

Red Hat
Summit

NYRHUG May 2024 Summit 2024 Recap

Patrick Ladd
Senior Technical Account Manager - FSI
NYRHUG Founder / Organizer



Red Hat
User Group Community



Red Hat

Red Hat Summit 2024

- ▶ Product & Technology Overview
- ▶ Customer & Partner News
- ▶ [Red Hat: Summit Connect](#) - events coming to 25 cities in North America red.ht/summitconnect2023



Red Hat
User Group Community



Red Hat

New from Red Hat

Products and innovation unveiled at Red Hat Summit 2024



Red Hat Enterprise Linux AI

(Dev preview) A foundation model platform to seamlessly develop, test and run Granite family large language models (LLMs) to power enterprise applications.



InstructLab

An open source community built around the Large-scale Alignment for chatBots (LAB) method and open source Granite models from IBM; the community aims to put LLM development into the hands of developers by making building and contributing to an LLM as simple as contributing to any other open source project.



Red Hat OpenShift AI

Red Hat's flexible, scalable MLOps platform adds new capabilities with support for AI at the edge, hardware accelerator profiles, distributed workload support with Codeflare and being able to use multiple model servers to support both predictive and GenAI



Podman AI Lab

Helps fuel developer adoption of GenAI for building intelligent applications or enhancing their workflow using AI-augmented development capabilities. AI Lab features a recipe catalog with sample applications that give developers a jump start on some of the more common use cases for LLMs.



Red Hat Lightspeed Expansion

(Coming soon) Extends intelligent, natural language processing capabilities to Red Hat Enterprise Linux and Red Hat OpenShift, enabling organizations to minimize skills gaps and maximize existing knowledge through GenAI.

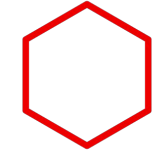
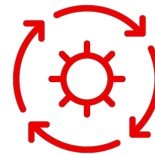


Image mode for RHEL

(Technology preview) A new deployment method for RHEL, image mode takes a container-native approach to deliver the OS as a bootc container image, creating a deployment model that fits seamlessly into container-native workflows.



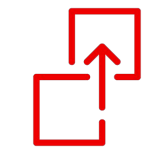
Policy as Code

(Coming soon) Enables Ansible Automation Platform users to better enforce governance, risk and compliance (GRC) policies across complex, AI-infused IT environments during application development and during automation runs.



Red Hat Connectivity Link

(Dev preview) A hybrid multicloud application connectivity solution based on the Kuadrant project, it provides a cohesive platform for enhanced application performance, global scalability and robust security capabilities through simplified deployment and management processes.



Konveyor GenAI

(Coming soon) Brings GenAI capabilities to the Konveyor community, helping to integrate application modernization workflows with LLMs to provide better, faster and more efficient application transformation efforts.



Red Hat Summit '24 product & technology news

- ▶ New features for **Red Hat OpenShift AI** (Tech Preview)
- ▶ **Red Hat RHEL AI**
- ▶ **InstructLab** (Generally Available)
- ▶ **Red Hat Lightspeed** (Ansible - Updated, OpenShift - Preview, RHEL - Vision)
- ▶ **Podman AI Lab** (Generally Available)
- ▶ **RHEL Image mode** (Generally Available)
- ▶ **Ansible Automated Policy as Code** (Vision Announcement)
- ▶ Announcing the introduction of **Gen AI for Konveyor**
- ▶ **Red Hat Connectivity Link**, (Developer Preview)
- ▶ **Extended Update Support (EUS) offerings for OpenShift (GA)**
- ▶ **OpenShift Dev Spaces 3.13.0** release (GA)



Red Hat Summit '24 product & technology news

- ▶ Announcing more **Red Hat OpenShift AI** model serving options, distributed workloads, and model serving at the edge (Tech Preview)
- ▶ Announcing **Red Hat Enterprise Linux AI** to help seamlessly develop, test, and run GenAI models for enterprise applications
- ▶ Introducing **InstructLab** to simplify GenAI model alignment tuning and community based approach to model development (Generally Available)
- ▶ Announcing **Red Hat Lightspeed** - infusing enterprise-ready AI across our portfolio. (Ansible - Updated, OpenShift - Preview, RHEL - Vision)
- ▶ Announcing **Podman AI Lab** enabling developers to easily bring AI to their applications on their laptop (Generally Available)
- ▶ Introducing **RHEL Image mode**, a new deployment method for Red Hat Enterprise Linux that delivers the platform as a container image (Generally Available)
- ▶ Introducing **Ansible Automated Policy as Code** helps you operate with less risk and more consistency to ensure AI can do its best work. (Vision Announcement)
- ▶ Announcing the introduction of **Gen AI for Konveyor**, the Red Hat-led open source community project aimed at helping organizations modernize their applications
- ▶ Introducing **Red Hat Connectivity Link**, a solution for hybrid and multi cloud environments that provides a cohesive set of application connectivity capabilities (Developer Preview)
- ▶ Expanded **Extended Update Support (EUS) offerings for OpenShift** (Generally Available)
- ▶ **OpenShift Dev Spaces 3.13.0** release (Generally Available)



Red Hat
User Group Community



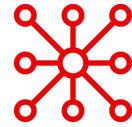
Red Hat



Red Hat Enterprise Linux AI

Foundation Model Platform

Seamlessly develop, test, and run Granite family large language models (LLMs) for enterprise applications.



Granite family models

Open source-licensed LLMs, distributed under the Apache-2.0 license, with complete transparency on training datasets.



