothesis to deployment faster than ever.



Rest easy with comprehensive security built in.

Joint Red Hat-Cisco AI solutions enhance security with a holistic approach and a secure, controlled, and auditable AI, including secure separation, device hardening, microsegmentation, encryption, and zero-trust architecture for data protection and threat mitigation.



AI your way.

Leverage Red Hat and Cisco's collaborative relationships with a diverse array of partners to build a complete AI solution – from data acquisition and preparation to monitoring and maintenance to hardware acceleration to storage – based on your unique use cases.



Increase team efficiency and productivity.

Coexist with your existing network and build, deliver and manage your own AI-enabled applications and services across any cloud, on-premise, or edge environment.



Save energy and meet sustainability goals.

See a lower physical and carbon footprint with high-density compute and storage.



Boost efficiency.

Red Hat-Cisco AI solutions make ongoing operations easy by automating resource management and providing comprehensive monitoring and alert systems.

Conversation starter

AI is a game-changing technology, and organizations who want to compete in the modern economy need to prioritize incorporating AI into their processes and applications. But how? And where do you start?

Enter Red Hat and Cisco. With joint solutions that bring together powerful AI capabilities, compute, networking, security, observability, collaboration, and automation, organizations can streamline deployment and management of AI, freeing their teams to innovate while also preparing for future growth and scale.

Increase efficiency, enhance customization, maximize flexibility and consistency when scaling across hybrid cloud, and accelerate agentic AI innovation. Leverage automation and detailed blueprints for foolproof deployment of AI platforms, reducing deployment time and ensuring that installation is done correctly and reliably. And rest easy with comprehensive security built in, including secure separation, device hardening, microsegmentation, encryption, and zero-trust architecture for data protection and threat mitigation.

With Red Hat and Cisco joint AI solutions, companies can take the risk out of AI adoption, boosting security, easing implementation, and seeing faster time to value.

We've cleared your secure path to AI adoption. The open road awaits.

Copy blocks

25 words

Together, Red Hat and Cisco offer secure, powerful, streamlined AI solutions that smooth the path to AI adoption and implementation.

50 words

Red Hat-Cisco joint AI solutions clear the path to AI adoption and implementation. Deploy with confidence, manage AI/ML workloads, simplify and operationalize AI alongside existing workloads, and rest easy with comprehensive security, reducing time-to-value, mitigating risk, and freeing teams to innovate while also preparing for future growth and scale.

100 words

With joint solutions that bring together powerful AI capabilities, compute, networking, security, observability, collaboration, and automation, organizations can streamline deployment and management of AI, freeing their teams to innovate while also preparing for future growth and scale. Take advantage of AI-ready infrastructure with pre-designed, validated, full-stack architecture bundles for targeted AI use cases.

Increase efficiency, enhance customization, maximize flexibility and consistency when scaling across hybrid cloud, and accelerate agentic AI innovation. Leverage automation and detailed blueprints for fast, foolproof deployment. And rest easy with comprehensive security built in. We've cleared your secure path to AI adoption. The open road awaits.

FAQ

What do the components of Cisco FlashStack for AI do?

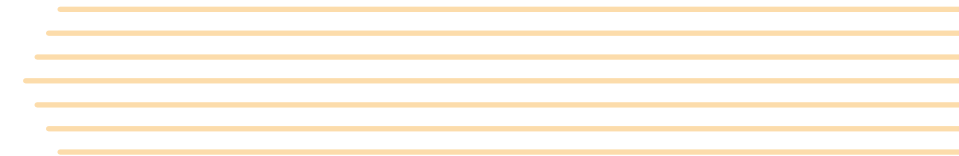
- The Cisco UCS® X-Series Modular System with X-Fabric technology allows for flexible CPU/GPU ratios, and Cisco Intersight simplifies management of traditional and AI workloads distributed across data centers and edge environments.
- Cisco Data Center Networking (DCN) AI/ML blueprint with Cisco Nexus® 9000 Series switches for enterprise networking delivers the high performance, throughput, and lossless Ethernet fabrics needed for AI/ML workloads.
- High-performance storage systems from Pure Storage provide a wide range of storage options with the scalability and efficiency that large training data sets and model serving require.

