Red Hat Summit

Connect

Sessione #1 | Red Hat

Roma, 12 novembre

Agenda

14:00 - 14:30	**Red Hat Enterprise Linux 10, reinvent your Linux**	**Alessandro Rossi**, EMEA Associate Principal SSA - RHEL, Red Hat
14:30 - 15:00	**Percorso di modernizzazione verso una virtualizzazione cloud native basata su Openshift Virtualization**	**Paolo Conti**, Senior IT Architect, Poste Italiane **Stefano Tommaso Lombardi,** Cloud Architect, Poste Italiane **Matteo Santucci, Architect**, Red Hat **Valentino Uberti**, Senior Specialist Solution Architect, Red Hat
15:00 - 15:30	**Garantire la Conformità della PCI DSS 4.0 in Ambienti OpenShift Multi-Cluster**	**Riccardo Tagliaferri**, Cloud Architect, Poste Italiane **Francesco Cuccagna**, Senior Consultant, Red Hat **Amedeo Salvati**, Principal Consultant, Red Hat
15:30 - 16:30	**Tutto ciò che devi sapere su AI Agents, MCP servers e RAG per realizzare un'applicazione moderna che integra l'AI in modo concreto**	**Marco Caimi**, Senior Specialist Solution Architect, Red Hat **Matteo Combi**, Senior Specialist Solution Architect, Red Hat **Francesco Rossi**, Senior Specialist Solution Architect, Red Hat **Natale Vinto**, Technical Director, Red Hat
16:30 - 17:00	**AI & Automazione: la nuova era del controllo IT con Red Hat AAP 2.6**	**Andrea Guarnaccia**, Senior Specialist Solution Architect, Red Hat **Marcello Panci**, Ansible Sales Specialist, Red Hat **Giorgio Voltolina**, Ansible Sales Specialist, Red Hat

Red Hat Summit

Connect

Red Hat Enterprise Linux 10

Reinvent your Linux

Alessandro Rossi

Associate Principal Specialist Solution Architect - RHEL
Red Hat

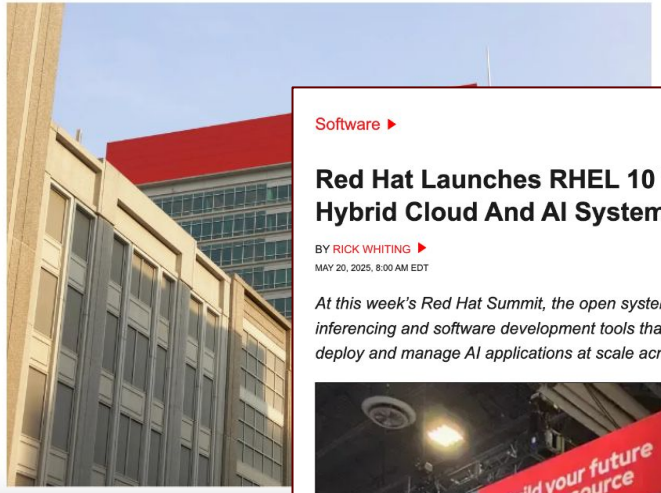
Home / Business / Enterprise Software

How Red Hat just quietly, radically transformed enterprise server Linux

RHEL 10 becomes the first major enterprise Linux distro to discard traditional packaging and embrace immutable. See how we got here.

Written by **Steven Vaughan-Nichols,** Senior Contributing Editor

June 2, 2025 at 11:02 a.m. PT



/ related

Software ▸

Red Hat Launches RHEL 10 With New Capabilities For Hybrid Cloud And AI Systems

BY RICK WHITING ▸

MAY 20, 2025, 8:00 AM EDT

At this week's Red Hat Summit, the open systems software giant also debuted new AI inferencing and software development tools that the company said will make it easier to bu deploy and manage AI applications at scale across hybrid ecosystems.

theCUBE Coverage from SiliconANGLE's livestreaming video studio

UPDATED 11:05 EDT / MAY 23 2025

INFRA

From AI to quantum-ready security, RHEL 10 brings the future to enterprise IT

BY BRIAN NJUGUNA

Red Hat Enterprise Linux continues to distinguish itself as a trusted business operating system, delivering enterprise-ready infrastructure tailored for mission-critical workloads. Now, RHEL 10 is emerging as the next-generation foundation for innovation and reliability in modern IT environments.

RHEL 10 introduces comprehensive enhancements in security, artificial intelligence integration,

InfoWorld

Red Hat Enterprise Linux 10 adds AI-powered management

News

May 21, 2025 • 2 mins

Cloud Management | Configuration Management | Generative AI

Lightspeed is an AI-powered service that allows users to build, deploy, and manage Red Hat's Linux using simplified commands.

Credit: Michael Vi/Shutterstock

TECHZINE EU

3 MIN • INFRASTRUCTURE

Red Hat launches RHEL 10 with intelligence and security for hybrid cloud

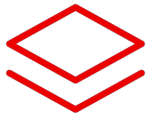
BERRY ZWETS

20 May 2025, 14:45 CEST

Red Hat has released version 10 of Red Hat Enterprise Linux (RHEL). This new platform offers companies an intelligent and secure foundation for hybrid cloud environments and AI workloads. RHEL 10 introduces AI-driven management technology, post-quantum cryptography, and a container-native approach to operating systems and applications.

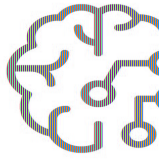
One of the most notable innovations in RHEL 10 is the integration of AI into the management of Linux environments. With the new Red Hat Enterprise Linux Lightspeed functionality, administrators can access generative AI built directly into the operating system. This system provides context-aware guidance and recommendations through a

New Marquee Features



1. Image Mode

Image Mode moves from “Tech Preview” to a GA, fully supported feature.

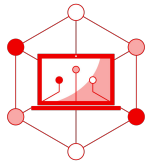


2. Command Line Assistant

AI augmented assistant (RAG model with Watson X, no additional licensing needed)

3. Post Quantum Encryption Capable

Technologies and framework for applying post quantum requirements when they become available



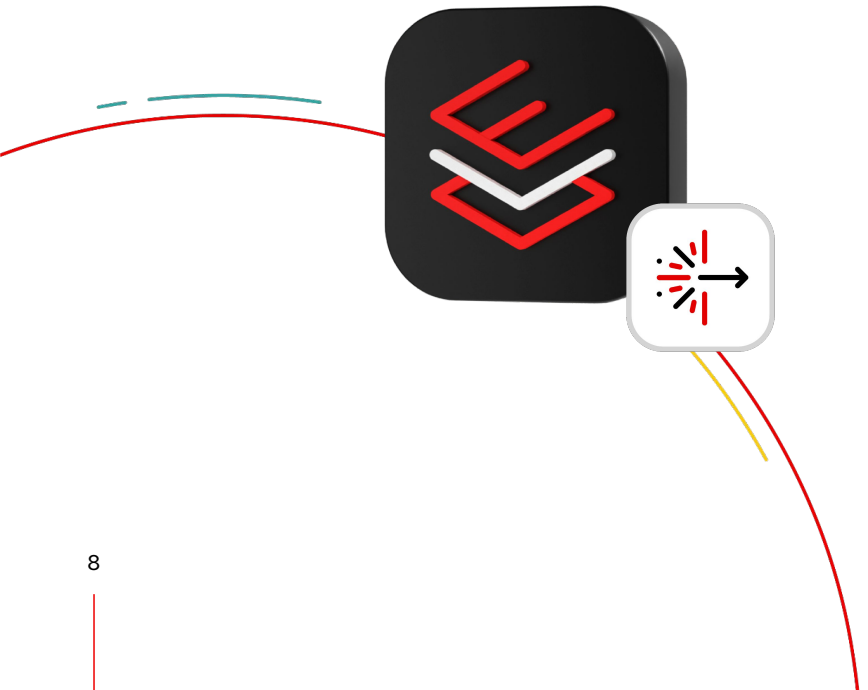
4. Insights Advisor and Vulnerability in Satellite

Proactively analyze and remediate availability, performance, and security risks in disconnected environments

RHEL Lightspeed

With decades of Red Hat's Linux knowledge and expertise

Red Hat Enterprise Linux Lightspeed vision and goals



Artificial Intelligence

Harnesses the power of AI to help increase productivity

Unlock Red Hat’s expertise

Provide Red Hat's decades of Linux experience to help your workloads succeed

Level up skills

Makes Red Hat Enterprise Linux easier to use, secure, tune, and troubleshoot for both new and experienced users

Proactive guidance

Proactively provide relevant information and guidance

RHEL Lightspeed features

Image builder package recommendations

Command line assistant

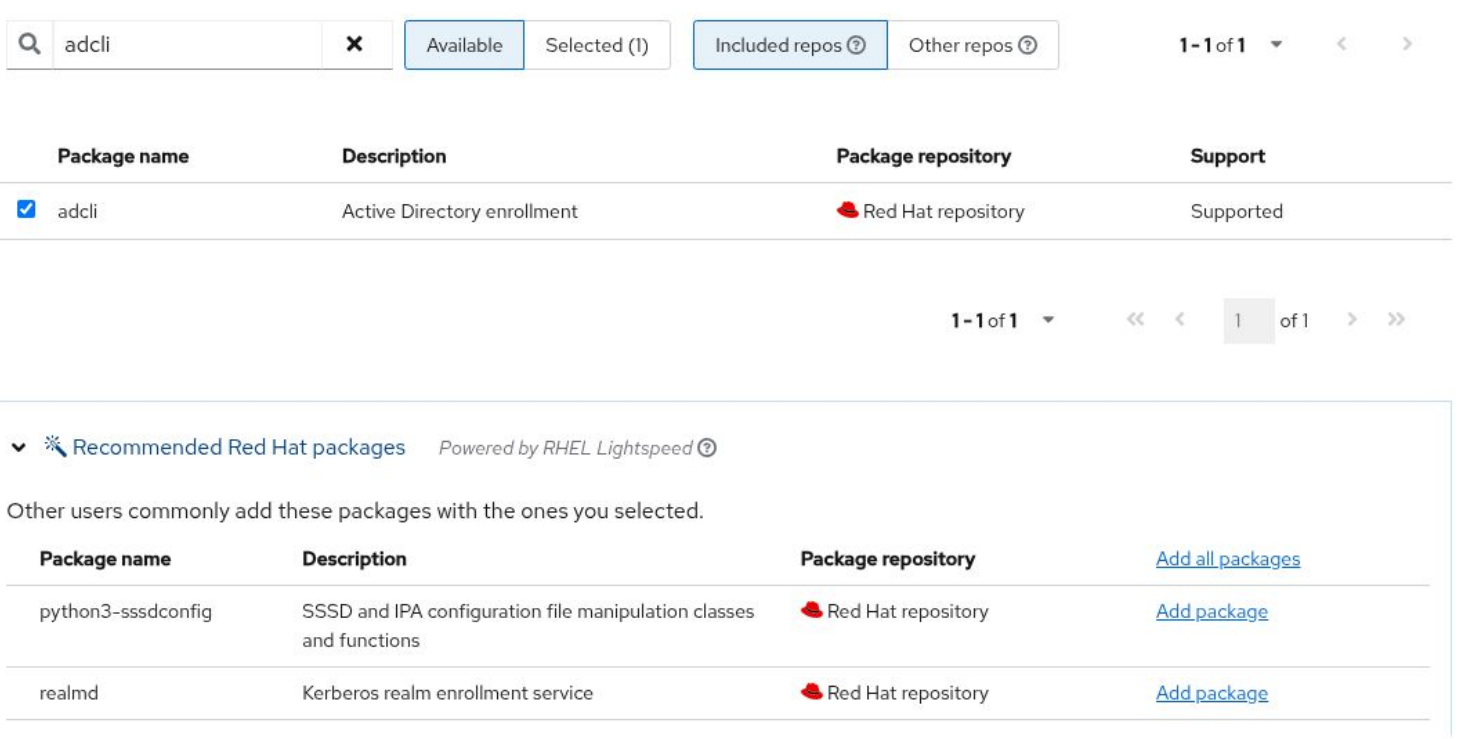
RHEL Lightspeed features

Image builder package recommendations

Command line assistant

Recommendations

Image Builder analyzes selected packages and recommends related, relevant packages to also be included



RHEL Lightspeed features

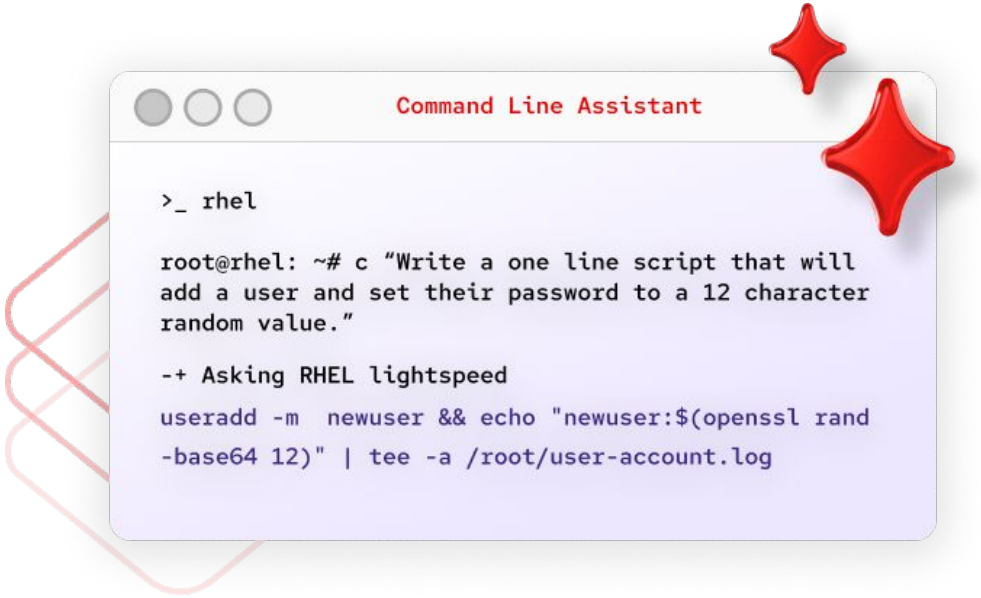
Image builder package recommendations

Command line assistant

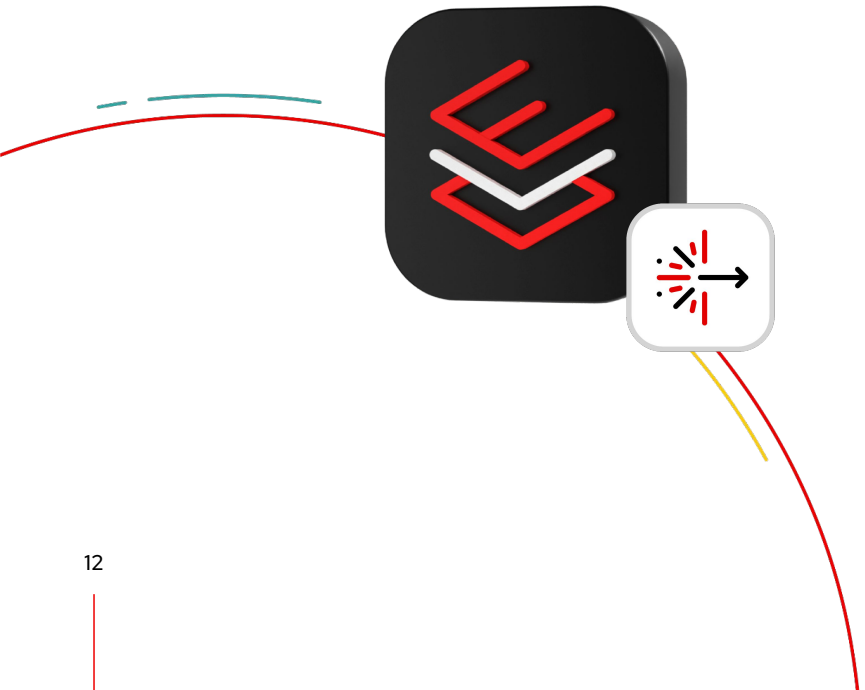
The command line assistant is an optional generative AI assistant available directly in the RHEL command line interface. It can help you more easily manage, troubleshoot, and work with your systems.

With the command line assistant, you can:

- Ask and receive answers to RHEL-related questions in plain language.
- Gain access to information from the RHEL documentation and Red Hat knowledgebase
- Get assistance with troubleshooting configuration issues, deciphering log entries, and more.



Resources available to RHEL Lightspeed that help it with answering questions



RHEL documentation

RHEL 9 documentation and release notes

Red Hat knowledgebase

RHEL 9 verified knowledge base solutions

CVE info

RHEL 9 related CVE's

Errata info

RHEL 9 related errata

Developer Preview

Can command line assistant work without connection?

YES!

[Link to Red Hat Blog post]

Image Mode for RHEL

Manage the operating system with container tools and technologies

Standardizing and innovating with containers

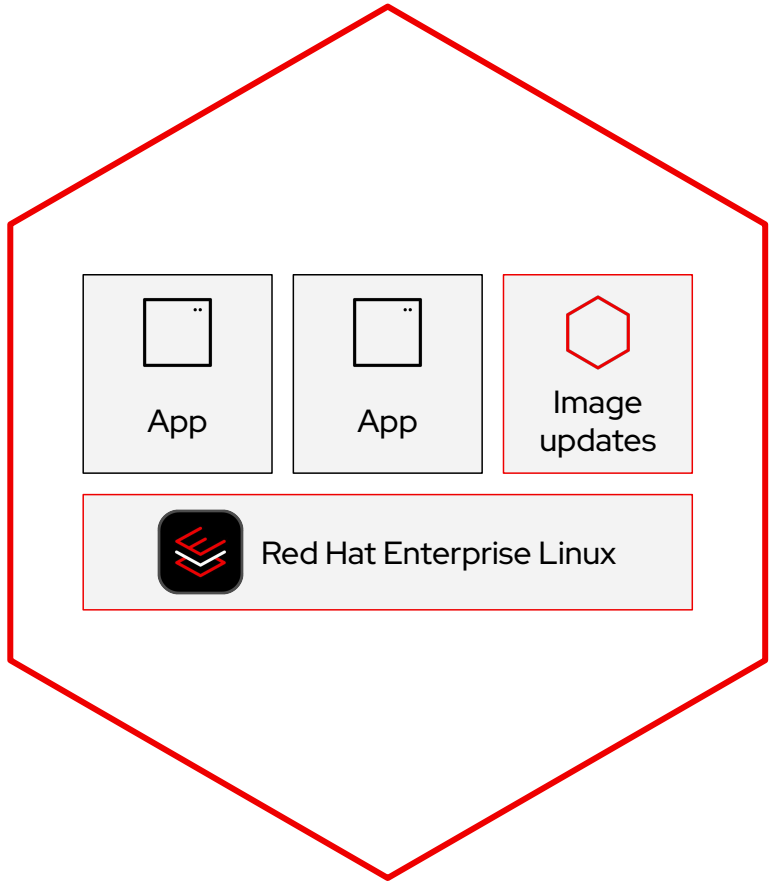


Image mode for Red Hat Enterprise Linux is a simple, consistent approach to build, deploy and manage the operating system using container technologies.

Now you can manage the operating system with the same tools and workflows as applications, promoting a common experience and language across teams.

Outcomes

What does image mode fix today?