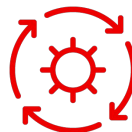
InstructLab model alignment tools

Scalable, cost-effective solution for enhancing LLM capabilities and making AI model development open and accessible to all users.



Optimized bootable model runtime instances

Granite models & InstructLab tooling packaged as a bootable RHEL image, including Pytorch/runtime libraries and hardware optimization (NVIDIA, Intel and AMD).



Enterprise support, lifecycle & indemnification

Trusted enterprise platform, 24x7 production support, extended model lifecycle and model IP indemnification by Red Hat.





Red Hat OpenShift AI

Integrated MLOps platform

Create and deliver GenAI and predictive models at scale across hybrid cloud environments.



Model development

Provides flexibility and composability by supporting multiple AI/ML libraries, frameworks, and runtimes.



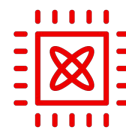
Model serving and monitoring

Deploy models across any OpenShift footprint and centrally monitor their performance.



Lifecycle management

Expands DevOps practices to MLOps to manage the entire AI/ML lifecycle.



Resource optimization and management

Scales to meet the workload demands of foundation models and traditional machine learning.

Available as

- Fully managed cloud service
- Traditional software product on-site or in the cloud!





Red Hat Ansible Lightspeed with IBM watsonx Code Assistant

Accelerate Ansible content creation

- Enable users regardless of background to automate consistently across domains.
- Generate multiple task suggestions using natural language prompts (full playbook generation coming June 2024).

Instill trust in generative AI

- Ansible-specific watsonx.ai foundation model trained on Ansible datasets.
- Provides transparency into origins of recommendations.
- On-premise option coming June 2024.

Enhance developer productivity

- Brings generative AI into existing content creation workflows.
- Simplify code maintenance with repo scans and recommendations.





Accelerating AI-enabled application development

Streamlined onboarding and workflow for developers to learn AI development, experiment with models, and create AI-enabled applications

Available as

- Local workstation tools
- On premises solutions for developer productivity Internal Developer Platform and AI assisted software development

Lab for AI development



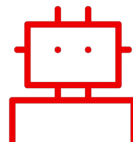
Podman AI Lab, an extension for Podman Desktop, gives developers the ability to build, test and run GenAI applications in containers using an intuitive UI on their local laptop or workstation

Internal Developer Platform



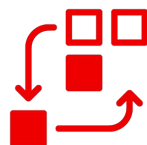
Red Hat Developer Hub enables developers to learn and onboard for AI development with software catalogs and templates, technical documentation and workflow automation across platforms via plugins

AI Assisted Modernization



Infusing GenAI capabilities to the Konveyor project enhancing the developer experience and productivity through recommended source code changes right in their IDE

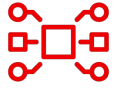
Streamlining Data for AI



Red Hat Application Foundations streamlines and automates data workflows for AI including data acquisition and preparation



Red Hat's AI certified partner ecosystem



Integrated ISVs

Anaconda
Elastic
IBM
Intel
NVIDIA
Pachyderm
Run:ai
Stability AI
Starburst



AI & general ISVs

C3.ai
Cloudera
Couchbase
Crunchy
Enterprise DB
Guise AI
H2O.ai
Hugging Face
Lightbend
Palantir
Redis Labs
SAS



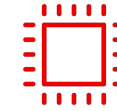
Delivery partners

Accenture
BCG
CapGemini
Deloitte
HCL Technologies
Infosys
Kyndryl
World Wide Technology



Cloud partners

AWS Google
Azure IBM Cloud



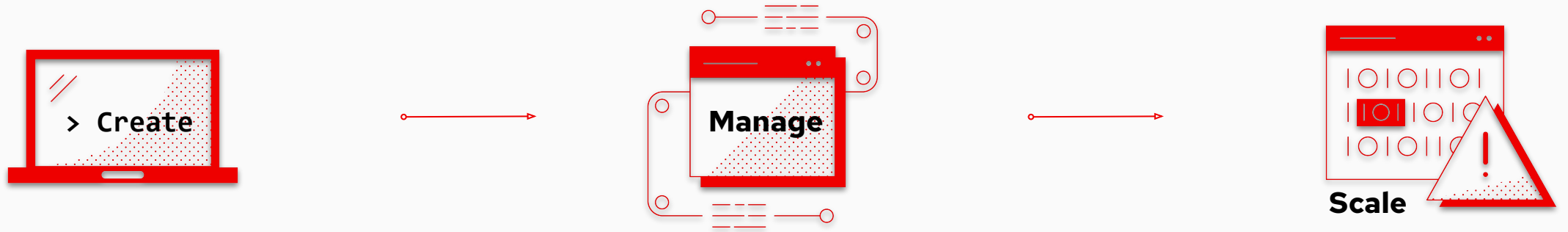
Hardware/OEMs

AMD
Dell
Cisco
HPE
Intel
Lenovo
NVIDIA
Qualcomm
Supermicro



Coming soon: Policy as Code

AAP as a solution for Governance, Risk and Compliance (GRC)



- > **Check policy enforcement during the development cycle**

- > Ansible Lightspeed integration to accelerate development
- > Integration into Ansible plugins for Integrated Development Environments (IDE)s

- > **Provide an integrated mechanism for policy enforcement based on unique needs**

- > Discretionary or mandatory checks before or during automation runs
- > Hook into a wider ecosystem for defining, selecting, finding and using policies for checks and enforcement