What do the components of Cisco FlexPod AI do?

- Cisco UCS X-Series Modular System with X-Fabric technology allows for flexible CPU/GPU ratios and cloud-based management for computing distributed anywhere across core and edge.
- Cisco Nexus delivers the high performance, throughput, and lossless fabrics needed for AI/ML workloads.
- High-performance storage systems from NetApp complete the picture with the scalability and efficiency that large, growing data sets demand. NetApp AI tools such as the DataOps Toolkit make it simple for developers, data scientists, DevOps engineers, and data engineers to perform various data management tasks within a Kubernetes cluster.

What does Red Hat OpenShift AI add to a Cisco FlashStack solution?

- OpenShift AI, an add-on to Red Hat OpenShift, is integrated into the FlashStack AI infrastructure to support a range of AI/ML use cases.
- OpenShift AI enables organizations to streamline and standardize their model delivery process, while providing flexibility with both integrated and custom options.
- OpenShift AI also provides pipeline automation and scalable model serving with support for inferencing servers such as Intel OpenVINO and NVIDIA Triton.
- OpenShift AI leverages the underlying OpenShift platform to provide the infrastructure resources for model development.
- The model pipeline can be integrated seamlessly with Red Hat application delivery ecosystem using GitOps pipelines making it easier for organizations to integrate these models with intelligent AI-enabled applications.



FAQ

What does Red Hat OpenShift AI add to a Cisco FlexPod solution?

- Red Hat OpenShift AI is deployed as the MLOps platform on FlexPod Bare Metal infrastructure with Red Hat OpenShift to support Enterprise AI/ML initiatives at scale.
- OpenShift AI provides a trusted, operationally consistent environment for Enterprise teams to experiment, serve models, and deliver ML-enabled applications. Data scientists, engineers, and app developers can collaborate in a single destination that promotes consistency, security, and scalability.
- OpenShift AI streamlines the ML model delivery process from development to production deployment (model serving) with efficient life cycle management and pipeline automation.
- OpenShift AI delivers enhanced support for predictive and generative AI model serving and improves efficiency of data processing and model training.
- OpenShift AI is compatible with leading AI tools and frameworks such as TensorFlow, PyTorch, and can work seamlessly with NVIDIA GPUs, to accelerate AI workloads.

What capabilities does Red Hat OpenShift add to Cisco data center infrastructure?

- Red Hat® OpenShift® is an enterprise grade open source application platform for accelerating the development and delivery of cloud-native applications in a consistent way across the hybrid and multi cloud, all the way to the edge.
- In addition to Kubernetes' core container orchestration, Red Hat OpenShift provides a broad set of capabilities that includes built-in CI/CD pipelines, enhanced security, developer-friendly workflows, and enterprise support enterprise-grade features and tools.
- Red Hat OpenShift simplifies container management, automates the application lifecycle, and incorporates robust security features like role-based access control, network policies, and image scanning.
- Cisco Intersight tightly integrates UCS-X with Red Hat OpenShift and Ansible to provide customers with the most secure systems that have zero touch install and best in industry day 2 operations.

Guide overview

Context

Messaging

Additional information

FAQ

Resources & Contacts

Resources

Red Hat and Cisco joint AI solutions:

Solution brief

 [Download](#)

Bare-Metal

[FlexPod for Accelerated RAG Pipeline with NVIDIA NIM and Cisco Webex](#)

[AI Inferencing on Cisco UCS X-Series with Intel AMX for Medical Image Analysis](#)

Virtualized on VMware vSphere

[FlexPod Datacenter with Generative AI Inferencing](#)

[Flashstack for Generative AI Inferencing Design Guide](#)

[FlashStack for AI: MLOps using Red Hat OpenShift](#)

[GenAI Inferencing on UCS X-Series with 5th Gen Intel Xeon Scalable Processors on Red Hat OpenShift AI](#)

[FlashStack for Cloud Native with Cisco Intersight, Red Hat OpenShift, and Portworx Enterprise Design](#)

Contacts

Questions? Looking for more information?
Contact us today!

Contact name or title

support@partnerpath.biz

1-888-123-4567