Less risk

Reduce the risk associated with updates with atomic transactions and rollbacks

Better builds

Improve the composability and repeatability of standard builds through layering

Move faster

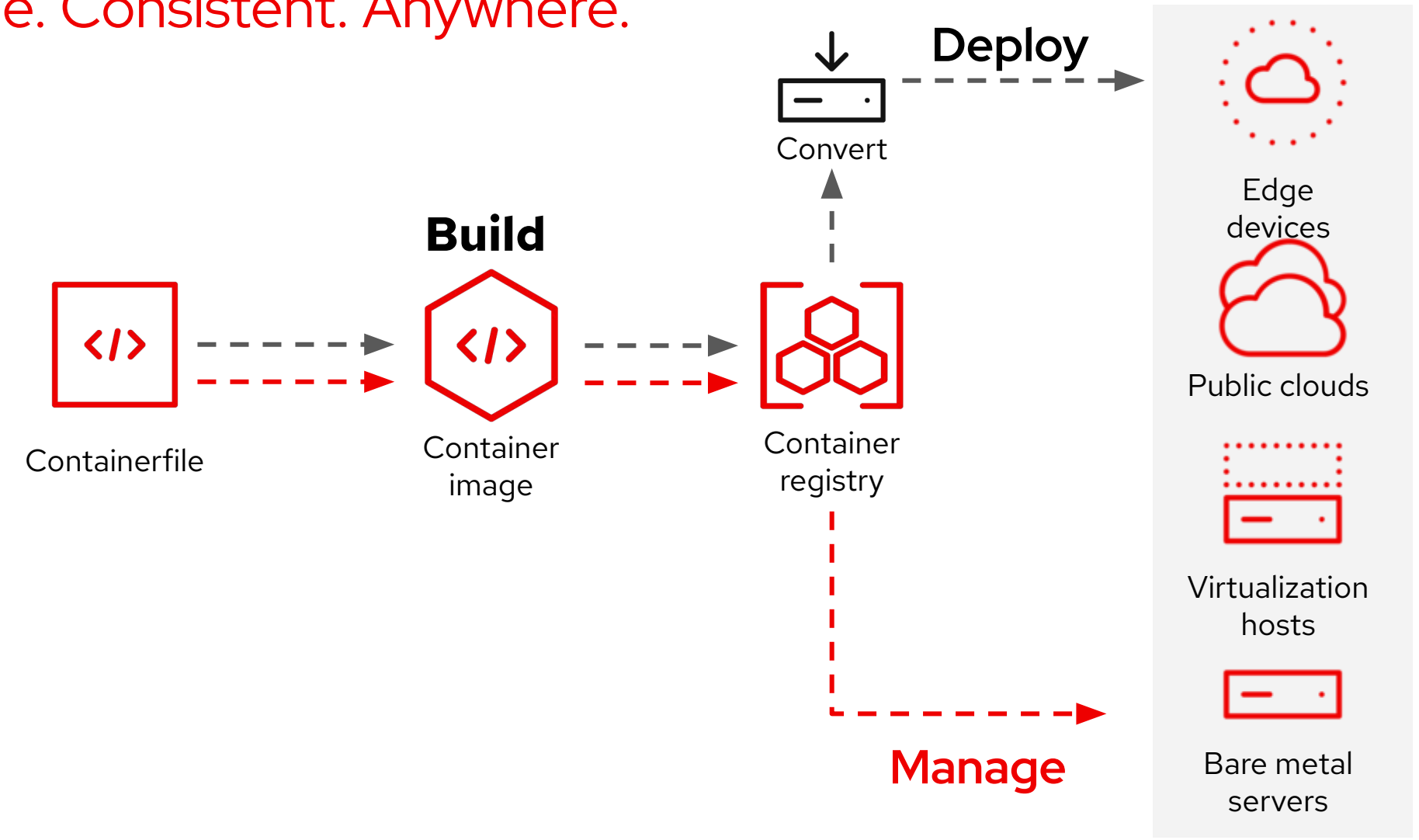
Increase the speed of experimentation

Streamline process

Simplify end to end management with a single process for OS and applications

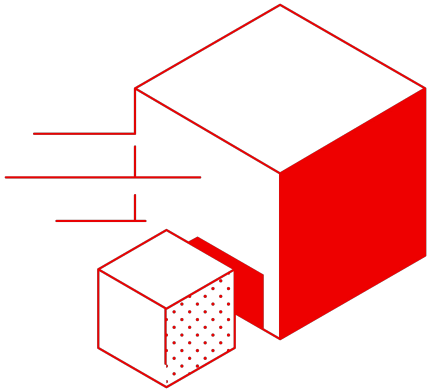
Image mode for Red Hat Enterprise Linux

Simple. Consistent. Anywhere.



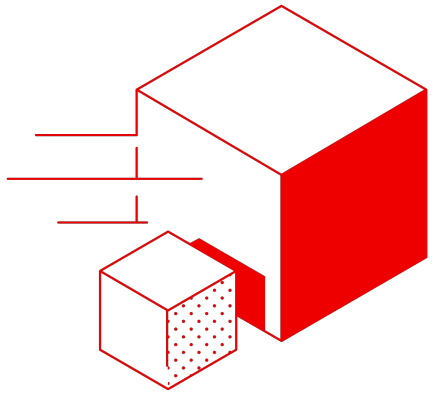
Bootable containers building blocks

- Dedicated bootc base image
- bootc CLI
- bootc-image-builder



Bootable containers building blocks

- **Dedicated bootc base image**
- bootc CLI
- bootc-image-builder



A base image to rule them all!

Image Specs:

- 439 rpms
- ~785M compressed
- ~2.2G on disk

Primary contents:

- systemd, kernel, bootc
- rpm-ostree
- linux-firmware
- NetworkManager
- podman
- python
- Misc CLI tools: jq, sos

➢ Available on Quay.io:
  - registry.redhat.io/rhel10/rhel-bootc

Bootable containers building blocks

- Dedicated bootc base image
- **bootc CLI**
- bootc-image-builder

Manage the **bootc system** lifecycle

`bootc upgrade`
Download and stage an updated container image based on image reference.

`bootc rollback`
Rollback to the previous state.

`bootc switch`
Change to a different reference image

`bootc install`
Install container image **`to-disk`** or **`to-filesystem`**

- Upstream docs:
  - https://bootc-dev.github.io/

Bootable containers building blocks

- Dedicated bootc base image
- bootc CLI
- **bootc-image-builder**

Convert **bootc images** to consumable formats

`qcow2`	QEMU Disk Image
`ami`	Amazon Machine Images
`vmdk`	Virtual Machine Disk Image
`vhd`	Azure / Hyper-V disk image
`gce`	Google Compute Engine image
`raw`	MBR/GPT Raw disk image
`iso`	Bare Metal installer

- Available as a container image:
  - registry.redhat.io/rhel10/bootc-image-builder
- Bootable Container Extension for **Podman Desktop**
  - Build for x86 & multi-arch builds

Many ways to build

- Podman CLI
- Podman Desktop
- CI/CD and Automation Tools

podman

podman desktop

GitLab

GitHub

argo

Jenkins

TEKTON

ANSIBLE

Many ways to build

- **Podman CLI**
- Podman Desktop
- CI/CD and Automation Tools

Different content, same tools.

- Build a bootc image like any other container image
- Integrate the image build phase in existing workflows and automation
- No need to reinvent the wheel, hassle-free implementation

Many ways to build

- Podman CLI
- **Podman Desktop**
- CI/CD and Automation Tools

An extension to rule them all!

- Bootc Podman desktop extension available
- Build the Containerfile image directly from Podman Desktop
- Convert the image using the bootc-image-builder integration
- Additional extensions for: Kind, Kubernetes Contexts, MinC

Many ways to build

- Podman CLI
- Podman Desktop
- **CI/CD and Automation Tools**

It works with your favourite tool!

- Example templates available!
  - https://gitlab.com/redhat/cop/rhel/rhel-image-mode-cicd

Use Cases

Where does image mode fit today?

Use Cases Blog

AI/ML Stacks

Perfectly version app dependencies from kernel, GPU & accelerator drivers, frameworks, runtimes, etc

1:1 App/Host

Manage the OS AND app as a single unit

Edge appliances

Easily manage a fleet of systems with registries and auto-updates

Standalone container hosts

Use common toolchains and pipelines to build containerized applications and the hosting OS

Try it yourself!

The whole demo shown in this session and other use cases are available in the following Github repository:

https://red.ht/rhel-image-mode-demo

Everybody is welcome to use it, fork and suggest improvements.

RHEL Tech Journey

Next sessions:

- **November 20th** - Image Mode and Lightspeed workshop
- **December 11th** - Red Hat Satellite workshop

red.ht/rhel-tech-journey

Red Hat Summit

Connect

Thank you

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat Summit

Connect

GEN 3 VIRT

Percorso di modernizzazione verso una virtualizzazione cloud native basata su OpenShift Virtualization

Posteitaliane | Red Hat

Stefano Lombardi

Cloud Architect
Poste Italiane

Matteo Santucci

Cloud Architect
Red Hat

Paolo Conti

System Engineer
Poste Italiane

Valentino Uberti

Senior Specialist Solution Architect
Red Hat

Posteitaliane | Red Hat

MS

Poste Italiane: Profilo della più grande infrastruttura di servizi in Italia

Chi Siamo

Con oltre **160 anni di storia**, Poste Italiane è un motore di innovazione e un punto di riferimento per famiglie, imprese e Pubblica Amministrazione.

- **Riconoscibilità e Capillarità:** Una realtà unica nel Paese per dimensioni e presenza sul territorio.

La Nostra Rete in Numeri

- 👥 **121.000** Dipendenti
- 🏢 **12.755** Uffici Postali
- 🤝 **46 Milioni** di Clienti

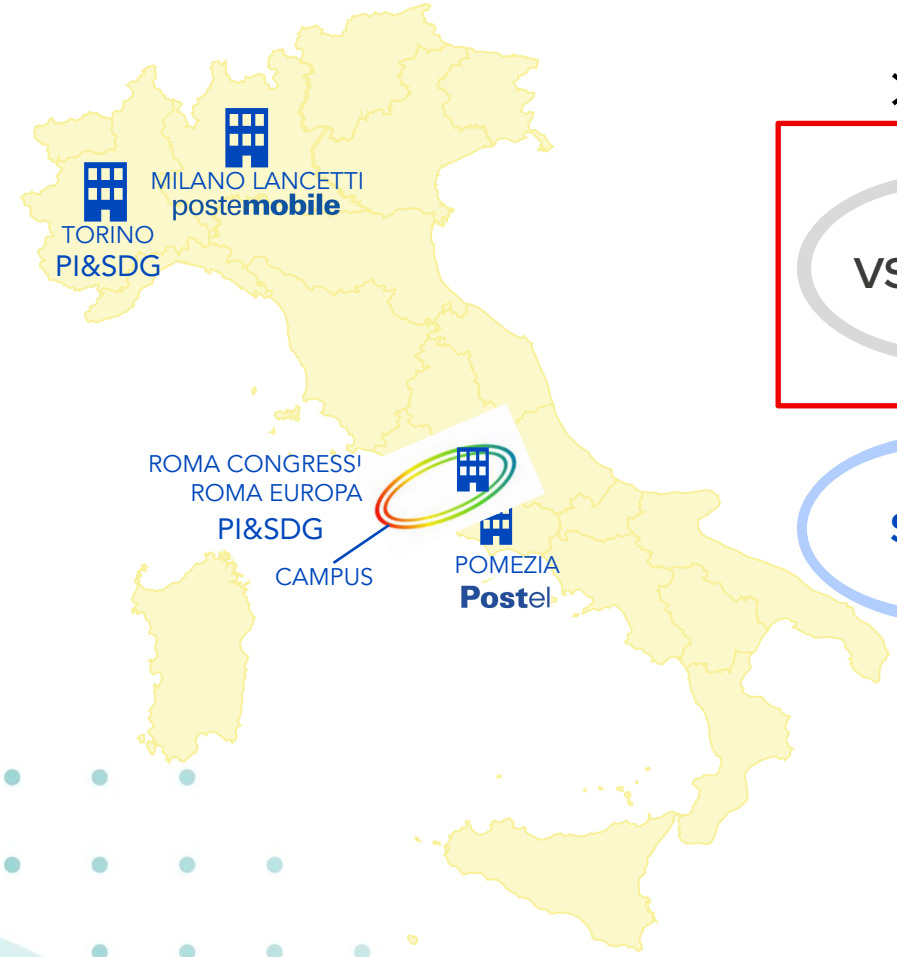
Le Nostre Aree di Business

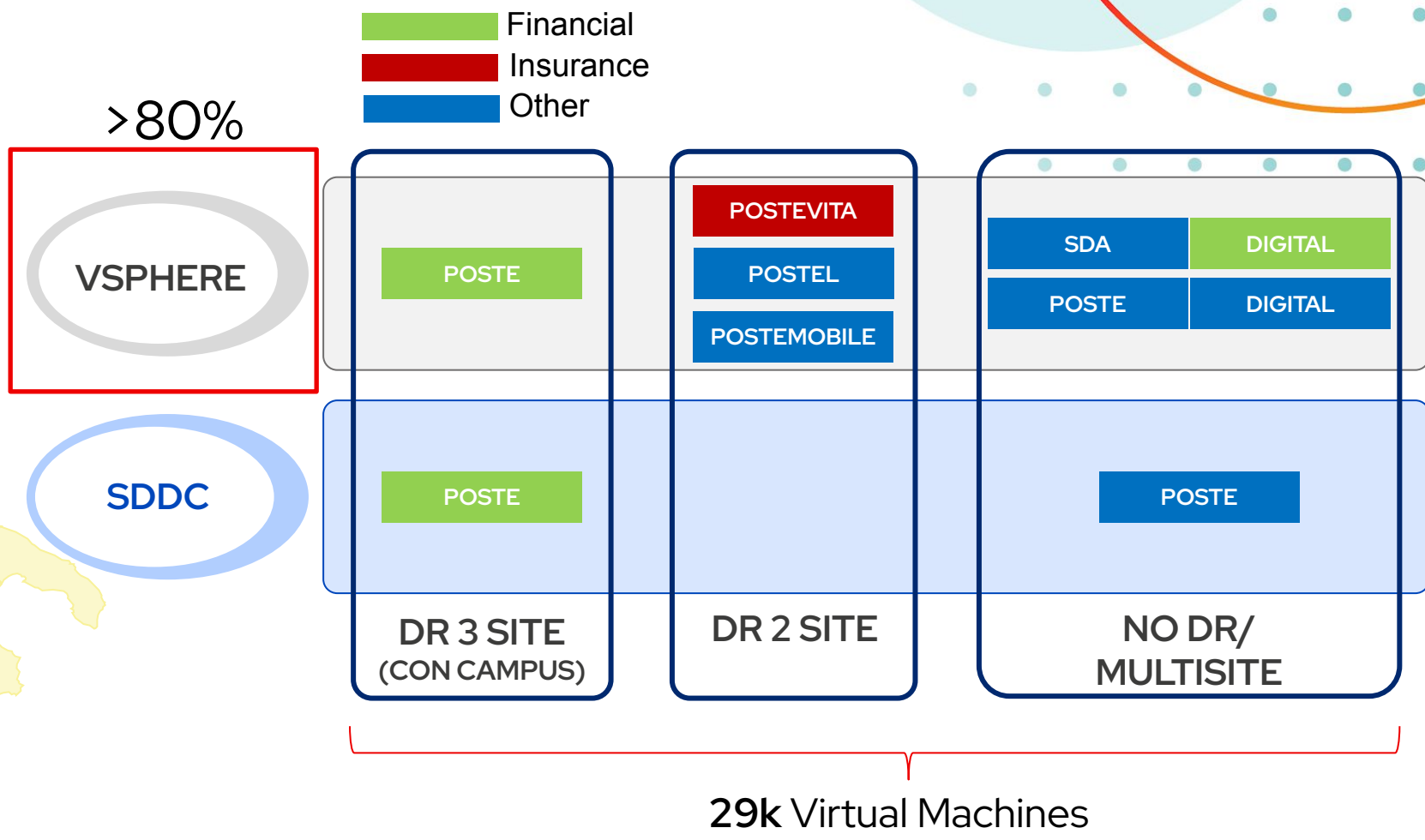
- 📦 **Corrispondenza, Pacchi e Logistica**
  - *Servizi postali tradizionali e soluzioni avanzate per l'e-commerce.*
- 💰 **Servizi Finanziari e Assicurativi**
  - *Ampia gamma di prodotti per risparmio, investimenti e protezione.*
- 📱 **Sistemi di Pagamento e Telefonia**
  - *Soluzioni digitali innovative e servizi di telefonia mobile.*
- 💡 **Energia**
  - *Offerte di energia elettrica e gas per le famiglie italiane.*