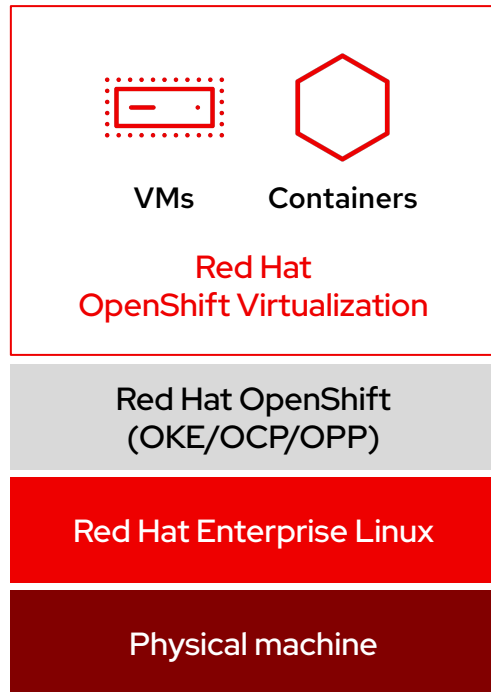
- > **Reporting mechanisms to aid in compliance status and audit tracking.**

- > Output to common output data schemas to allow third party tools easy access to results
- > Customizable alerts and notifications for policy status



Red Hat OpenShift Virtualization

The modern option for general purpose virtualization



- ▶ **Unified platform**
for virtual machines and containers
- ▶ **Consistent management**
tools, interfaces, and APIs incl. ACM and AAP integrations
- ▶ **Performance and stability**
of Linux, KVM, and qemu
- ▶ **Healthy open source community**
the KubeVirt project is a top 10 CNCF active project, with 200+ contributing companies
- ▶ **Diverse ecosystem**
of Red Hat & partner operators

Included feature
of all OpenShift subscriptions

Includes Red Hat Enterprise Linux
guest entitlements

Supports Microsoft Windows
guests through Microsoft SVVP

Inbound guest migration
using Ansible Automation Platform +
Migration Toolkit for Virtualization,
Training and Consulting

Virt admin focused training
[DO316](#), [EX316](#)



Market momentum



Financial services



E-commerce



Telco



Manufacturing and energy



Government and military



Media and technology

Red Hat Summit 2024 Sessions This Week

Dell

Monday

Emirates NBD

Monday, Wed

Ally

Tuesday

Morgan Stanley

Wednesday

Goldman Sachs

Wednesday

SiriusXM

Wednesday

Key Partners:

 NetApp

 NVIDIA

 aws

 KASTEN K10

 Hewlett Packard Enterprise

 TRILI

 DELL Technologies

 portworx
by Pure Storage

 storware
BACKUP AND RECOVERY

 IBM



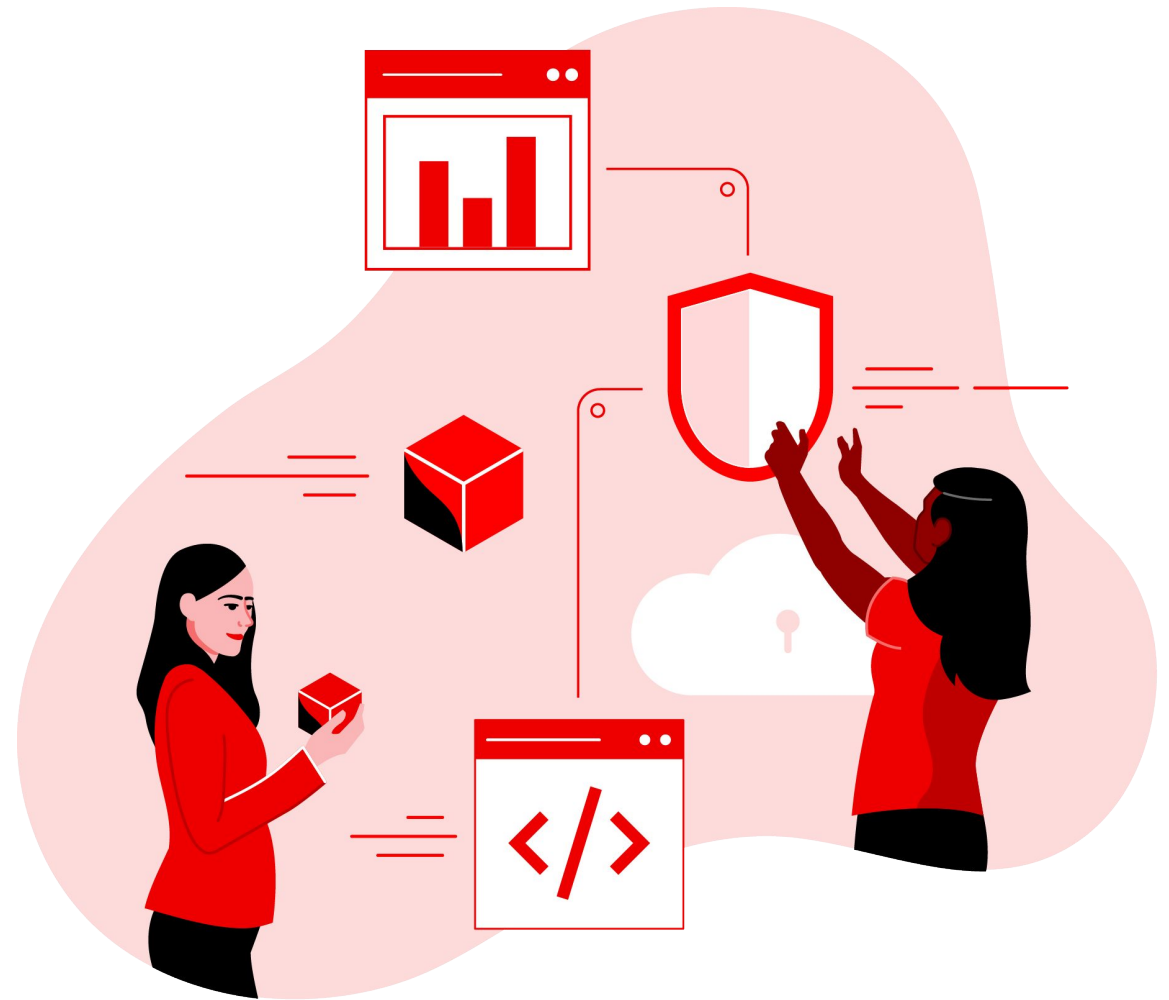
Red Hat
User Group Community



Red Hat

Red Hat Summit

Build a **trusted**
software supply chain
with Red Hat



Software supply chain attacks: a matter of when, not if

Ransom paid but a mere fraction to the overall
downtime and recovery costs of a data breach



742%

average annual increase in
software supply chain
attacks over the past 3 years¹

45%

of organizations worldwide
will experience supply
chain attacks by 2025²

1 in 5

data breaches are due
to a software supply
chain compromise³

71%

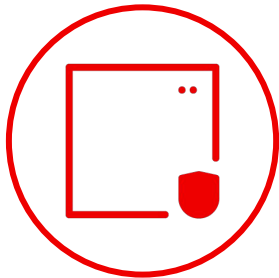
YoY increase in cost
of average ransom
payment⁴



Devs struggle to stay code compliant given high cognitive load¹

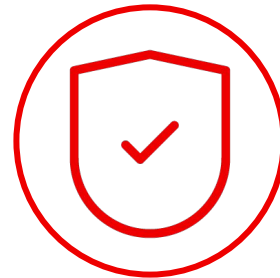
Platform engineer for security and compliance into the SDLC to mitigate and reduce risks

Prevent & identify
malicious **code**



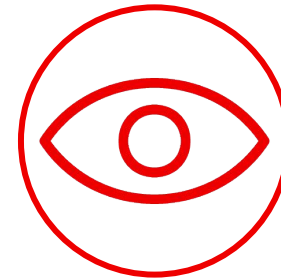
Confidence in
software integrity

Safeguard **build**
systems early



Adherence to
regulation/compliance

Continuously **monitor**
security at runtime



Enhanced trust with
customers/stakeholders

^[1] [Salesforce productivity survey](#)



Evolution of Red Hat Trusted Software Supply Chain

Leveraging the preview for intelligence gathering

Software Supply Chain
Security **is valued**



Validation of RH position: Customer awareness of risks and challenges is high. Preparedness, current solutions are insufficient

SaaS is NOT desired
mode of deployment



On-prem or hybrid deployment is a necessity. A fully-hosted offering was a non-starter for the organizations we engaged.

Comprehensive capabilities are needed



Interviews across Developers, Platform Teams, Ops indicate the importance of Security across the SDLC, breadth of needs and concerns to address

Modularity is a
must-have

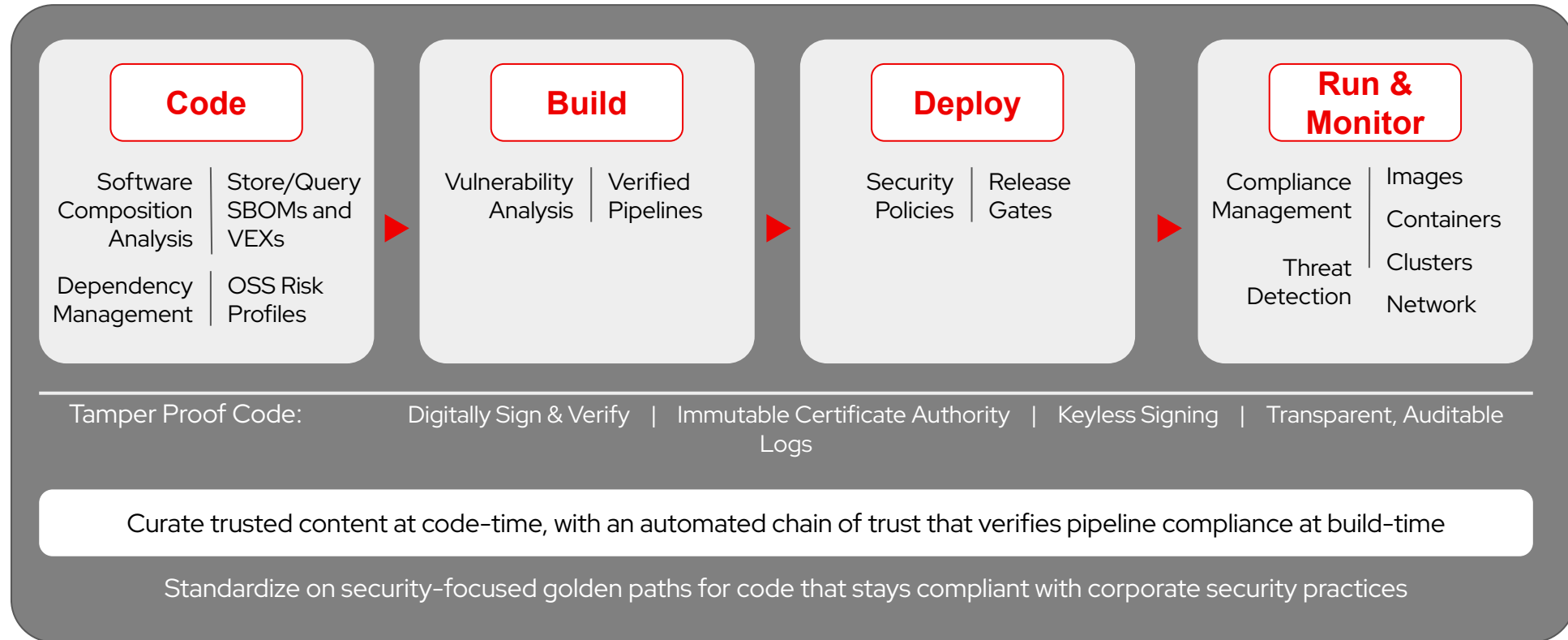


Monolithic, all-in-one solutions not desirable. Teams need flexible, composable solutions. Must interoperate with incumbent approaches and perform consistently across different environments



Accelerate Innovation that Safeguards User Trust

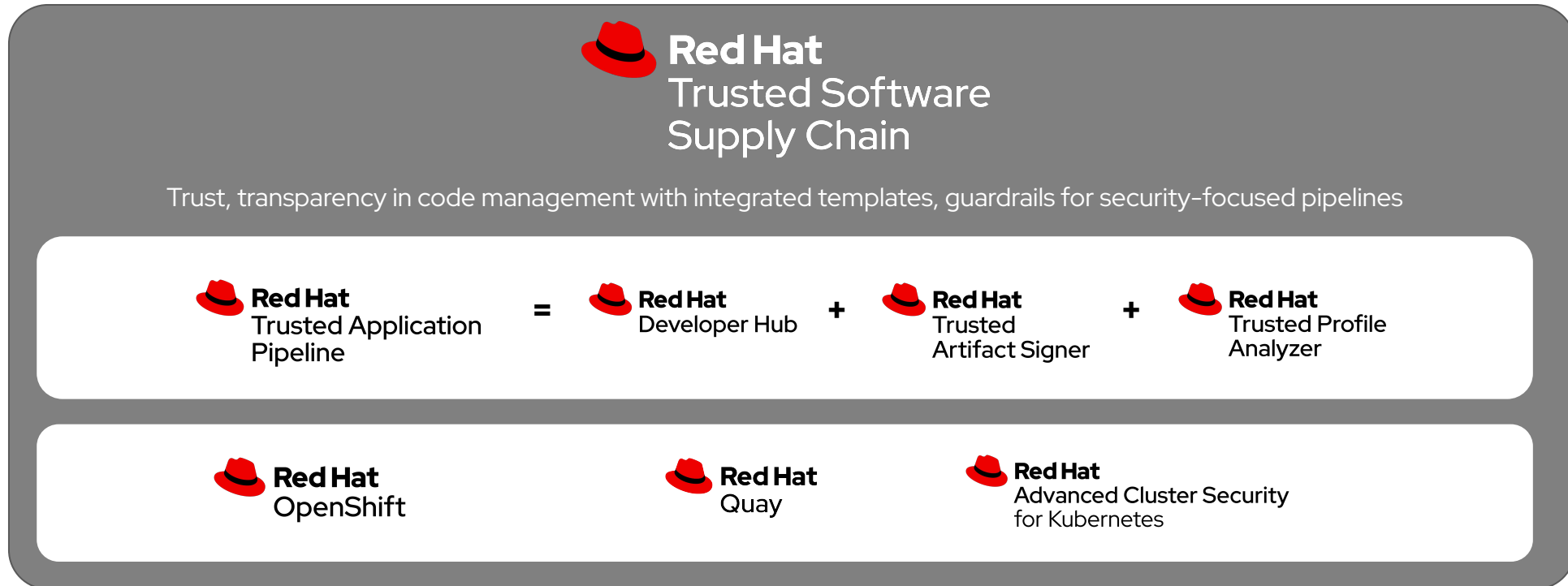
Delivered with integrated security guardrails at every phase of the software development lifecycle



Build and deploy platform, pipeline and applications as-code to an auditable, declarative state that's continuously monitored

Shift Left Security in the Software Supply Chain

Protect the components, processes and practices early in your software factory



*Note: **Red Hat Trusted Application Pipeline** is a single product SKU that includes RHDH, RHTAS, RHTPA.

Red Hat Application Pipeline, Red Hat Trusted Artifact Signer, Red Hat Trusted Profile Analyzer are in service/technical preview



Red Hat® Summit: Connect 2024

Join us as we explore the future of AI, hybrid cloud, open source technology, and IT.

WEST

- ▶ Bellevue, WA - Sept. 19
- ▶ Sacramento, CA - Sept. 26
- ▶ Scottsdale, AZ - Oct. 3
- ▶ Los Angeles, CA - Oct. 16
- ▶ San Jose, CA - Oct. 23

EAST

- ▶ Arlington, VA - Sept. TBD
- ▶ Waltham, MA - Sept. 10
- ▶ New York City, NY - Sept. 12
- ▶ Philadelphia, PA - Oct. 3
- ▶ Detroit, MI - Oct. 9
- ▶ Alpharetta, GA - Nov. 19
- ▶ Jacksonville, FL - TBD

CENTRAL

- ▶ Chicago, IL - Oct. 1
- ▶ Austin, TX - Oct. 3
- ▶ Minneapolis, MN - Oct. 10
- ▶ Dallas, TX - Oct. 15
- ▶ St. Louis, MO - Oct. 30