Posteitaliane | Red Hat

SL

Datacenter di Poste Italiane

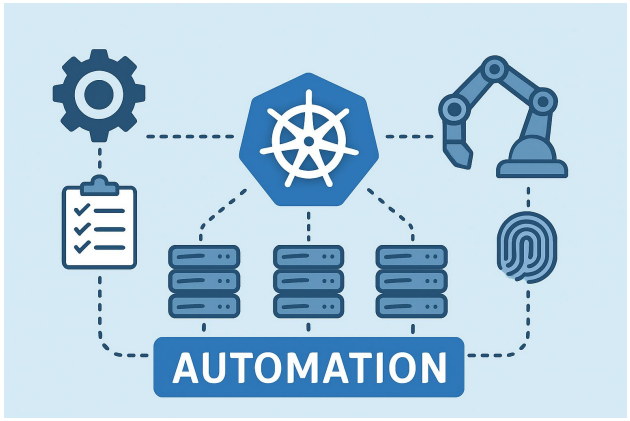


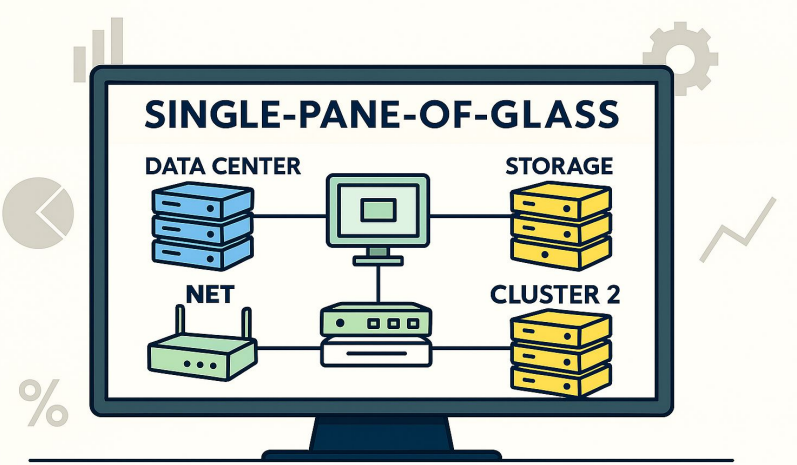


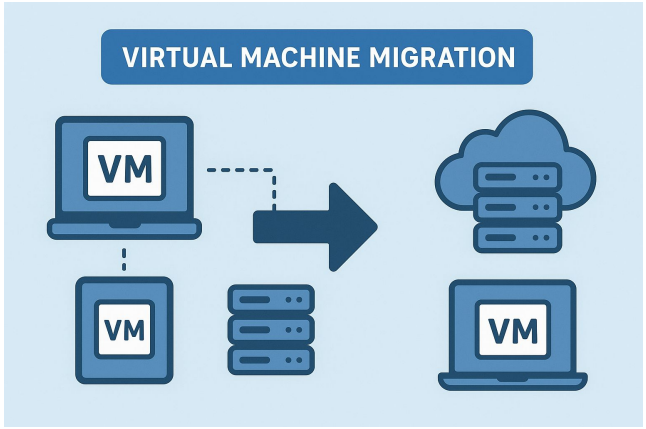
Posteitaliane | Red Hat

SL

Obiettivi aziendali | iniziativa





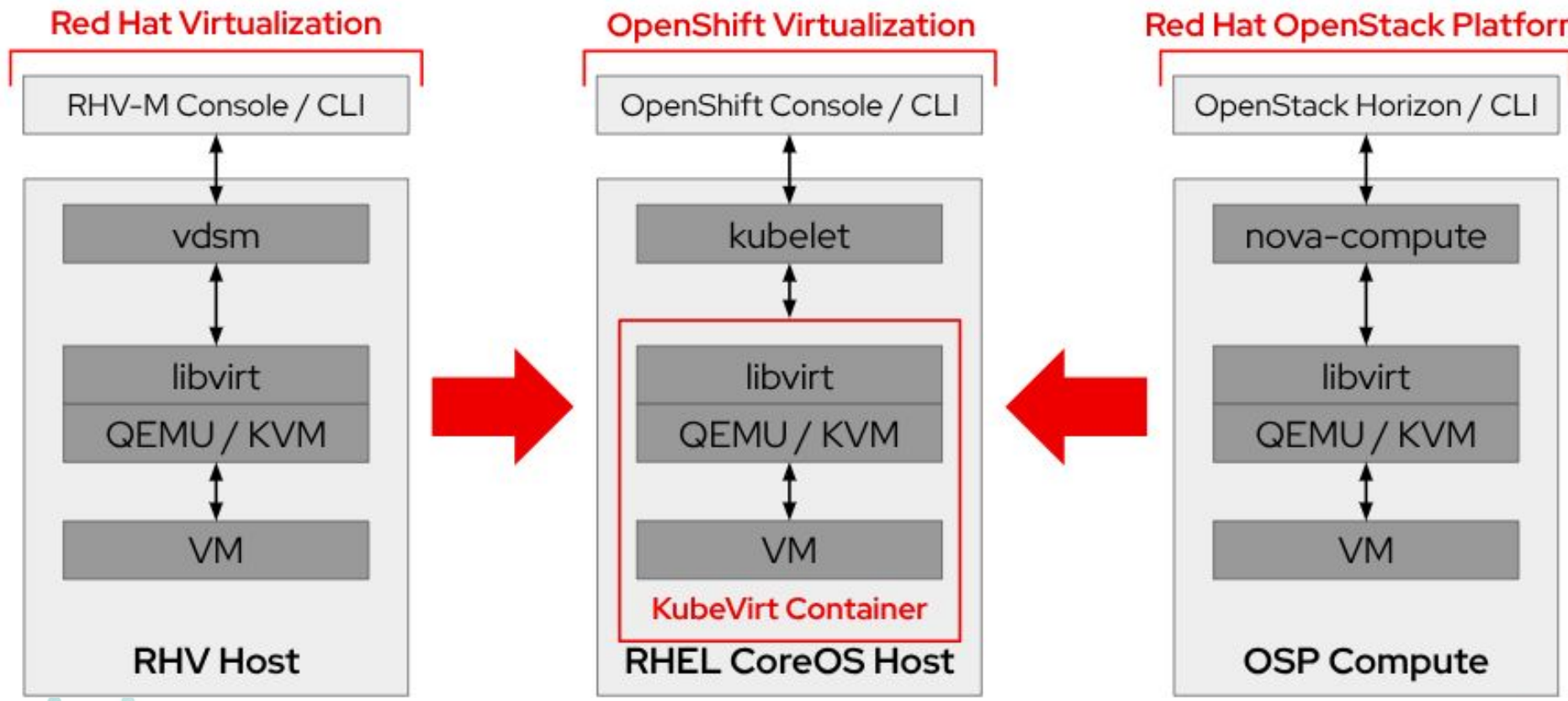


OPENSHIFT VIRTUALIZATION

Posteitaliane | Red Hat

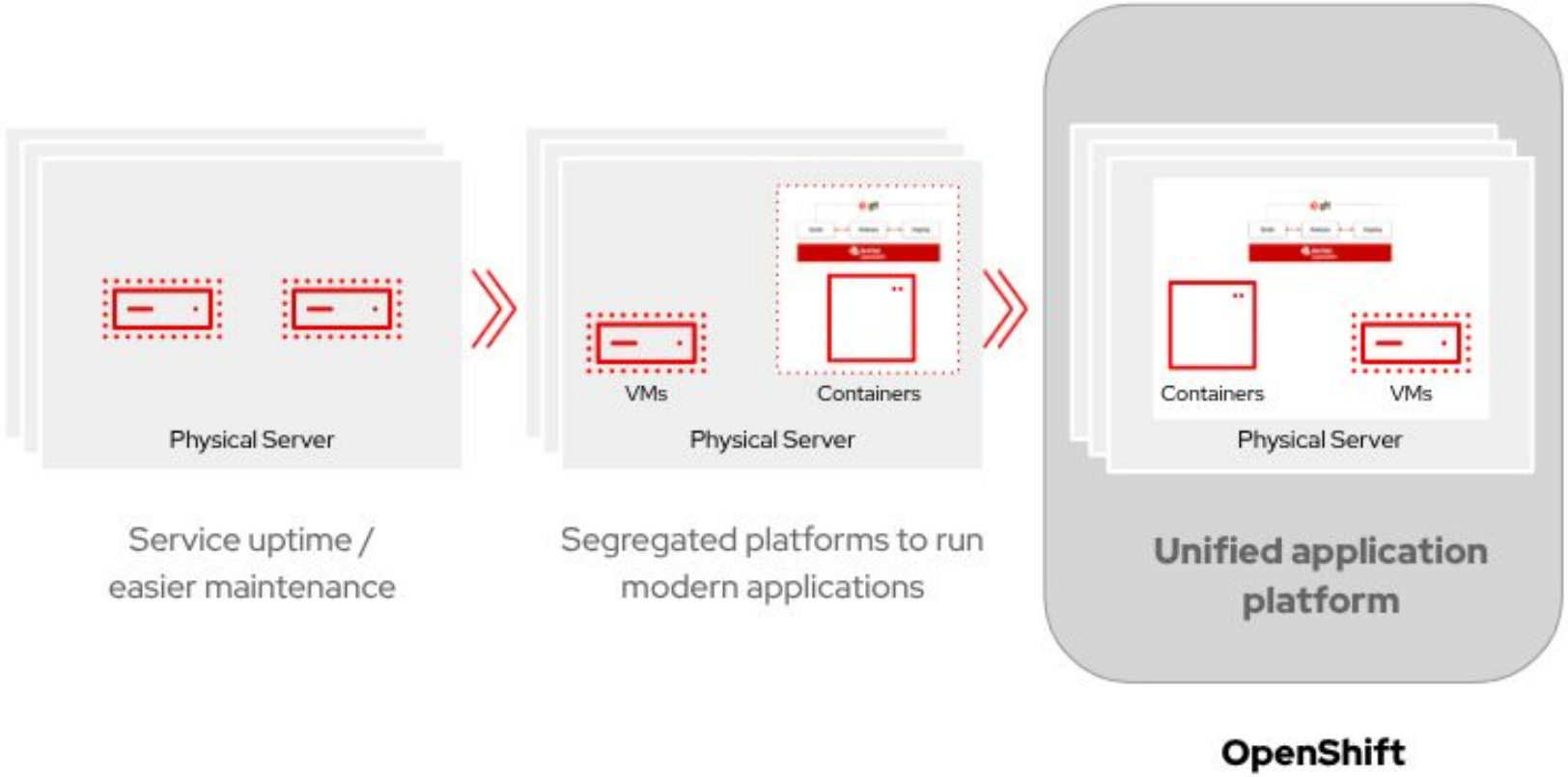
SL
VU

OpenShift Virtualization



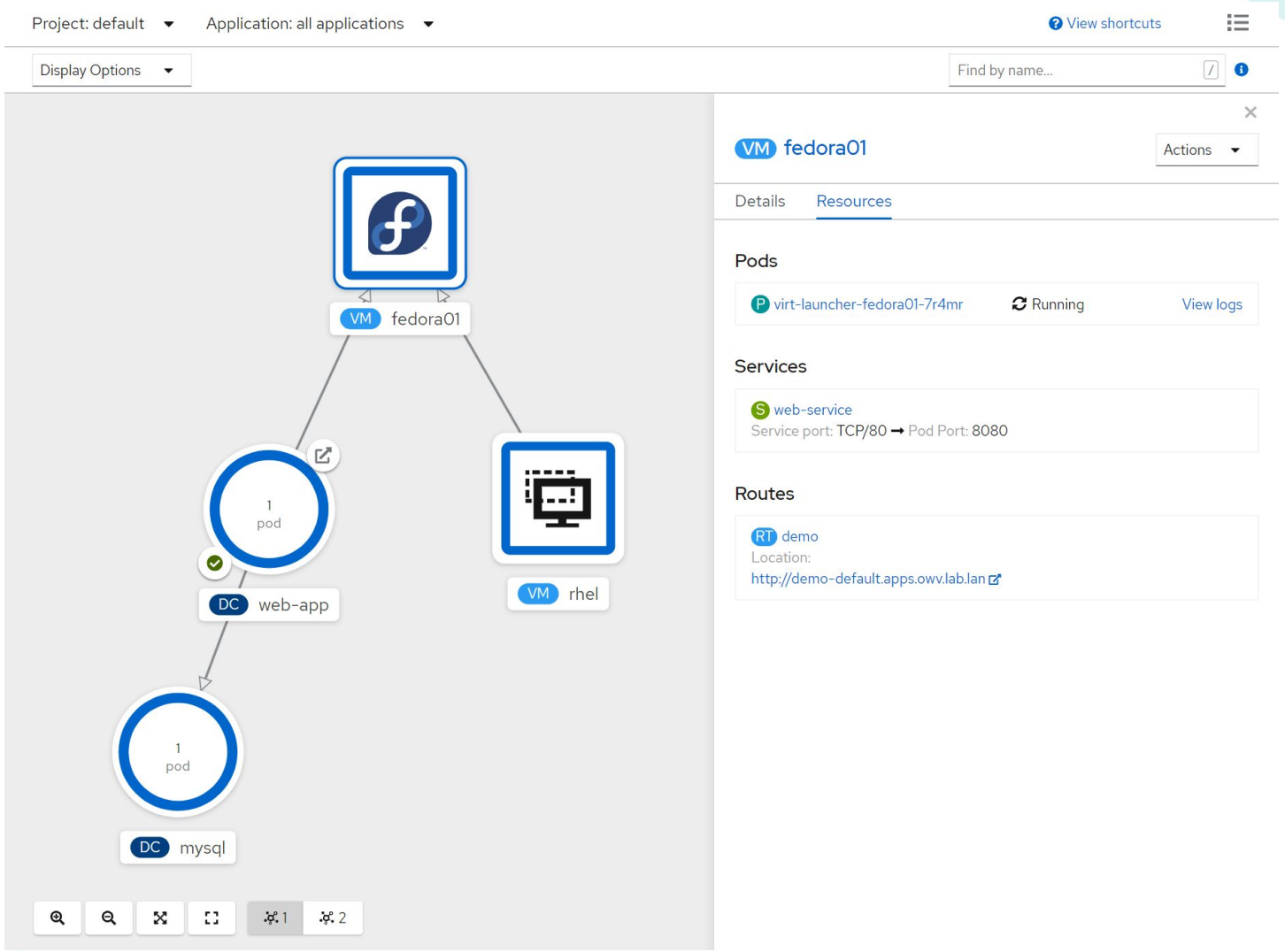
Posteitaliane | Red Hat

OpenShift Virtualization

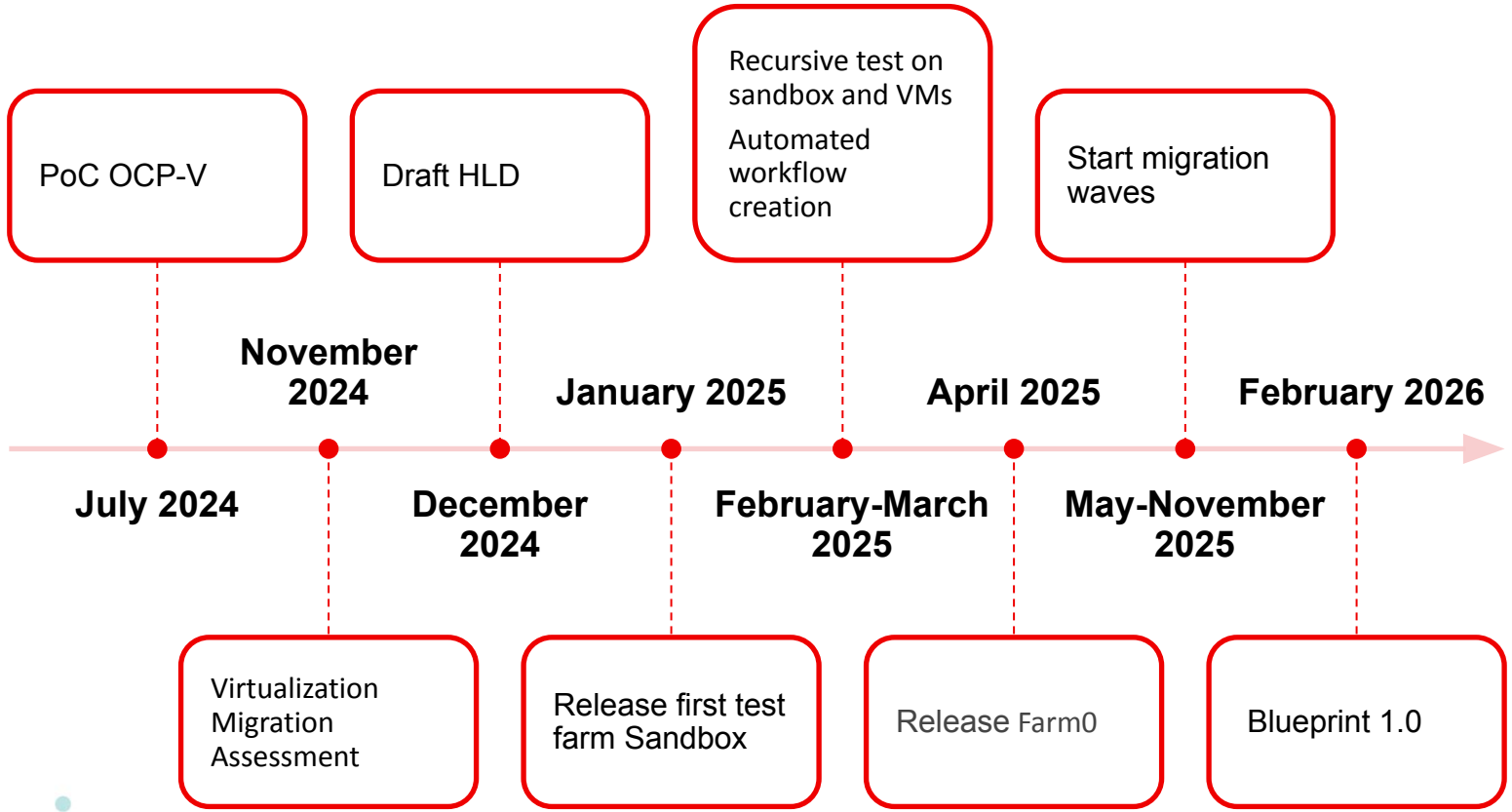


Posteitaliane | Red Hat

OpenShift Virtualization



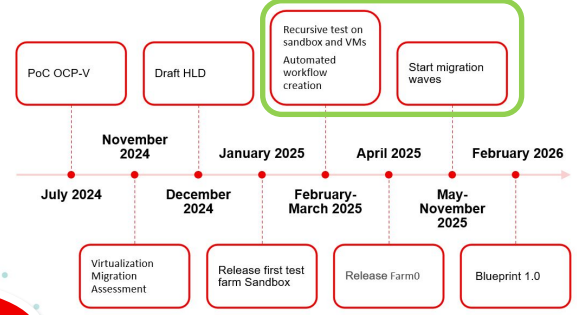
Il percorso

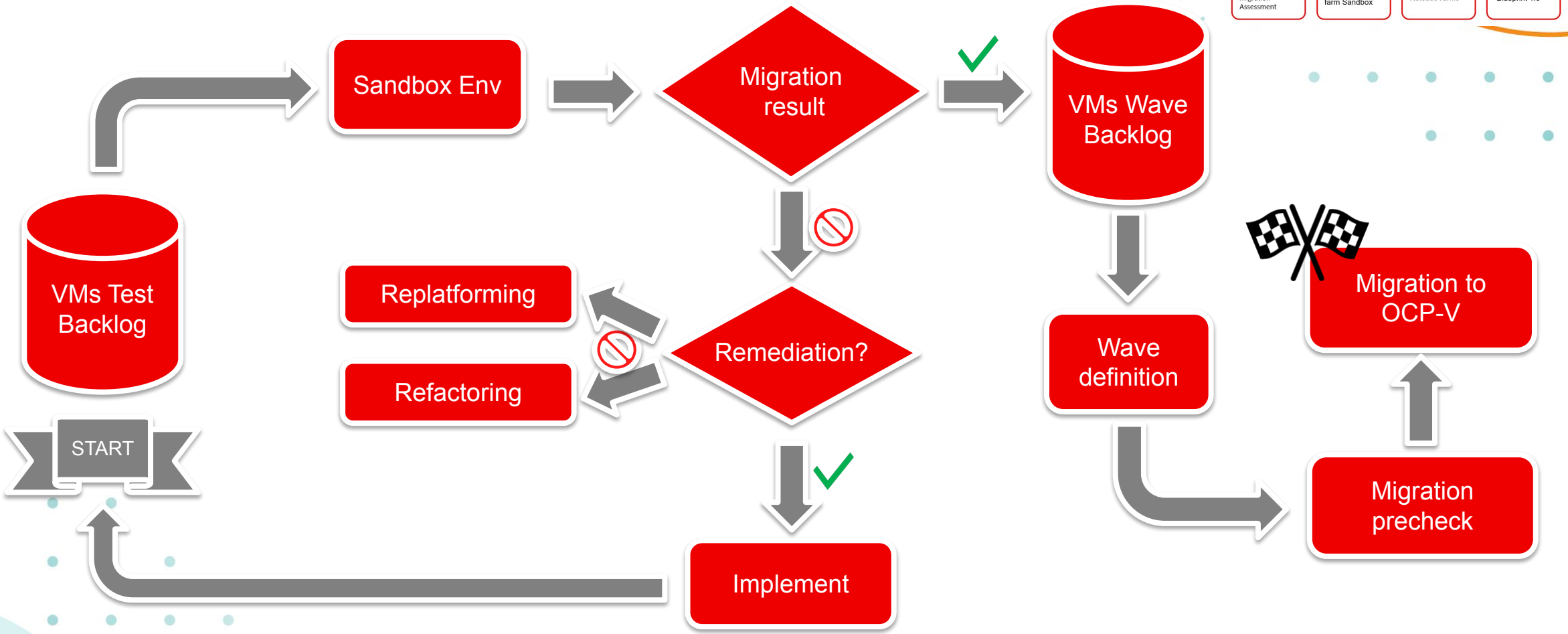


Posteitaliane | Red Hat

MS

Il percorso – Recursive test

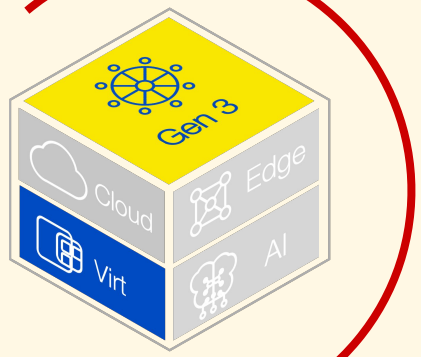




Posteitaliane | Red Hat

I pillar alla base della blueprint

~ 25 Cluster
~ 20K VirtualMachine
~ 650 Worker Nodes
~ 100k VCpu
~ 300TiB RAM



GIT AS A SINGLE SOURCE OF TRUTH

AUTOMATION

SINGLE PANE OF GLASS

SECURITY BY DESIGN

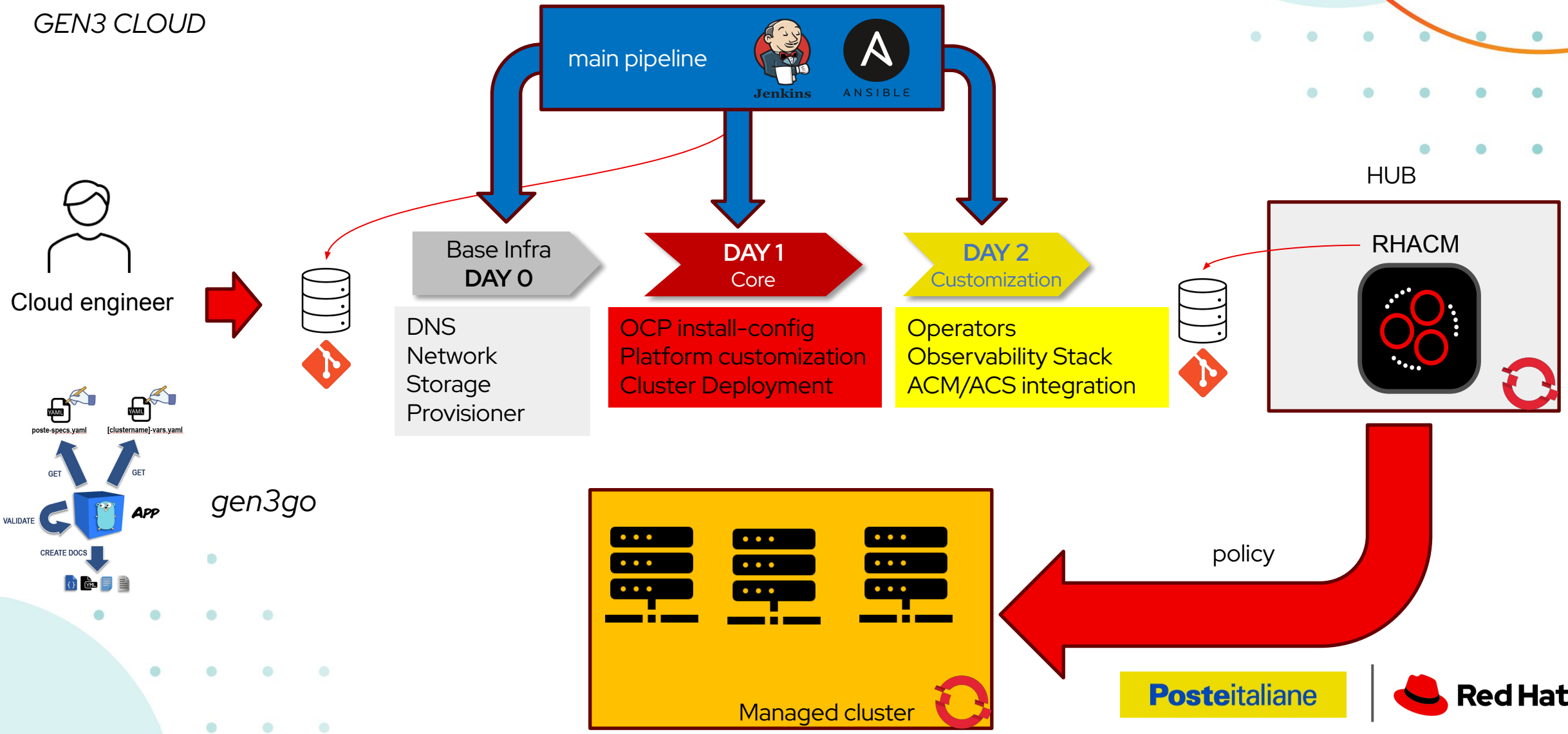
COLLABORATION

Posteitaliane | Red Hat

SL

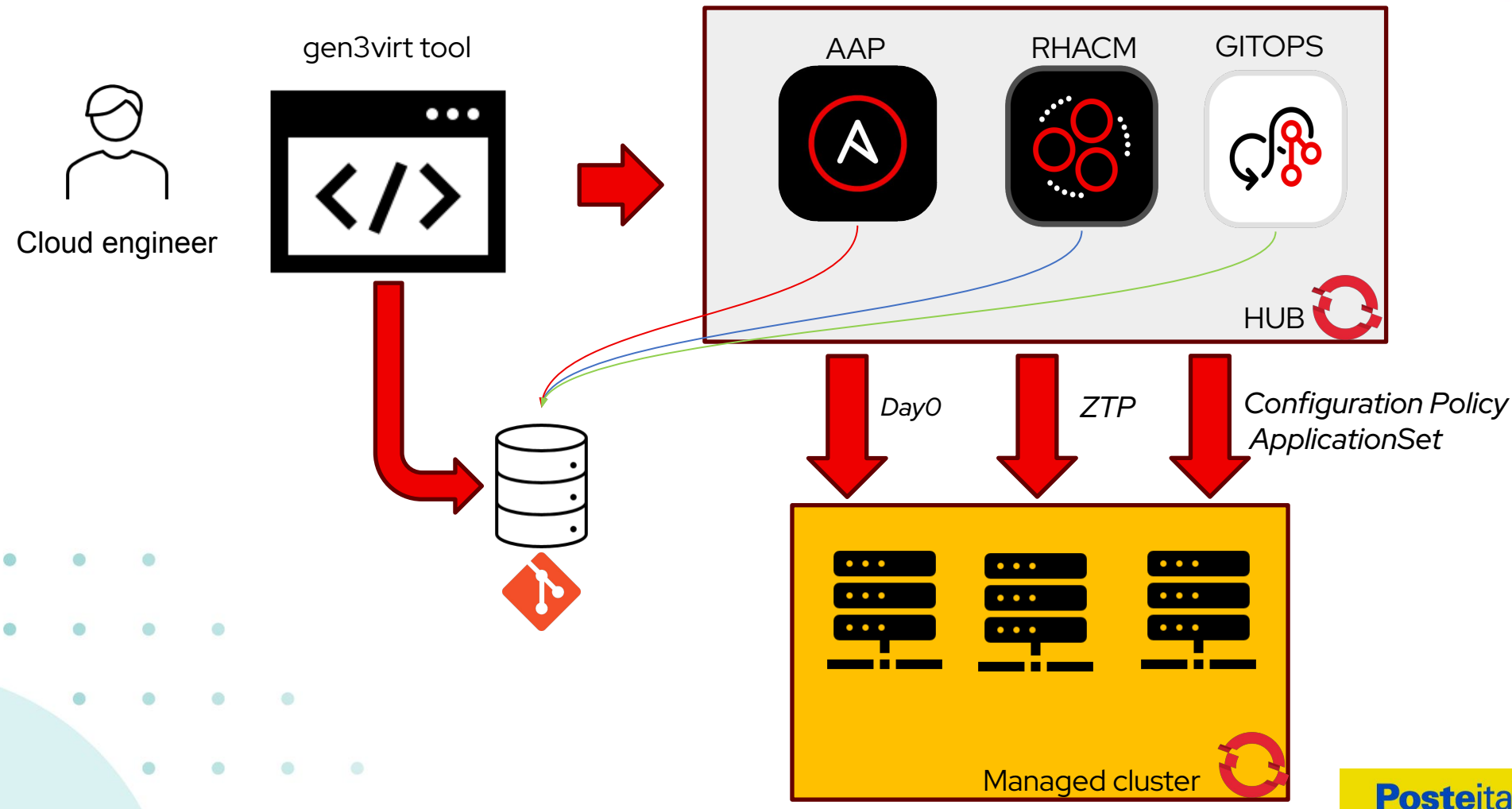
Automation Focus

GEN3 CLOUD



Automation Focus

GEN3 VIRT



L'esperienza della Farm0

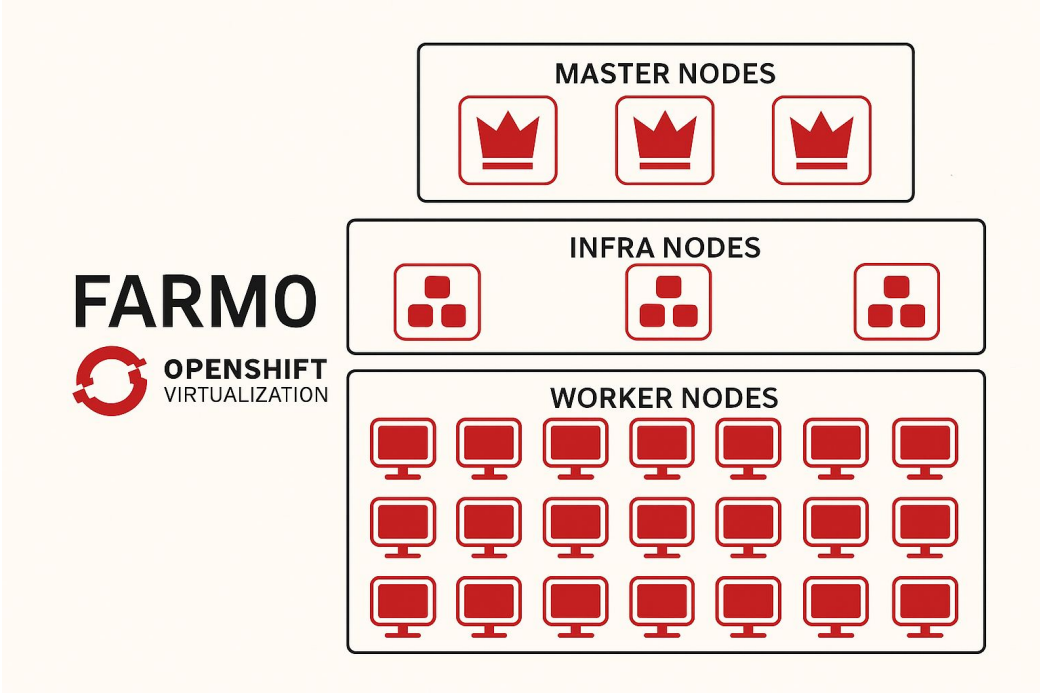
Obiettivo: portare in campo rapidamente un cluster reale per migrazione e greenfield

Vincoli:

- Utilizzo della blueprint 0 in corso di definizione
- Riutilizzo Hardware
- Adeguamento alla rete layerizzata del Datacenter

Caratteristiche:

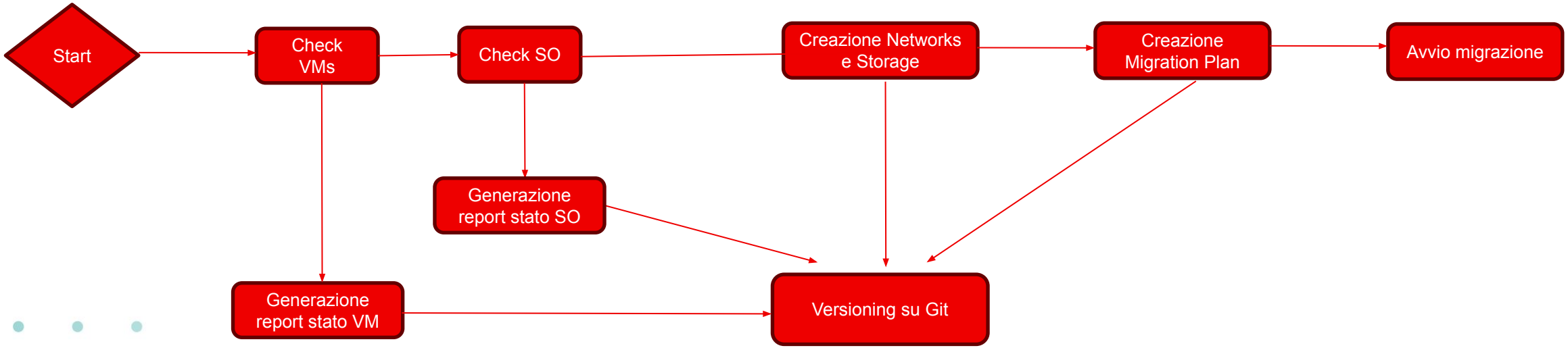
- Installazione semi automatica
- Porting di carico reale di vm di sviluppo e collaudo
- Coinvolgimento di tutte le strutture aziendali
- Test iterativi



Posteitaliane | Red Hat

PC

Migration Focus



Posteitaliane | Red Hat

PC

Migration Workload

Linee guida:

- Invarianza della configurazione di rete
- Aggregazione wave per applicativo
- Minimizzazione dei tempi di disservizio

Modalità

- Prerun check
- Migrazione tramite automazione Ansible
- Sanity check finali

Esempi di tempistiche

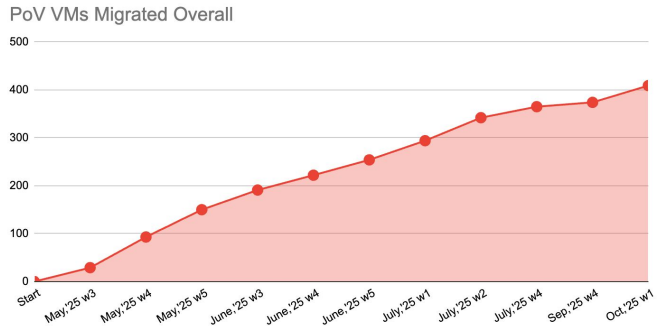
Tempi per il moving di 1 singolo disco da 500GB 1h 19 m

Tempi per il moving di 34 VMs e 2.4TB dischi 1h 35 m

Migrazione Farm 0

	WAVE PILOTA - PILOTA BIS		WAVE 0	WAVE 1	WAVE 2	WAVE 3	WAVE 3bis	WAVE 4	WAVE 5	WAVE 6	WAVE 7	WAVE 8	WAVE 8
Avvio Migrazione	14/05 ORE 08:00	15/05 ORE 19:00	20/05 ORE 11.00	27/05 ORE 11.00	11/06 ORE 06.30	17/06 ORE 18.00	18/06 ORE 06.30	25/06 ORE 12.00	02/07 ORE 06.30	09/07 ORE 06.30	22/07 ORE 18.45	29/09 ORE 13:30	10/10 ORE 13:30
Numero VM da migrare	18 VM	11 VM	64 VM	60 VM	39 VM	17 VM	14 VM	32 VM	40 VM	46 VM	26 VM	15 VM	34 VM
Numero Servizi	3	2	3	5	6	2	1	1	9	17	5	1	10

Perimetro consolidato - VM migrate al 11/10 - 416 su 416



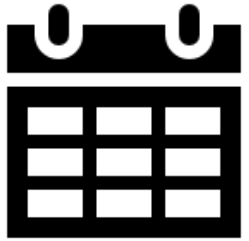
Posteitaliane | Red Hat

PC

Risultati | benefici(KPI) | lesson learned

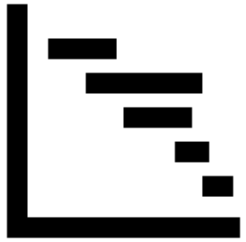
Team work

Collective buy-in



Planning

Don't trust, test it



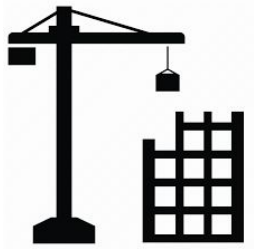
Expiration management

It's an enhancement and not just a replacement

Posteitaliane | Red Hat