CANADA

- ▶ Ottawa, ON - Sept. 17
- ▶ Montreal, QC - Oct. 9
- ▶ Vancouver, BC - Oct. 24
- ▶ Toronto - Nov. 5

Learn more:

red.ht/SummitConnect24



Red Hat
User Group Community



Red Hat

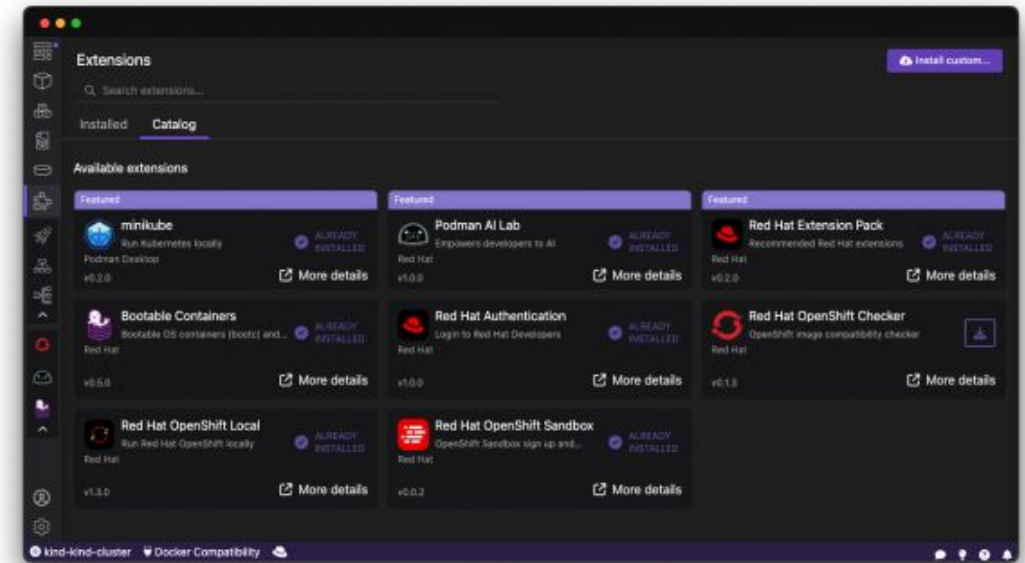
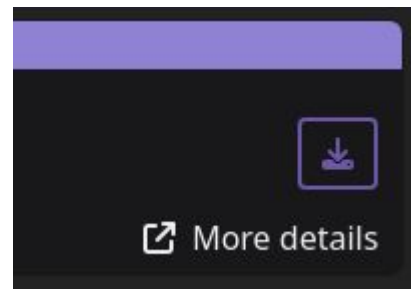
Installation

Install Podman Desktop (1.10 or later)

- ▶ <https://podman-desktop.io/downloads>
 - Available for Linux (binary & flatpack), Mac, and Windows

Find Podman AI Lab extension

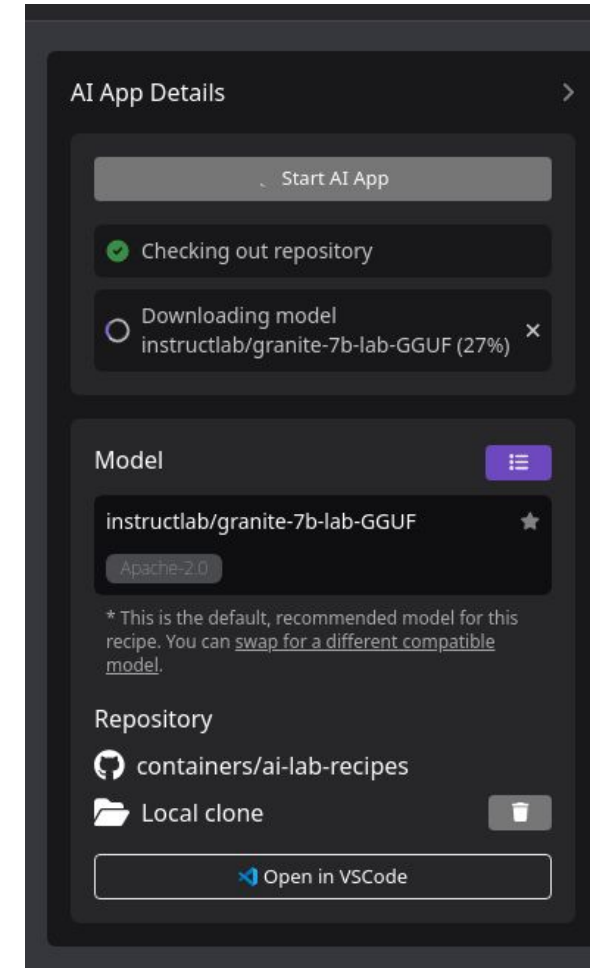
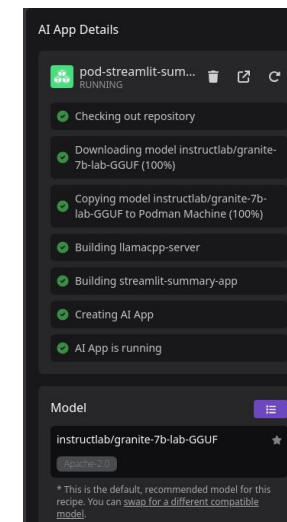
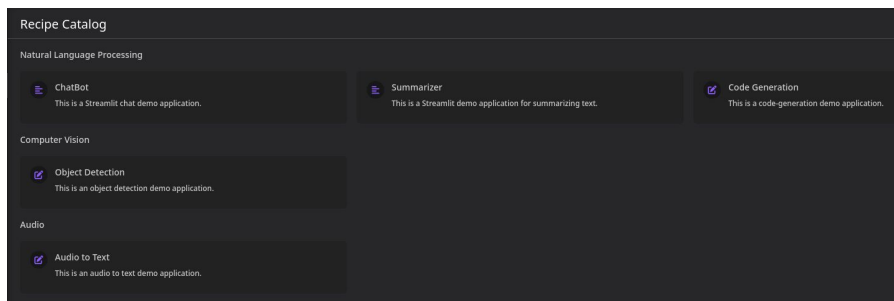
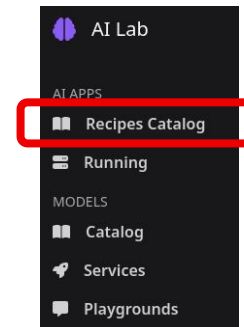
- ▶ Navigate to Extensions -> Catalog
- ▶ Hit the Download button