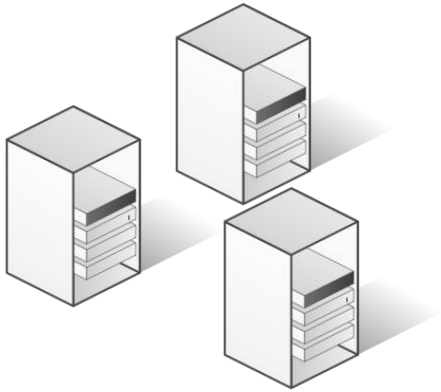
MS
PC

Next steps

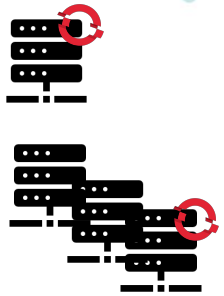


New Farm Planning

Disaster Recovery



Multiple Storage Provider



Single/Three node Flavour

NSX-T migration

- BGP
- Importing and exporting routes with the provider network
- VRF-LITE
- EVPN

Posteitaliane | Red Hat

SL
VU

Red Hat Summit

Connect

Grazie

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat Summit

Connect

Garantire la Conformità della PCI DSS 4.0 in Ambienti OpenShift Multi-Cluster

Un caso d'uso reale di implementazione in Poste Italiane

Posteitaliane | Red Hat

Riccardo Tagliaferri

Hybrid Cloud Architect
@Poste Italiane

Email: riccardo.tagliaferri@posteitaliane.it
Linkedin: https://it.linkedin.com/in/riccardotagliaferri

Francesco Cuccagna

Senior Cloud Consultant
Red Hat

Email: fcuccagn@redhat.com
Linkedin: https://it.linkedin.com/in/fcuccagna

Amedeo Salvati

Principal Cloud Consultant
Red Hat

Email: asalvati@redhat.com
Linkedin: https://it.linkedin.com/in/amedeos/en

Posteitaliane | Red Hat

Agenda

- Presentazione di Poste Italiane
- Obiettivi aziendali e iniziative di sicurezza
- Centralizzazione della sicurezza con Red Hat
- Il ruolo del Compliance e del File Integrity Operator
- Risultati, benefici e prossimi passi

Posteitaliane | Red Hat

Poste Italiane: Profilo della più grande infrastruttura di servizi in Italia

Chi Siamo

Con oltre **160 anni di storia**, Poste Italiane è un motore di innovazione e un punto di riferimento per famiglie, imprese e Pubblica Amministrazione.

- **Riconoscibilità e Capillarità:** Una realtà unica nel Paese per dimensioni e presenza sul territorio.

La Nostra Rete in Numeri

- 👥 **121.000** Dipendenti
- 🏢 **12.755** Uffici Postali
- 🤝 **46 Milioni** di Clienti

Le Nostre Aree di Business

- 📦 **Corrispondenza, Pacchi e Logistica**
  - *Servizi postali tradizionali e soluzioni avanzate per l'e-commerce.*
- 💰 **Servizi Finanziari e Assicurativi**
  - *Ampia gamma di prodotti per risparmio, investimenti e protezione.*
- 📱 **Sistemi di Pagamento e Telefonia**
  - *Soluzioni digitali innovative e servizi di telefonia mobile.*
- 💡 **Energia**
  - *Offerte di energia elettrica e gas per le famiglie italiane.*

Posteitaliane | Red Hat

Poste Italiane: Strategia, Performance e Valori

€ 12,6 mld

RICAVI TOTALI

€ 2,96 mld

RISULTATO OPERATIVO

€ 2 mld

UTILE NETTO

Strategia e impatto

Un percorso chiaro di sviluppo, innovazione e trasformazione digitale.

Impatto sul PIL (2018-2024): Generati circa € 90 Miliardi

Missione e Sostenibilità

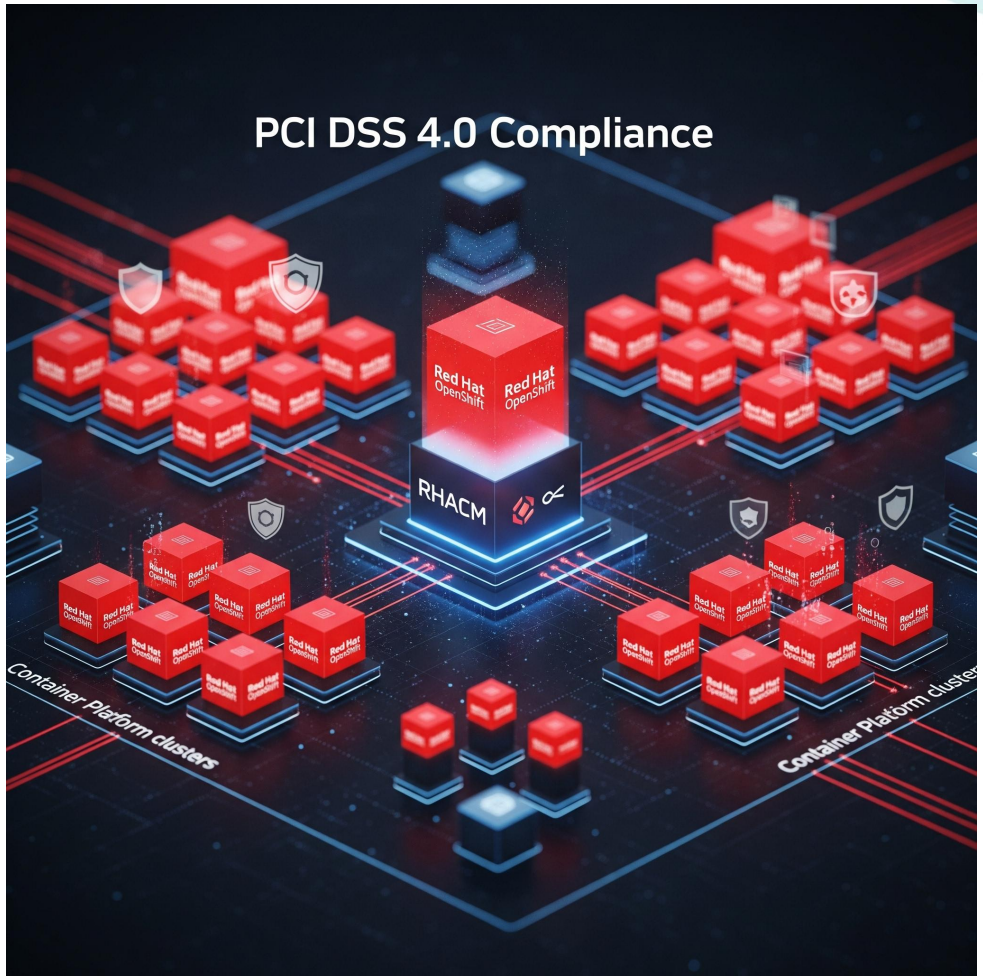
"Crescere responsabilmente per il successo sostenibile, l'innovazione, la digitalizzazione e la coesione sociale del Paese".

Un piano integrato per contrastare il cambiamento climatico, valorizzare le persone e agire con etica e trasparenza.

Posteitaliane | Red Hat

Obiettivi aziendali | iniziativa

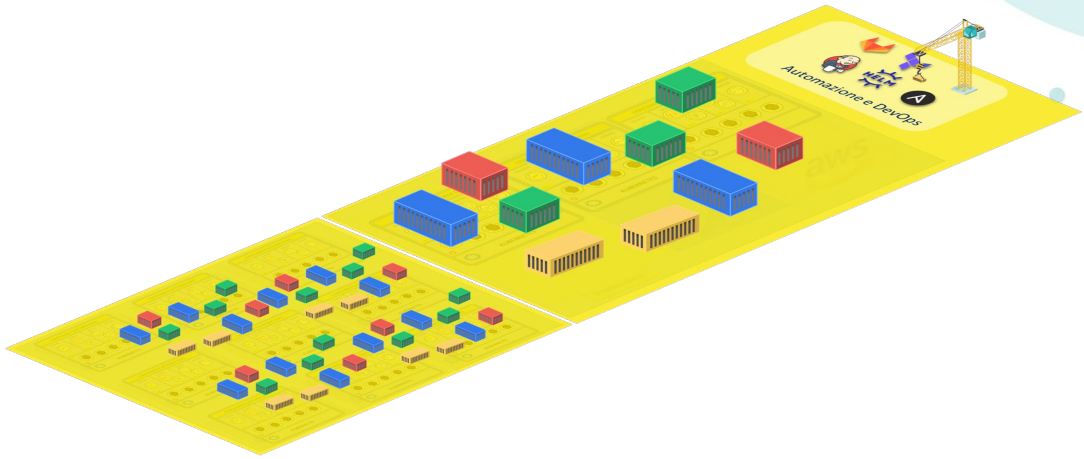
- Aumentare la sicurezza della Container Platform.
- Ottenere e garantire la conformità PCI DSS 4.0.
- Implementare soluzioni innovative in ambienti multi-cluster.
- Consolidare la fiducia dei clienti nei servizi.

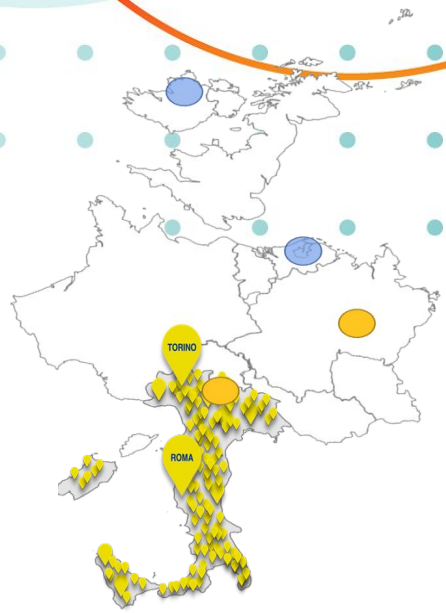


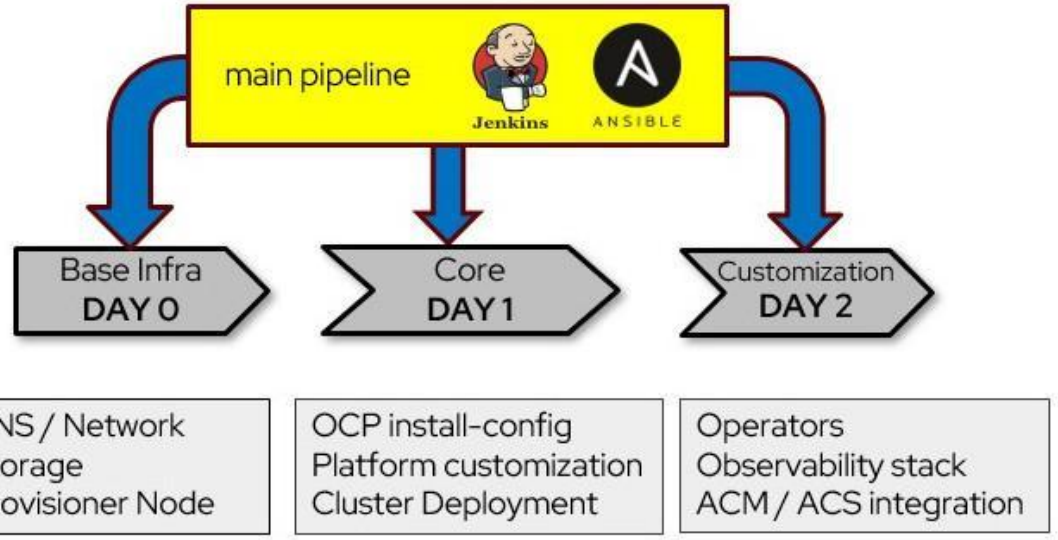
Posteitaliane | Red Hat

Container Platform OCP 4 / Gen 3

- La Container Platform di Poste Italiane è completamente automatizzata in tutte le fasi del ciclo di vita.
- Offre un'infrastruttura standard ed immutabile per il development ed il delivery di workload ibridi.
- Dispone di un "single pane of glass" per il mantenimento delle configurazioni, garantendo così consistenza, resilienza e contenimento dei costi.



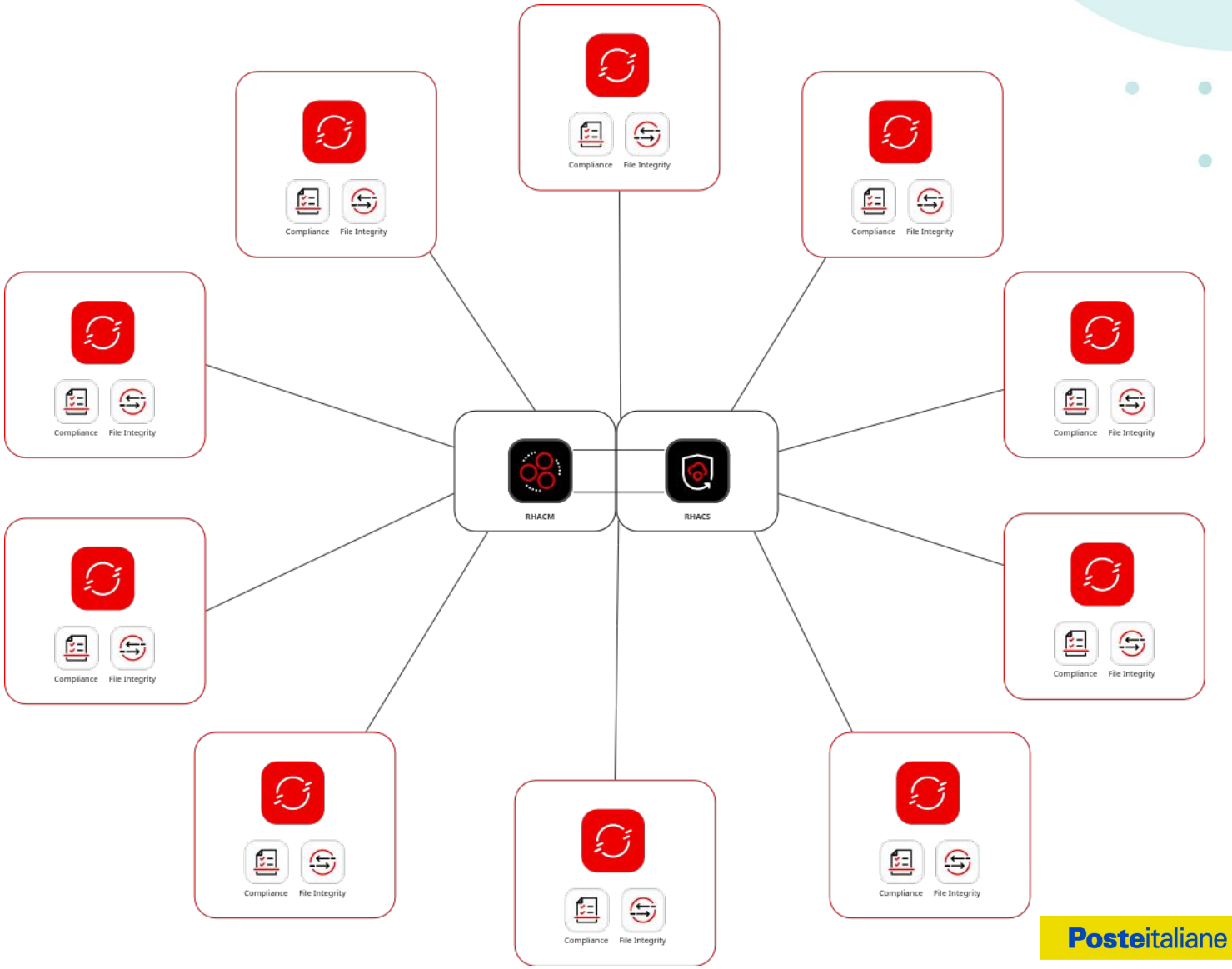




Posteitaliane | Red Hat

Centralizzazione tramite ACM e ACS

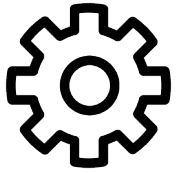
- Centralizzazione della gestione della sicurezza e della conformità
- Miglioramento della visibilità e del controllo sull'infrastruttura
- Automazione delle politiche di sicurezza per ridurre gli errori
- Riduzione del tempo e dei costi per raggiungere la conformità



Posteitaliane | Red Hat

Compliance Operator

Automating compliance checks for OpenShift and CoreOS



compliance-operator

L'operatore consente all'amministratore di descrivere lo stato di conformità desiderato di un cluster e fornisce una panoramica delle lacune e dei modi per correggerle.

OpenSCAP

Strumento certificato NIST per analizzare e applicare le policy di sicurezza fornite dal contenuto.

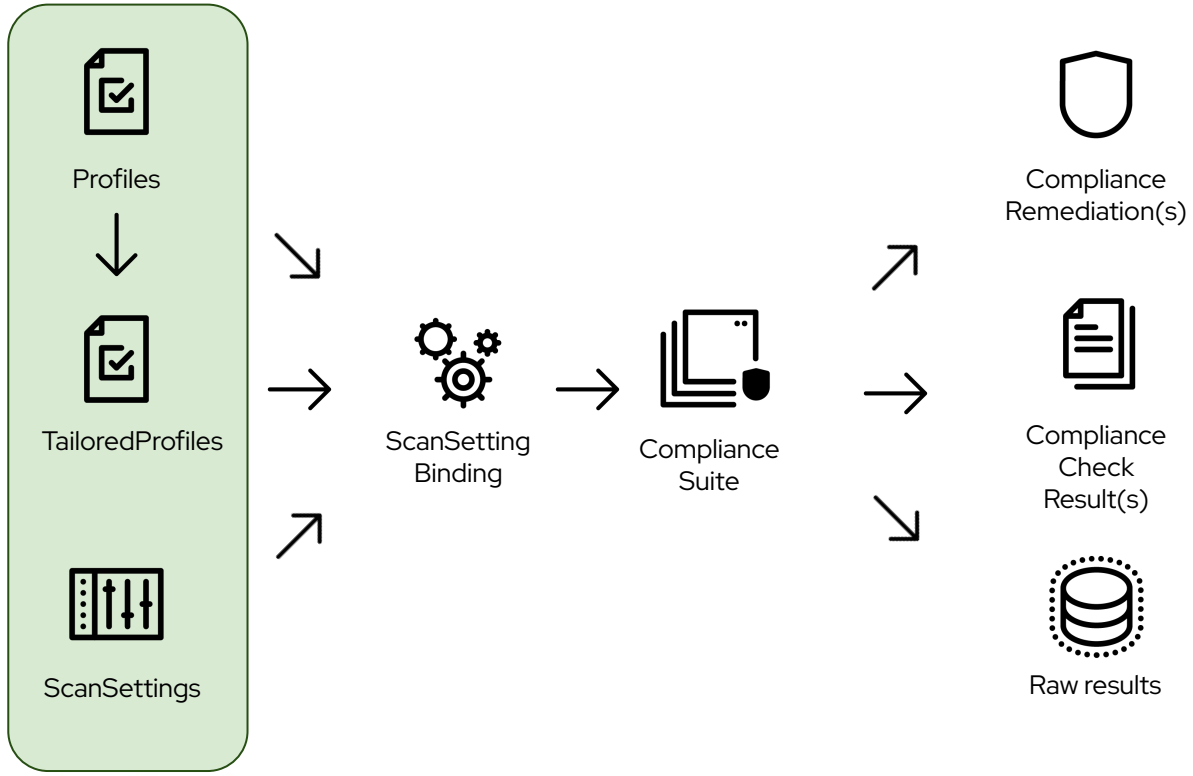
Content

I controlli di conformità stessi vengono forniti tramite contenuti SCAP, con un ciclo di vita indipendente dall'operatore o dallo scanner OpenSCAP. Questo avviene come parte del progetto **ComplianceAsCode**.

Posteitaliane | Red Hat

Compliance Operator

Overview delle componenti

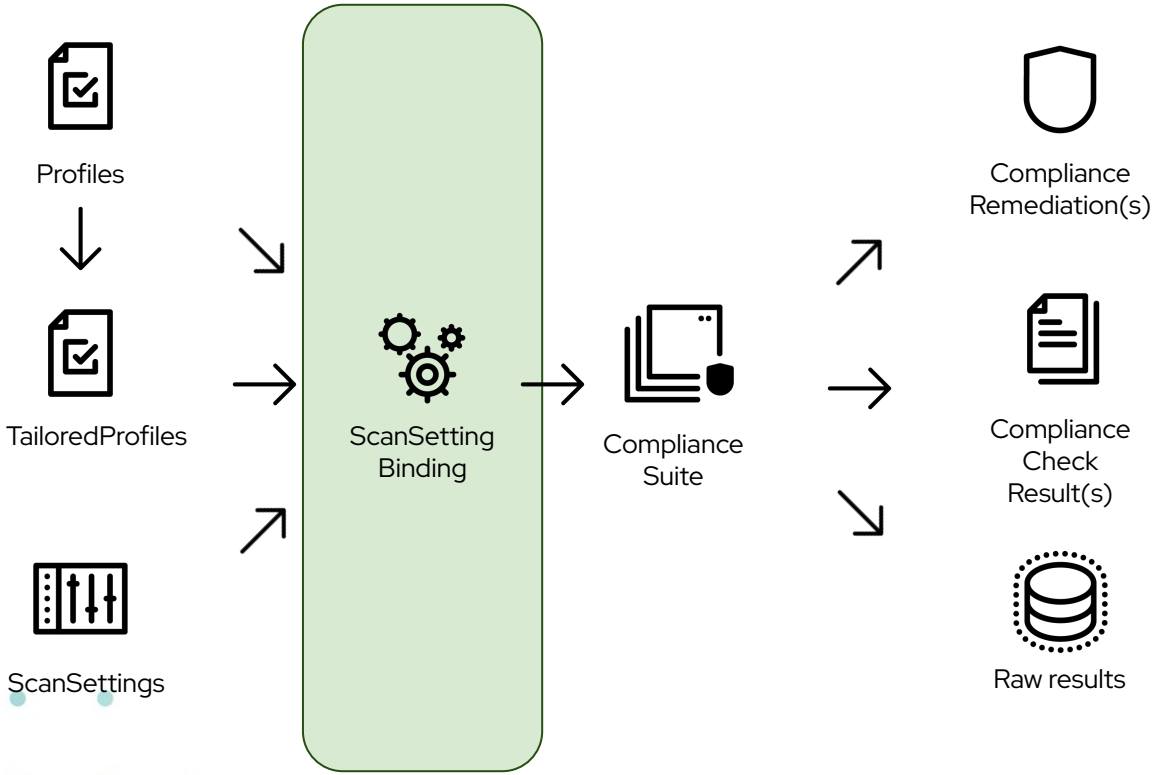


Con cosa devi essere conforme?
Seleziona una policy oppure creane una personalizzata che soddisfi le tue esigenze.
Qual è la policy dell'organizzazione in materia di scansione e monitoraggio dei sistemi?

Posteitaliane | Red Hat

Compliance Operator

Overview delle componenti

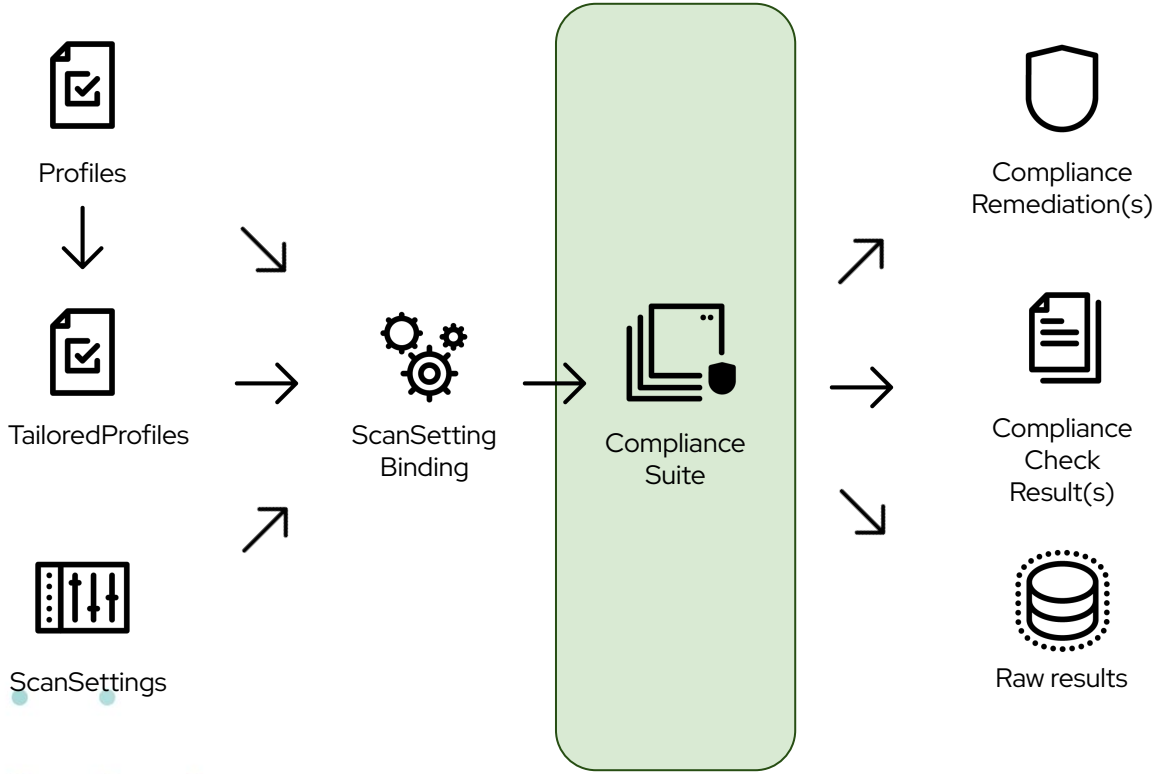


Chiedi all'operatore ciò che desideri.

Posteitaliane | Red Hat

Compliance Operator

Overview delle componenti



Controlla lo stato degli scan

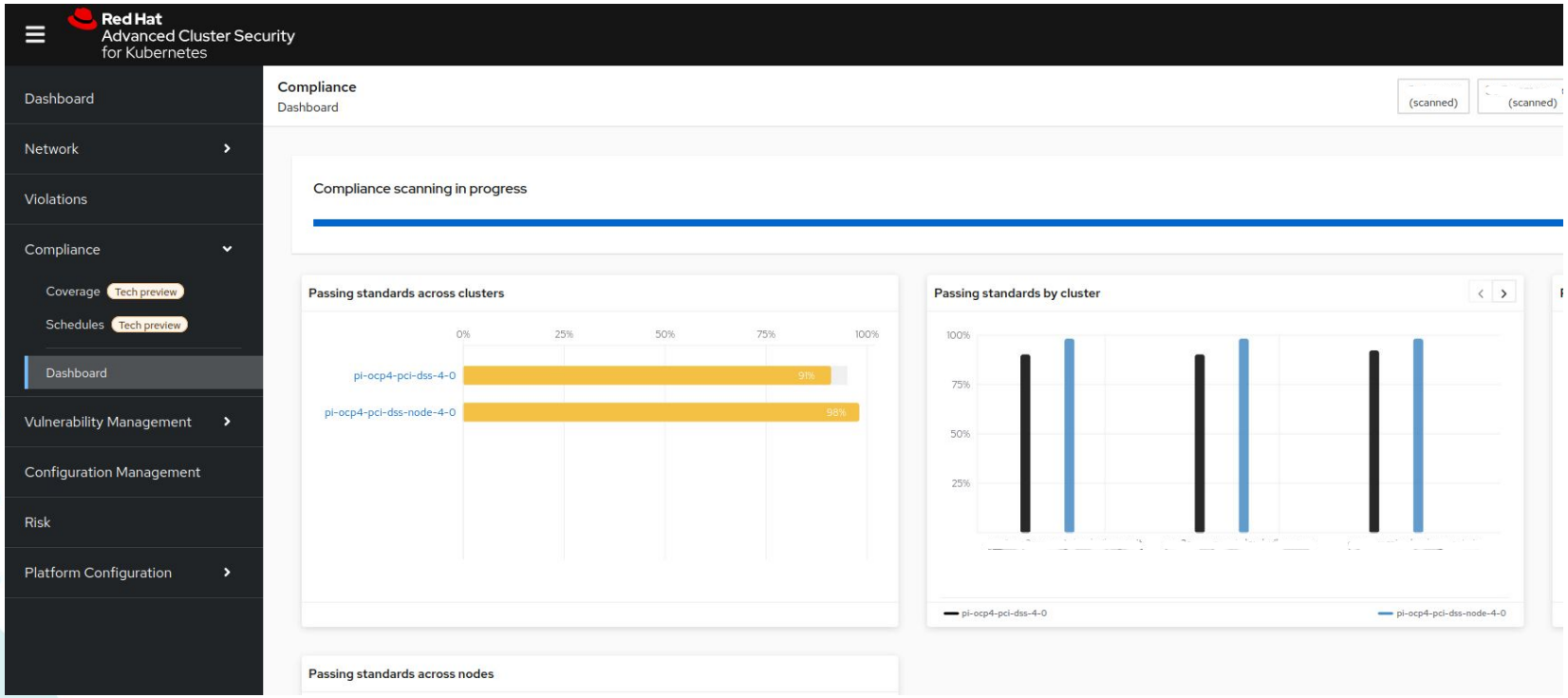
Posteitaliane | Red Hat

Compliance Operator

Compliance Suites

Monitor the status of the scans.