Start Experimenting!

Build and deploy a model

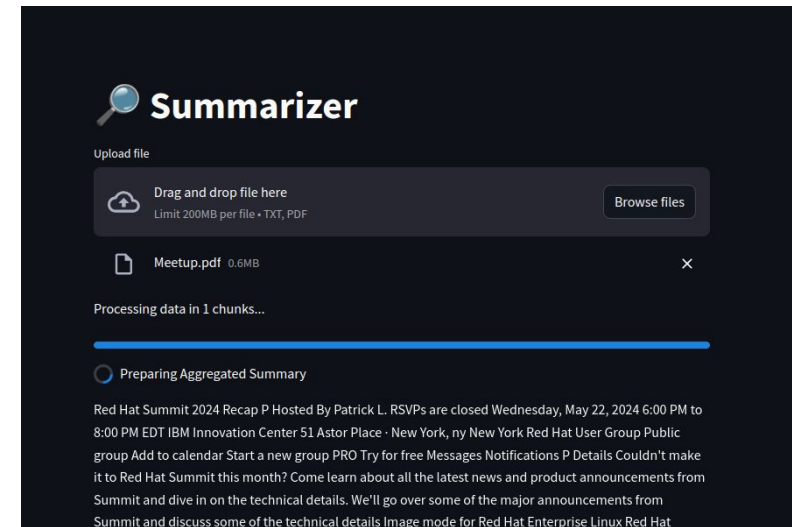
- ▶ Don't know where to start? Try a Recipe from the Recipes Catalog
- ▶ Select a recipe
- ▶ Hit "Start AI App" button on the right
- ▶ Wait for models to download and build
- ▶ Follow the quick start



Experiment some more

Use the model

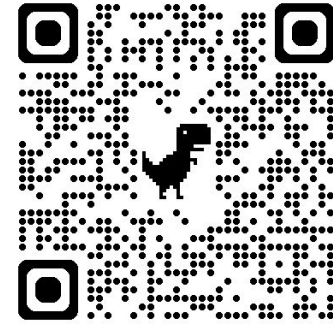
- ▶ Look for
- ▶ You may need to deploy a more powerful podman machine image - 2GB default is **very slow** :)



Podman AI Lab

Learn More

- ▶ <https://developers.redhat.com/articles/2024/05/07/introducing-podman-ai-lab>
- ▶ <https://developers.redhat.com/articles/2024/05/07/podman-ai-lab-getting-started>



RHEL Image Mode

About

Simple. Consistent. Anywhere.

Learn More:

- ▶ <http://red.ht/imagemode>

What is it?

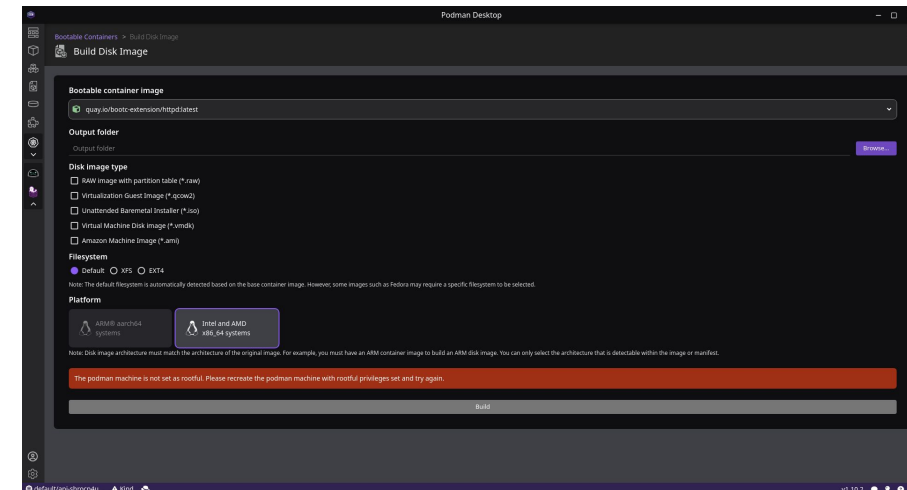
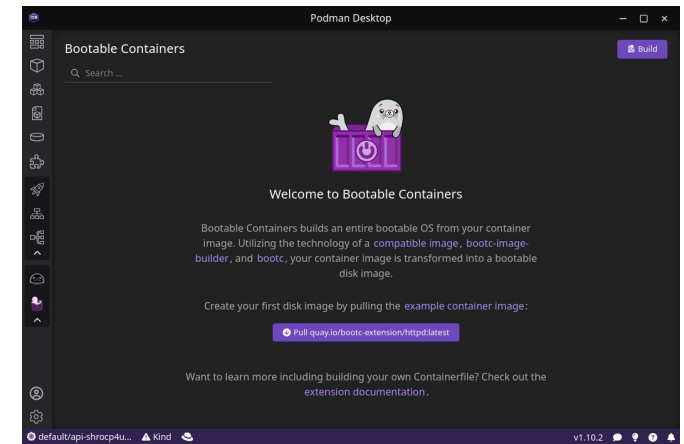
- ▶ New deployment method that takes a container-native approach to deliver the OS as a bootc container image
- ▶ A complete inventory of standard images and environments with history and version control
- ▶ Tracking of OS images and changes across all deployed bootc images
- ▶ Simple updates and rollbacks needing only a reboot
- ▶ A way to experiment faster with new dependencies and capabilities without duplicating work
- ▶ An opportunity to explore containerized Ci/CD and other tools to augment the workflows and gain new functionality