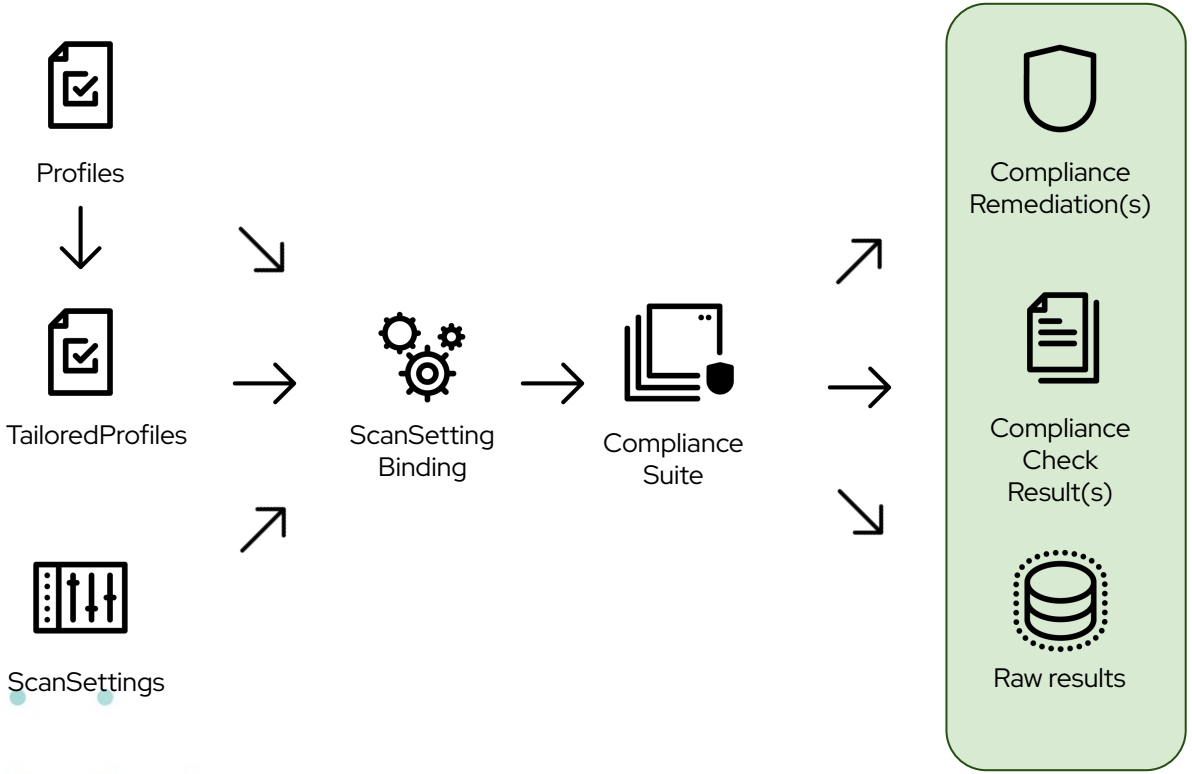
```
$ oc get compliancesuites.compliance.openshift.io -n openshift-compliance pi-compliance-pci-dss-4-0
NAME                          PHASE   RESULT
pi-compliance-pci-dss-4-0     DONE    NON-COMPLIANT
```



Posteitaliane | Red Hat

Compliance Operator

Overview delle componenti



Visualizza lo stato di conformità del sistema.

Verifica le possibili correzioni e le applica se necessario.

Scarica i risultati grezzi e li fornisce agli auditor.

Posteitaliane | Red Hat

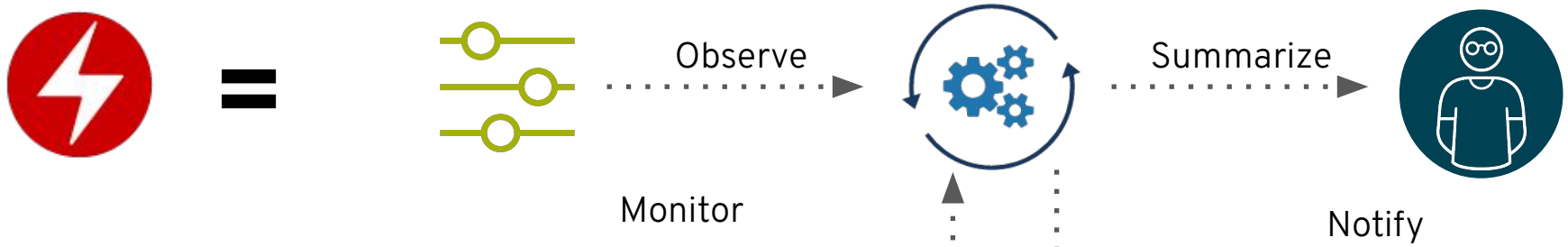
File Integrity Operator

AIDE

Enable FileIntegrity Checking

1 L'operatore esegue lo scan sui nodi selezionati e popola il database di AIDE

3 Gli Amministratori possono esaminare il risultato degli scan



2 Gli scan successivi confrontano i risultati sul database di AIDE

Red Hat OpenShift Container Platform

AIDE Configuration

Deploy AIDE Pods

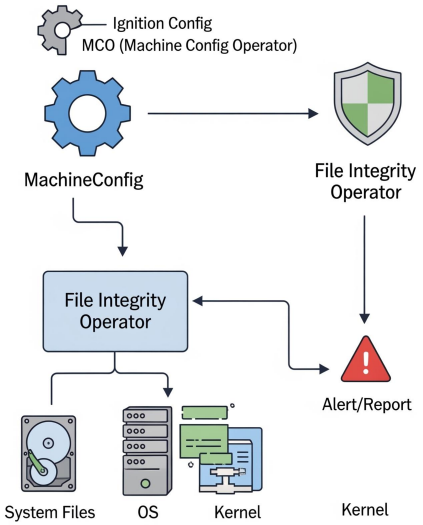
Scan Nodes

Notification (fileIntegrityNodeStatus)

Posteitaliane | Red Hat

File Integrity Operator: MachineConfig

- Le MachineConfig definiscono lo stato desiderato dei nodi.
- Il Machine Config Operator applica le configurazioni ai nodi.
- Le modifiche previste non attivano avvisi d'integrità.
- L'aggiornamento dei nodi sospende le scansioni AIDE.



Posteitaliane | Red Hat

Risultati | benefici(KPI) | lesson learned

- Integrazione di un tool antimalware
- Monitorare costantemente la conformità e le minacce
- Ridurre le potenziali superfici di attacco
- Reagire tempestivamente a violazioni della piattaforma

Posteitaliane | Red Hat

Prossimi passi

- Definire la roadmap per l'implementazione sulla Generative AI
- Continuare a collaborare con i team per il garantire la compliance alle evoluzione delle normative
- Shift left dei controlli di conformità PCI ed allarmi di intrusione

Posteitaliane | Red Hat

Red Hat Summit

Connect

Grazie

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat Summit

Connect

Tutto ciò che devi sapere su AI Agents, MCP servers e RAG per realizzare un’applicazione moderna che integra l’AI in modo concreto

Natale Vinto

Developer Experience, Evangelism Director
Red Hat

Marco Caimi

Senior Specialist Solution Architect - AI
Red Hat

Red Hat AI

Accelerate the development and delivery of AI solutions across hybrid-cloud environments

Increase efficiency with **fast, flexible and efficient inferencing**

Simplified and consistent experience for **connecting models to data**

Accelerate Agentic AI deployments

Flexibility and consistency when **scaling AI across the hybrid cloud**

Introducing OpenShift AI 3

Flexible and Efficient Inference

- GA distributed inference (llm-d)
- New validated and optimized models
- vLLM enhancements
- LLM Compressor GA

Connecting Data to Models

- Modular and extensible approach for: data ingestion, synthetic data generation, tuning, evaluations.
- RAG enhancements & partner integrations
- Feature Store GA

Agentic AI

- AI experiences: AI hub and gen AI studio
- Model Context Protocol support & MCP Server access in gen AI studio
- Llama Stack API integration

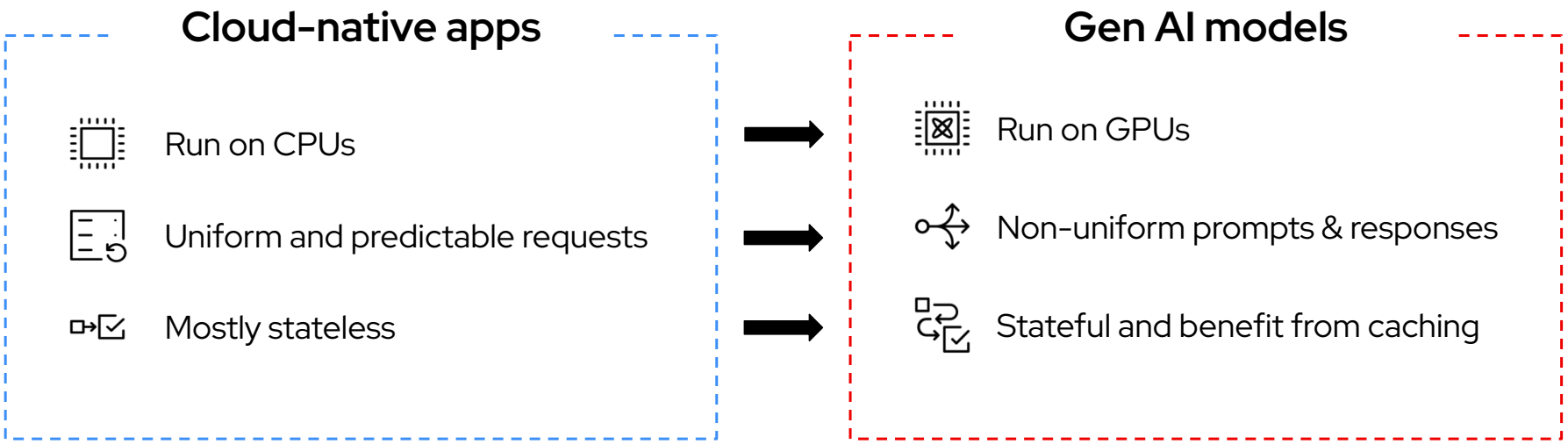
AI Platform

- Model catalog and registry GA
- Model as a Service provider enhancements and API Mgt integration
- GPU as a Service enhancements

Single platform to run any model, on any accelerator, on any cloud

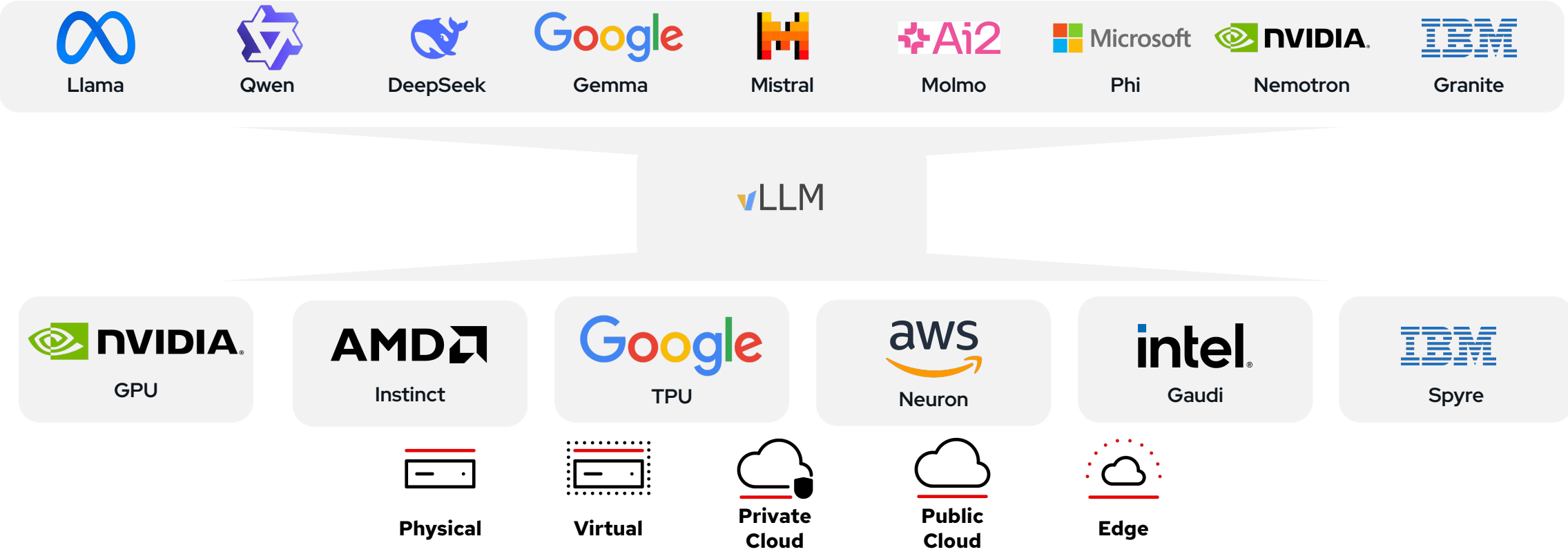
Overcoming the generative AI challenges

Running AI efficiently



vLLM Inference Server

vLLM v1 with enhanced performance, expanded model and hardware support



Single platform to run any model, on any accelerator, on any cloud

Red Hat

Inference at scale everywhere

Distributed, scalable gen AI inference for Enterprise AI

Red Hat AI

Now includes llm-d

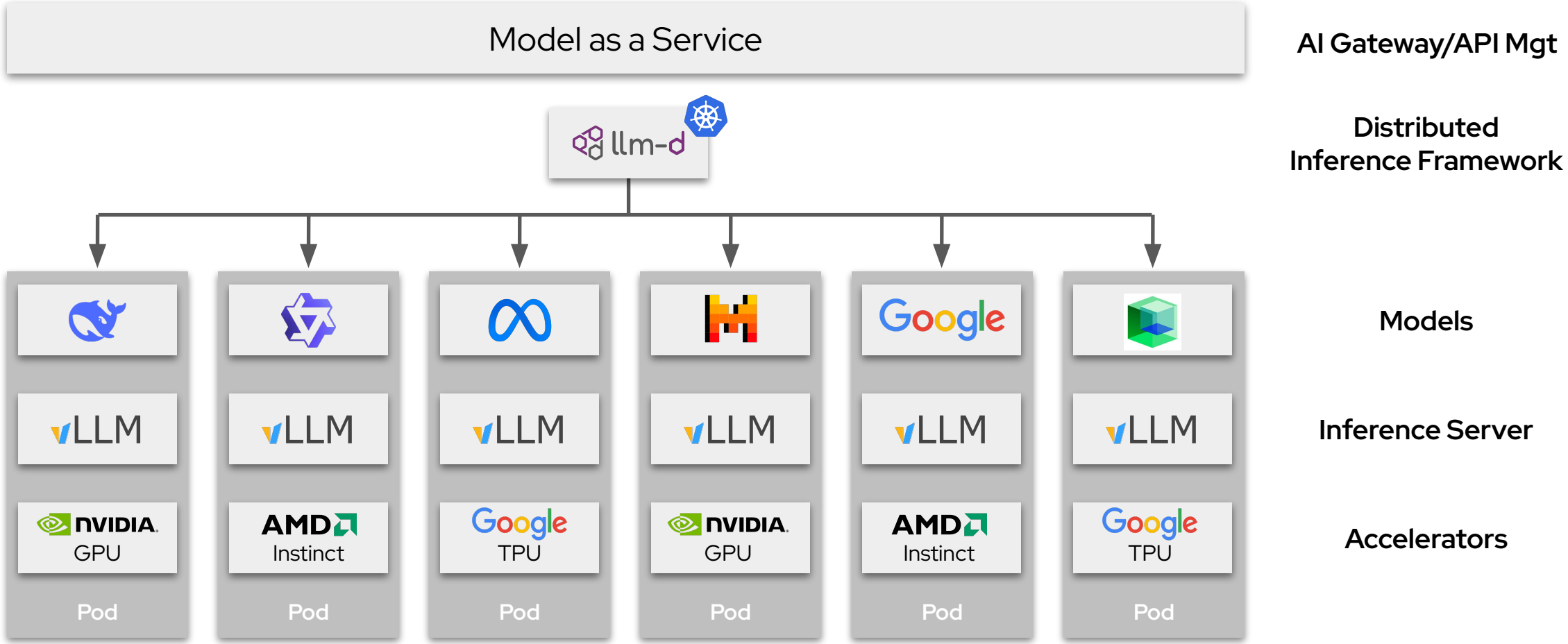
llm-d reimagines how LLMs run on Kubernetes

- Lower infrastructure spend for AI
- Great performance at larger scale
- Seamless scaling for unpredictable AI demand

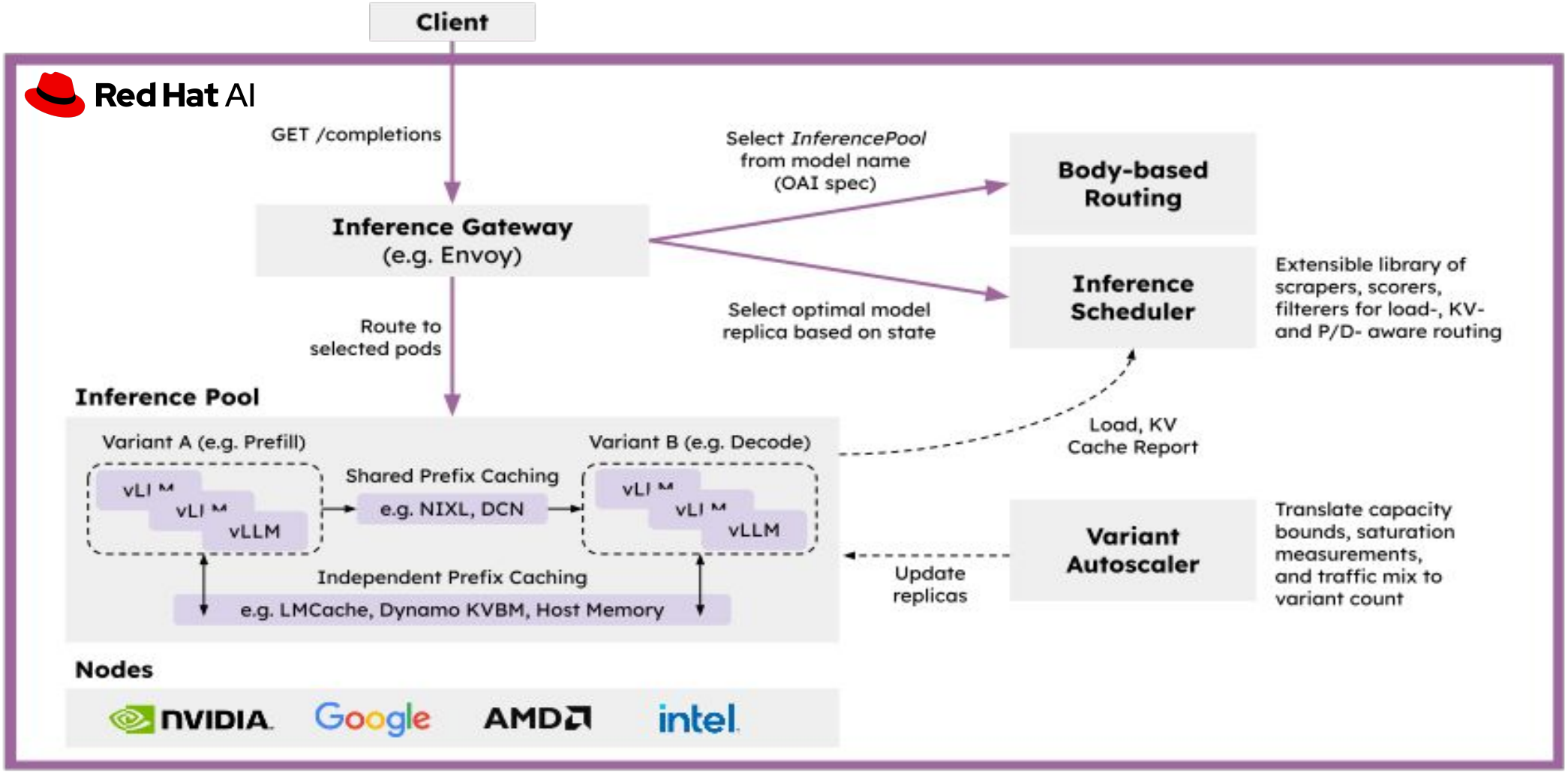
Deliver faster, cheaper, and more manageable AI systems for enterprise production

Enterprise GenAI Inference Platform

Holistic approach to optimize and operationalize deployment and scaling of open-source LLMs



Inference at scale everywhere



Red Hat AI Model Repository

New validated and optimized models and LLM Compressor now GA

Broad Collection of models

Llama

Qwen

Gemma

Mistral

DeepSeek

Phi

Molmo

Granite

Nemotron

Choice of Models

- Transformers (Dense, MOE), Multi-modal LLMs, Embeddings Models, Hybrid / Novel Attention, Vision
- Hugging Face compatible (safe tensors), OCI-compatible containers

Validated models

- Tested using realistic scenarios
- Assessed for performance across a range of hardware
- Done using GuideLLM benchmarking and LM Eval Harness

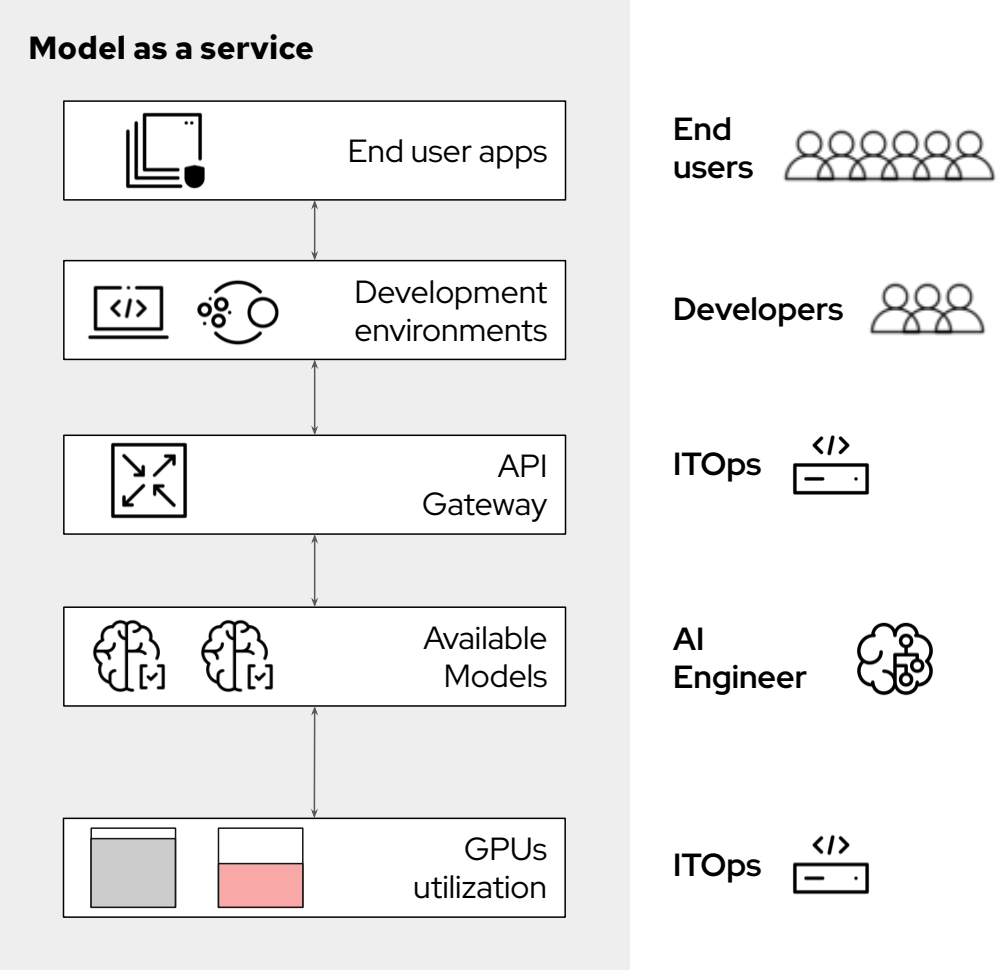
Optimized models

- Compressed for speed and efficiency
- Designed to run faster, use fewer resources, maintain accuracy
- Done using LLM Compressor with latest algorithms

Model-as-a-Service in OpenShift AI (Dev Preview)

Offering AI **models as the service** to a larger audience

- IT serves common models centrally
  - Generative AI focus, applicable to any model
  - Centralized pool of hardware
  - Platform Engineering for AI
- Models available through the RHOAI console
- Developers consume models, build AI applications
  - For end users (private assistants, etc)
  - To improve products or services through AI
- Shared Resources business model keeps costs down

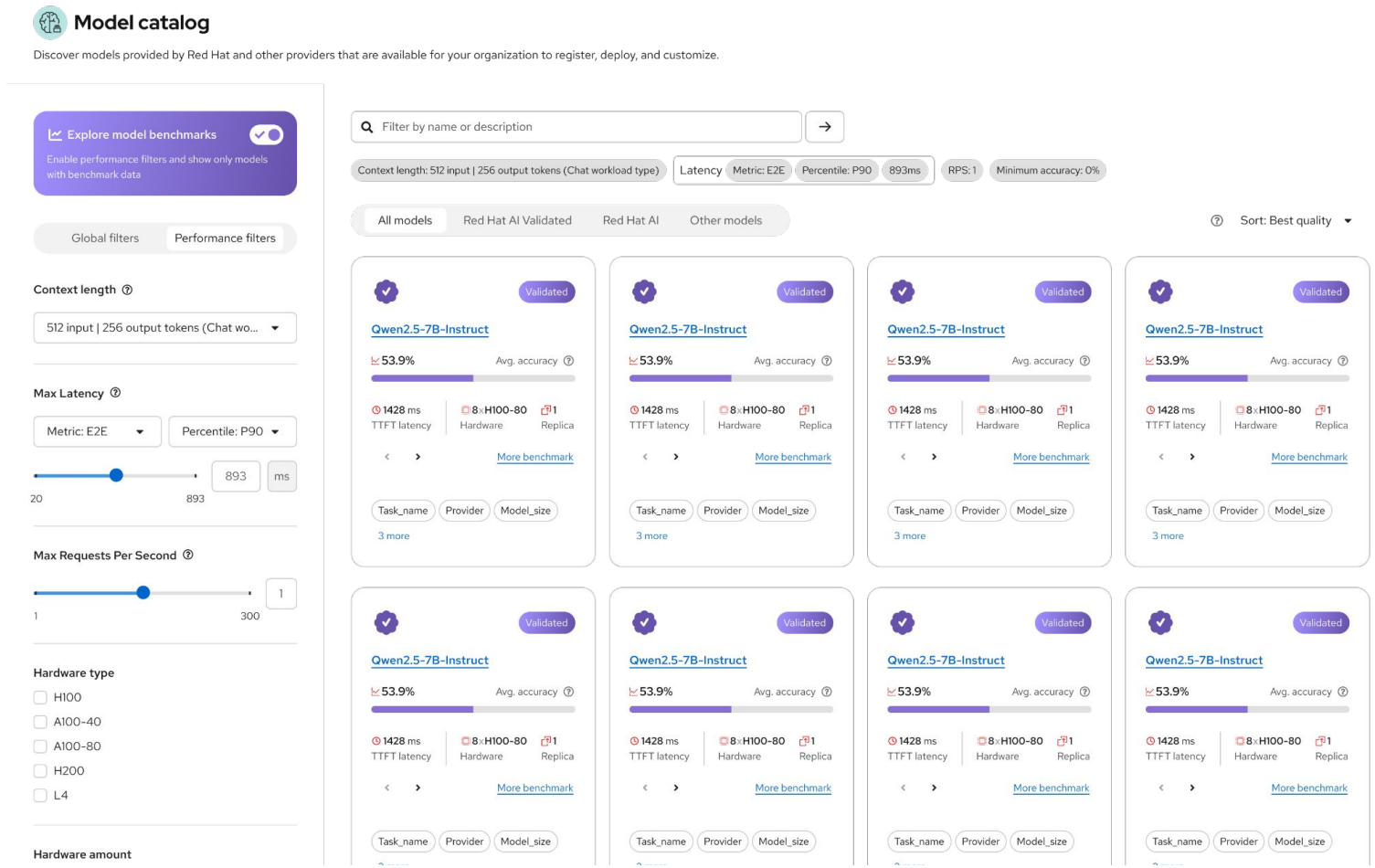


AI Hub

- AI Assets:
  - Starting with Models in 3.0.

    Future Assets: MCP

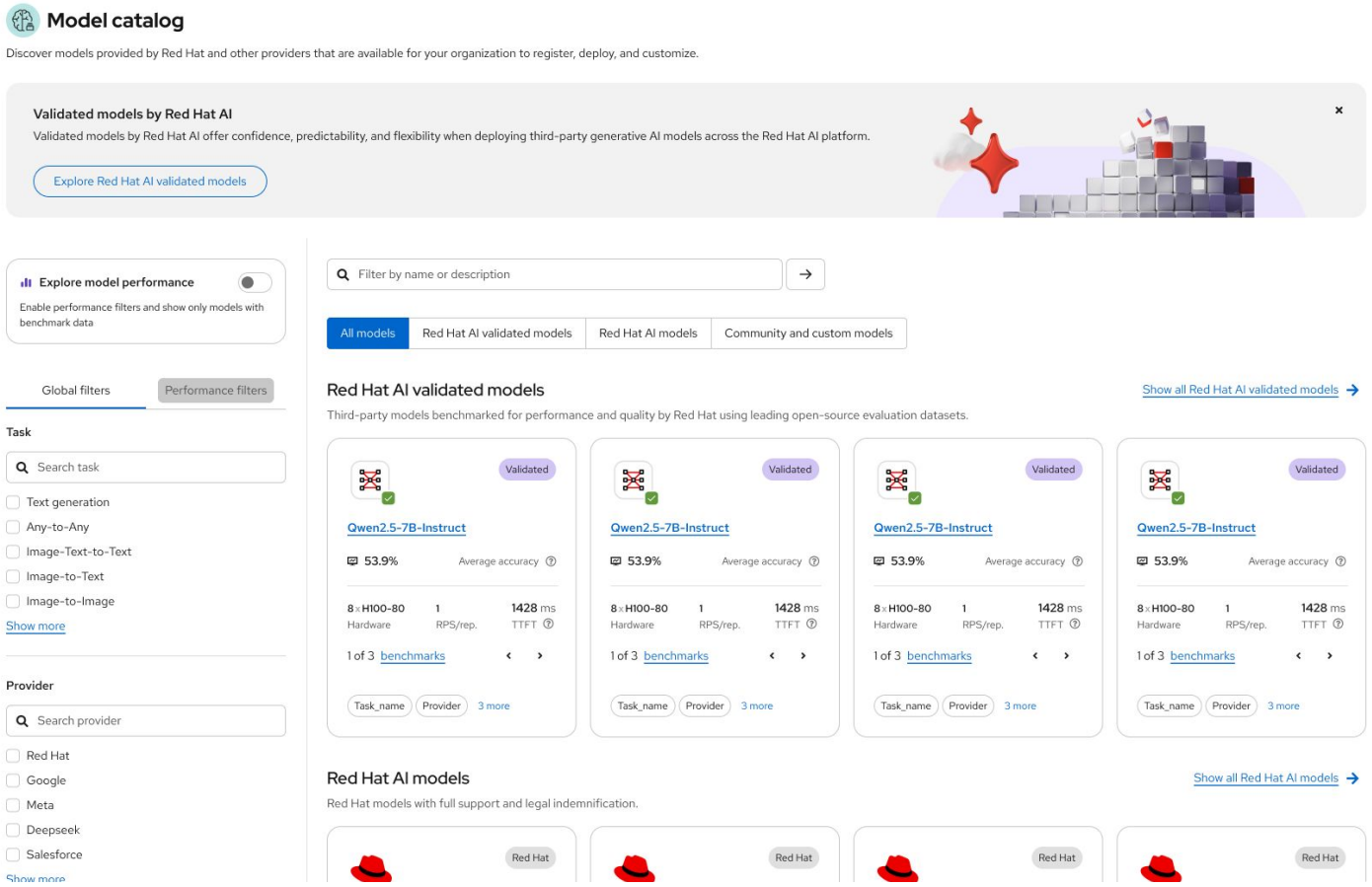
    Servers, agents, prompts
- Components:
  - Catalog
  - Registry
  - Deployments



Catalog Mockup

Catalog (Models)

OpenShift AI users now have the ability to discover Red Hat models, Red Hat AI validated models, and popular community and custom 3rd party models to be easily compressed, tuned, and deployed from within OpenShift AI.



Agentic AI: The Four Pillars

A unified framework for building, connecting, managing, and using AI agents

The Engine

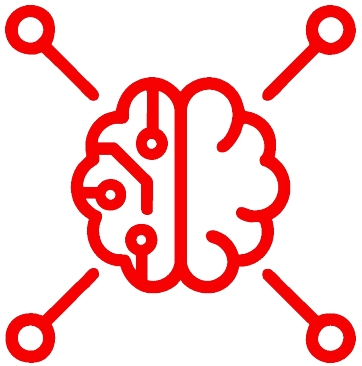
Backbone for Agentic AI

- Secure, scalable API server(Llama Stack)
- Model inference, memory, RAG, compliance

The Platform (AgentOps)

Lifecycle management for agents

- DevOps → AgentOps practices
- Governance, monitoring, versioning



Agentic AI

The Connectors

Interoperability across ecosystems

- Open protocols (MCP, A2A, etc.)
- Catalogs, registries, partner integrations

The Experience

Seamless developer on-ramps

- GenAI UI (playgrounds, prompt mgmt)
- Agent starter kit + LangChain/LlamaIndex

Introducing Model Context Protocol (MCP)

Give AI Engineers the power to connect models to real-world tools – right now

What's Coming in RHOAI v3.0

- **MCP Creation Guide** – Clear steps to design & configure MCP Servers
- **Try Before You Buy** – Evaluate MCP Servers with loaded models directly in the Playground
- **Available AI Assets Listing** – Quickly see and launch what's ready to use
- **Llama Stack Tool Calling** – Seamless tool invocation from your models

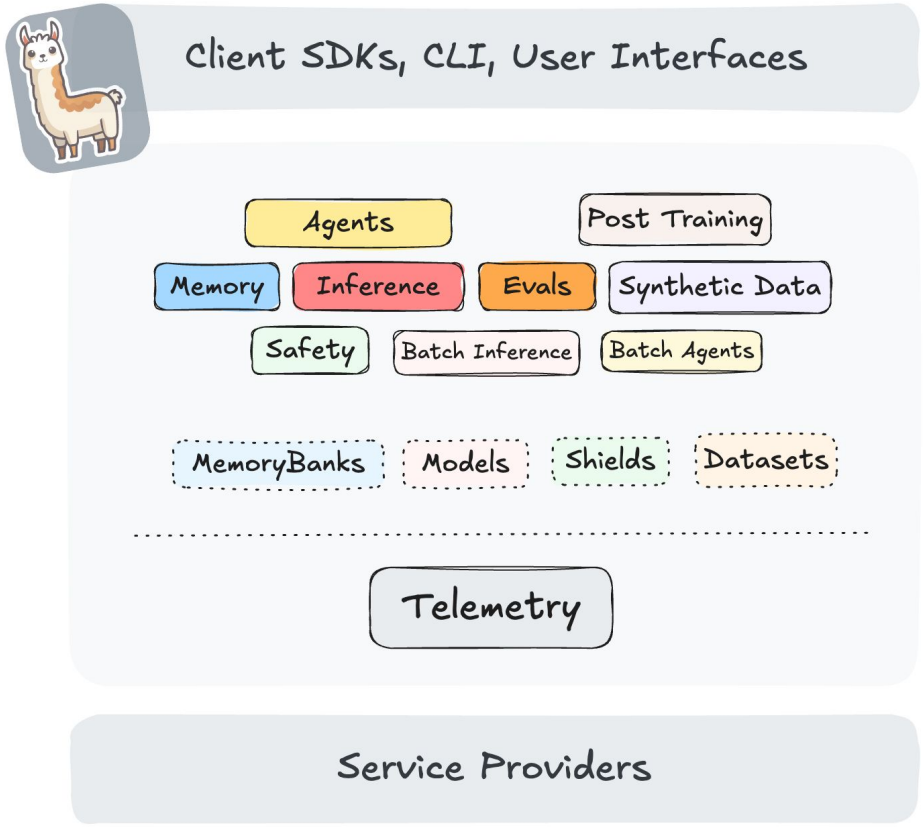
Why it Matters

- **Instant Experimentation** – Start building today
- **Model + Tool Synergy** – Bring real-world capabilities to your AI workflows
- **Foundation for RHOAI 3.next** – Governed today, certified assets tomorrow

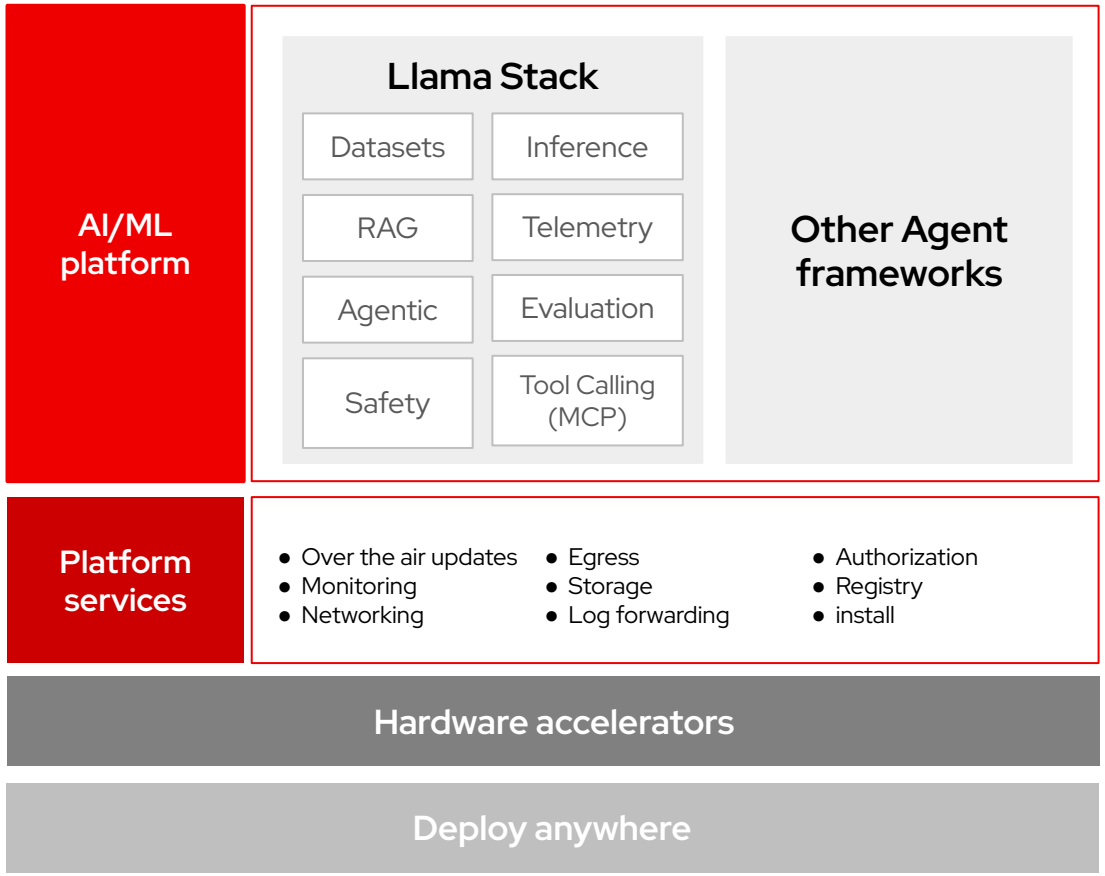
Red Hat

Unified Agentic AI Framework

- **Orchestration framework** for building AI applications.
- **Unified** set of **APIs** enabling seamless transitions between development and production environments.
- **Batteries Included:** Inference, VectorDB, Datasets, Embeddings, Agents, Safety Shields, Evaluation, etc...