RHEL Image Mode

Using Podman Desktop

1. Add Red Hat extension pack and get Red Hat SSO extension for Podman Desktop
2. Click on Bootable Containers
3. Click “pull quay.io/bootc-extension/httpd:latest”
4. Click “build quay.io/bootc-extension/httpd:latest”
5. Build image



RHEL Image Mode

Using Podman CLI

1. Make sure system is registered to Red Hat subscription manager and updated

2. `podman pull registry.redhat.io/rhel9/rhel-bootc:9.4`

3. Create dockerfile:

```
#install the lamp components

RUN dnf module enable -y php:8.2 nginx:1.22 && dnf install -y httpd mariadb mariadb-server php-fpm
php-mysqlnd && dnf clean all

#start the services automatically on boot

RUN systemctl enable httpd mariadb php-fpm

#create an awe inspiring home page!

RUN echo '<h1 style="text-align:center;">Welcome to RHEL image mode</h1> <?php phpinfo(); ?>' >>
/var/www/html/index.php
```

4. Build container:

```
$ podman build -f containerfile -t
quay.io/[my_account]/lamp-bootc:latest
```



RHEL Image Mode

Using RHEL Kickstart

1. **Download Installation Media:** Locate and download the 9.4 Boot ISO compatible with your system architecture from the [designated website](#).
2. **Create Kickstart Configuration File:** Develop your Kickstart configuration. Below is a sample file that you can modify to suit your needs:

```
# network configuration
network --bootproto=dhcp --device=link --activate

# Partitioning Setup
clearpart --all --initlabel --disklabel=gpt
reqpart --add-boot
part / --grow --fstype xfs

# --- Container Image Installation ---
ostreecontainer --url quay.io/[your_account]/chatbot-bootc:1.0 #
# --- Basic Security ---
firewall --disabled # Disable the firewall (consider hardening this later)
services --enabled=sshd # Enable SSH service for remote access
# --- Optional User Setup ---
user --name=cloud-user --groups=wheel --plaintext --password=changemme
sshkey --username cloud-user "ssh-ed25519 AAAAC3Nza...."
# --- Root User SSH Key (If Desired)---
rootpw --iscrypted locked #Disable direct root password login
# Add SSH key for root (replace with your key)
sshkey --username root "ssh-ed25519 AAAAC3Nza...."
# --- System Reboot ---
reboot
```

3. **Check and manage with:**

```
systemctl status bootc-fetch-apply-updates.timer
```

```
bootc upgrade
```

<https://developers.redhat.com/learn/rhel/rhel-image-mode-kickstart>



Ansible Policy as Code

Create - Manage - Scale

- **Aims:**
 - **Enforce policies during the creation of automation code**
 - **Integrate policy enforcement before or during automation runs**
 - **Simplify reporting and easily identify out-of-policy areas**
- **Get Involved:**
 - **Working Group:**
<https://forum.ansible.com/tag/policy-as-code>
 - **Webinar June 18th:**
<https://www.redhat.com/en/events/webinar/automating-policy-as-code-for-consistency-and-compliance>



Lightspeed Generative AI

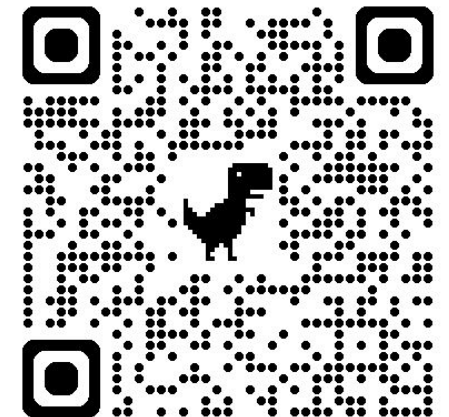
Extending Product Coverage of Lightspeed

- **Ansible Lightspeed already released and available**
- **New technologies being added using GenAI and IBM Watson AI**
- **OpenShift Lightspeed due late 2024**

- **Sign up here:**

<https://app.smartsheet.com/b/form/13e63c33ec1842ab850e75a87a5c4d65>

- **RHEL Lightspeed in the works as well**



Red Hat
Summit

Thank you



[linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)



[facebook.com/redhatinc](https://www.facebook.com/redhatinc)



[youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)



twitter.com/RedHat

Continue to build out your community and learnings with a
[Red Hat User Group \(RHUG\)](#) in your area.



Red Hat
User Group Community



Red Hat