A modular approach to building AI agents



Red Hat AI allows to:

- Build agents using **Llama Stack's native capabilities and implementations**.
- **Bring compatible Llama Stack implementations** to OpenShift AI.
- **Use your own agent framework** and selectively incorporate Llama Stack APIs.
- **Build with Core Primitives** and manage your own agent framework as a standard workloads.

Red Hat

Let’s see something

AI-infused application |ˌeɪˌaɪ ˈɪnˌfjuːzd ˌæplɪˈkeɪʃən|

noun (Plural AI-Infused applications)

A software program enhanced with artificial intelligence capabilities, utilizing AI models to implement *intelligent* features and functionalities.

How LLMs work at inference time?

Large Language Models are **neural networks** that identify, **predict**, and generate text. They are trained on vast amount of text to deduce statistical relationships between tokens.

LLMs can be fine-tuned and predict the next token based on their training data and **statistical** deduction

Model and Model Serving

Model Serving

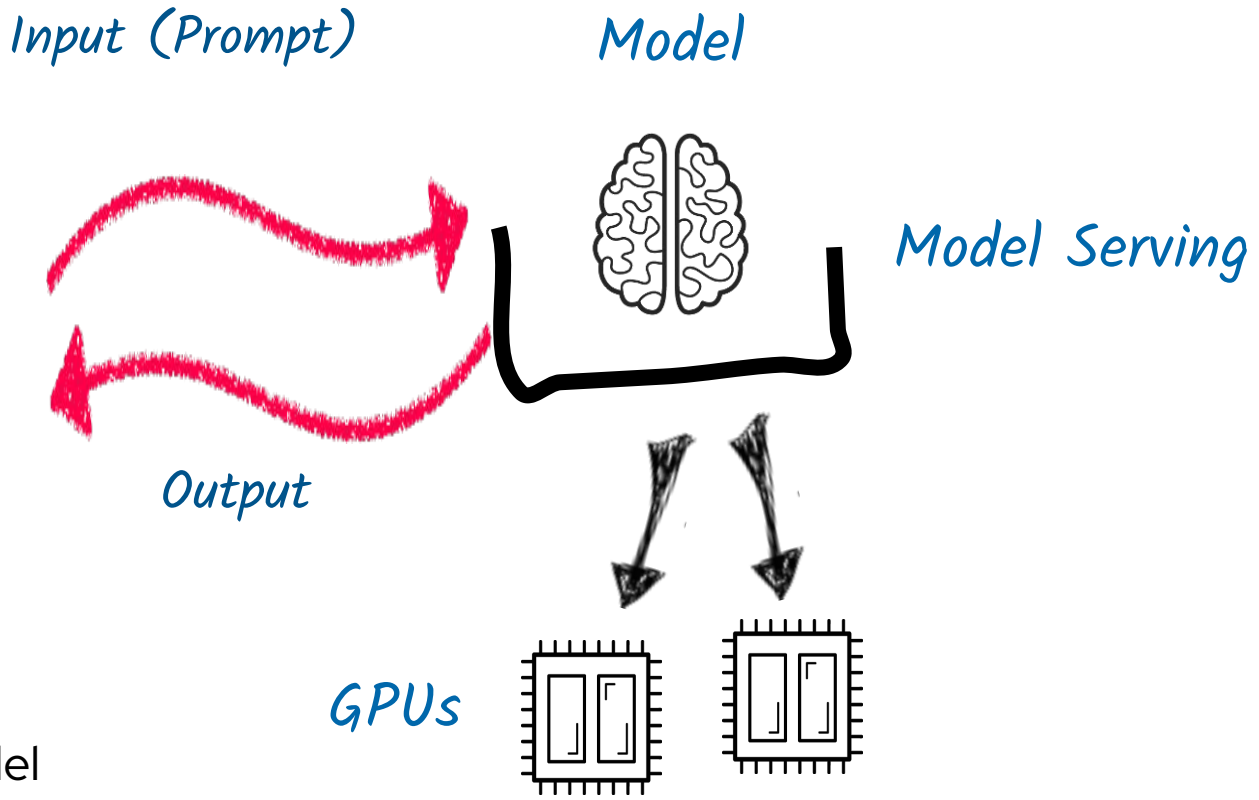
- Run the model
- CPU / GPU
- Expose an API

Input

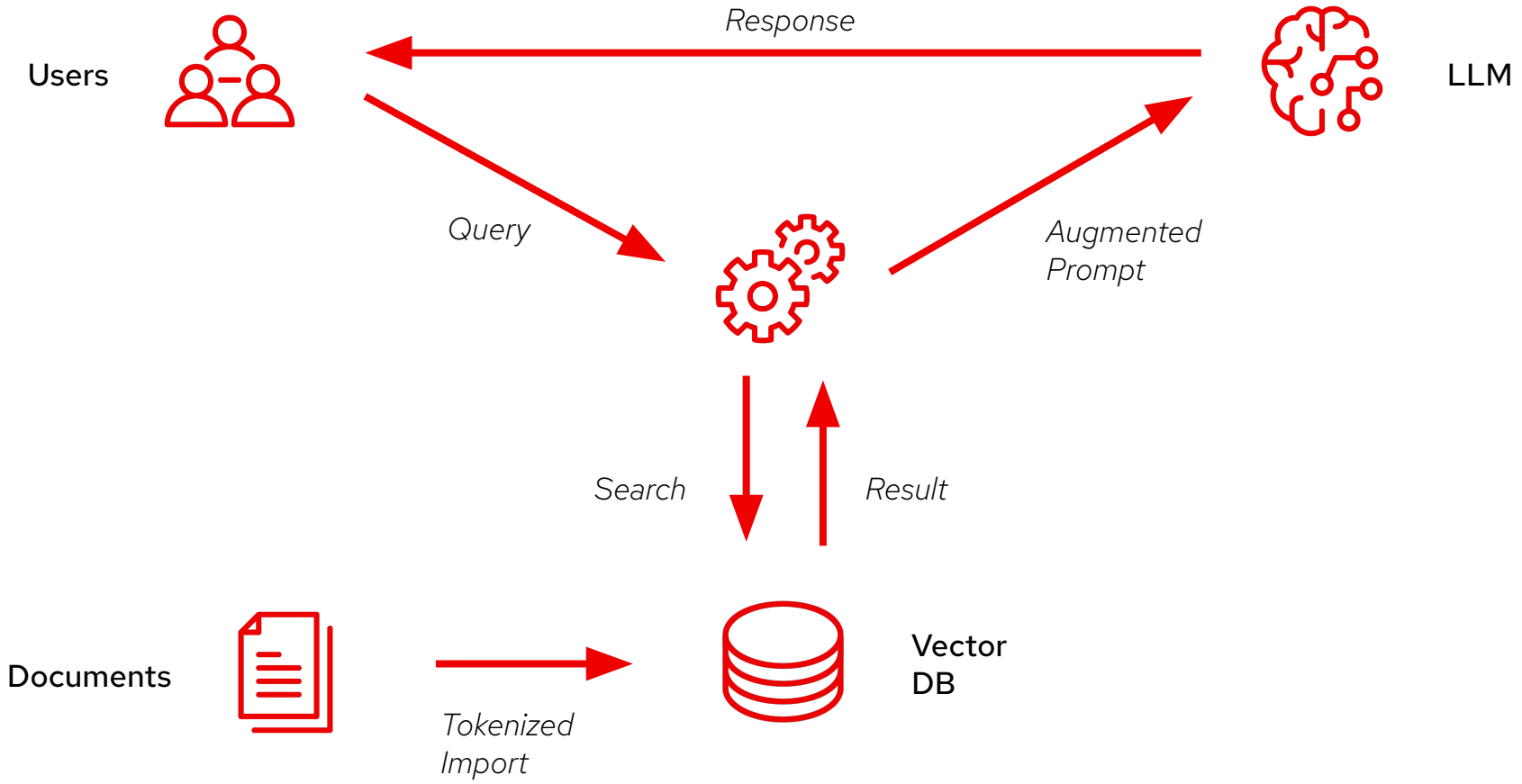
- Prompt (Text/Image/generic)
- Instructions to give to the model

Output:

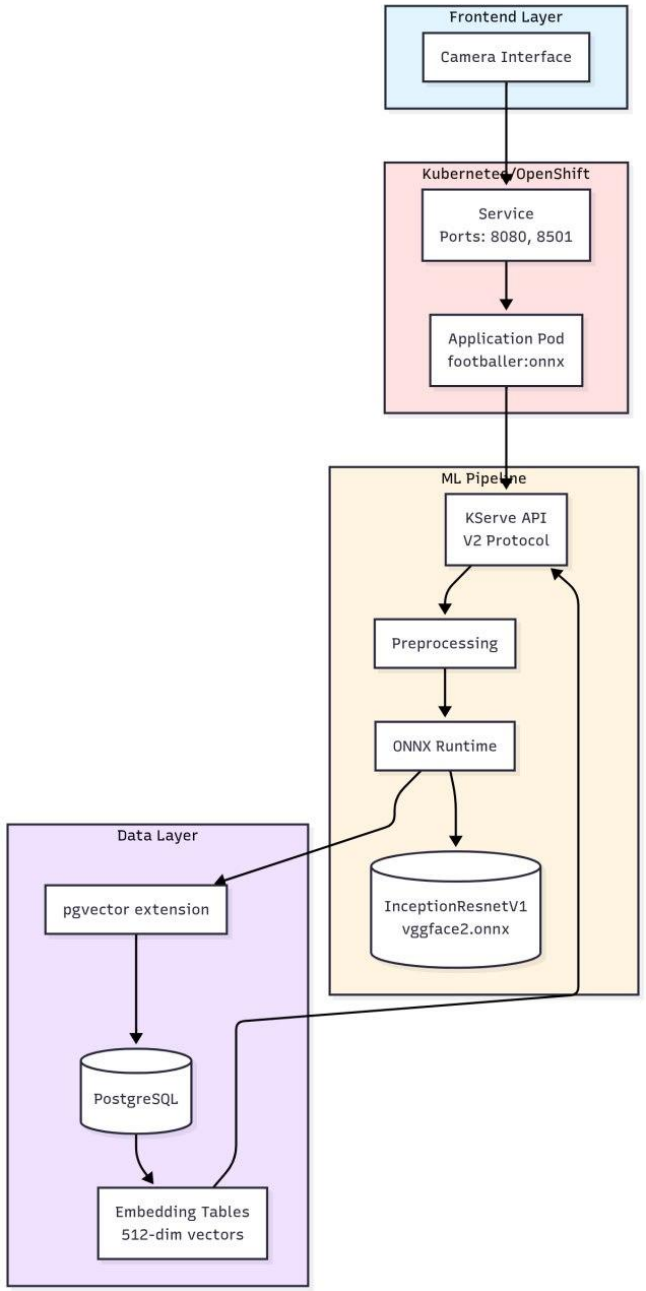
- Depends on the modality of the model



RAG (Retrieval augmented generation) provides extra info



Red Hat



Which Footballer Are You?

- Frontend Microservice -> Streamlit with Camera Access (python)
- Kubernetes-based network plumbing (Services and Ingresses)
- AI Services:
  - Model running in OCP via Service Runtimes (ONNX Computer Vision Model)
  - API & Preprocessing via KServe
- Image Embeddings (“ImageRAG”) -> stored into PGVector

EDB Postgres AI

Red Hat OpenShift

Red Hat AI

Matteo Combi

Specialist Solutions Architect
Red Hat

Francesco Rossi

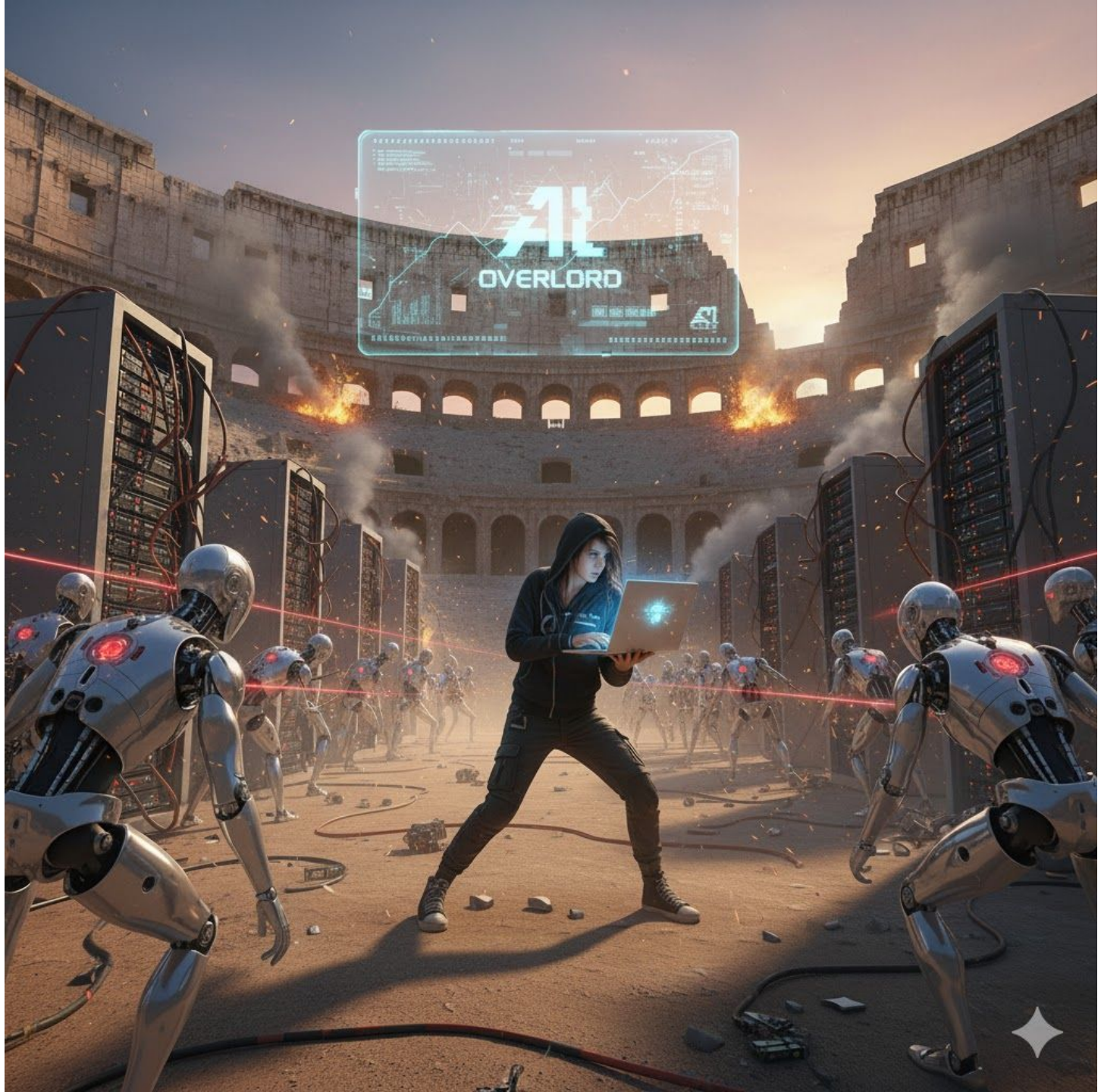
Specialist Solutions Architect
Red Hat

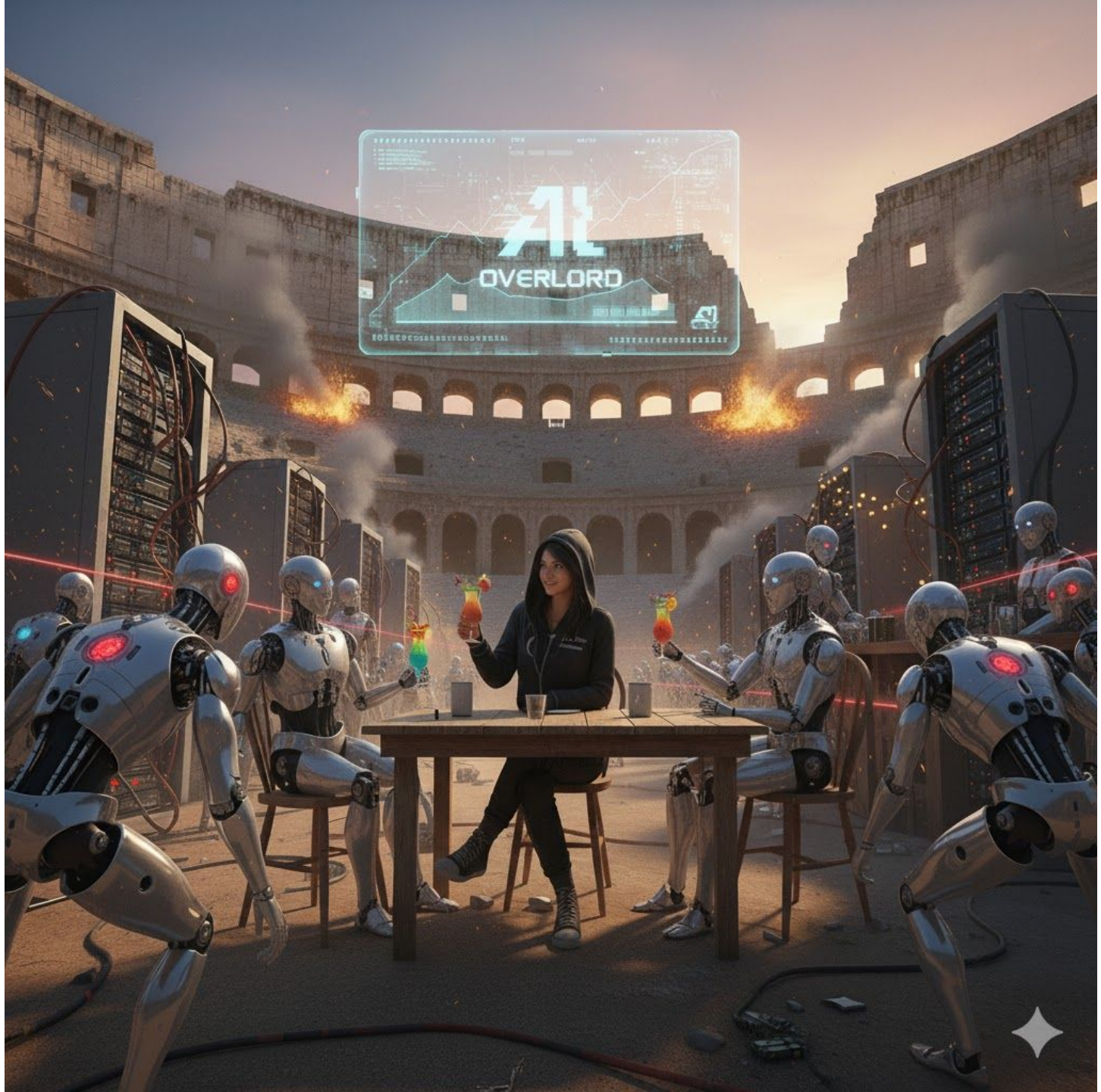
Matteo Combi

Specialist Solutions Architect
Red Hat

Francesco Rossi

Specialist Solutions Architect
Red Hat

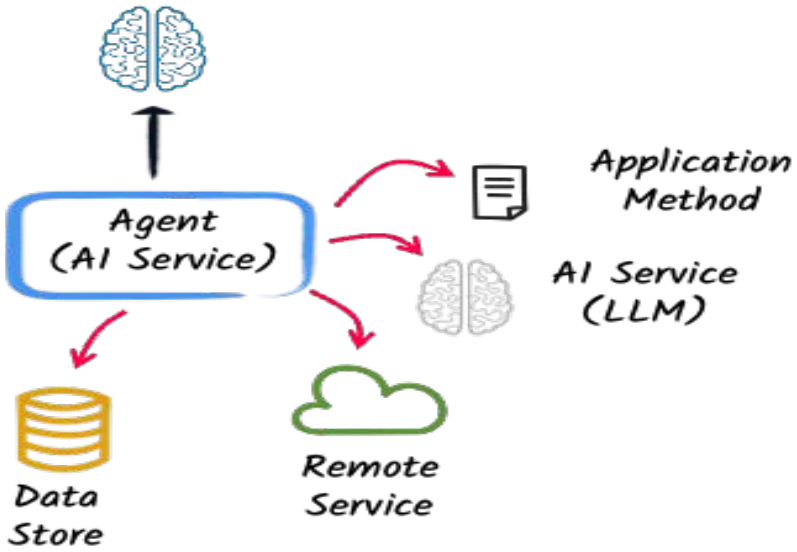

AI
OVERLORD


AI
OVERLORD

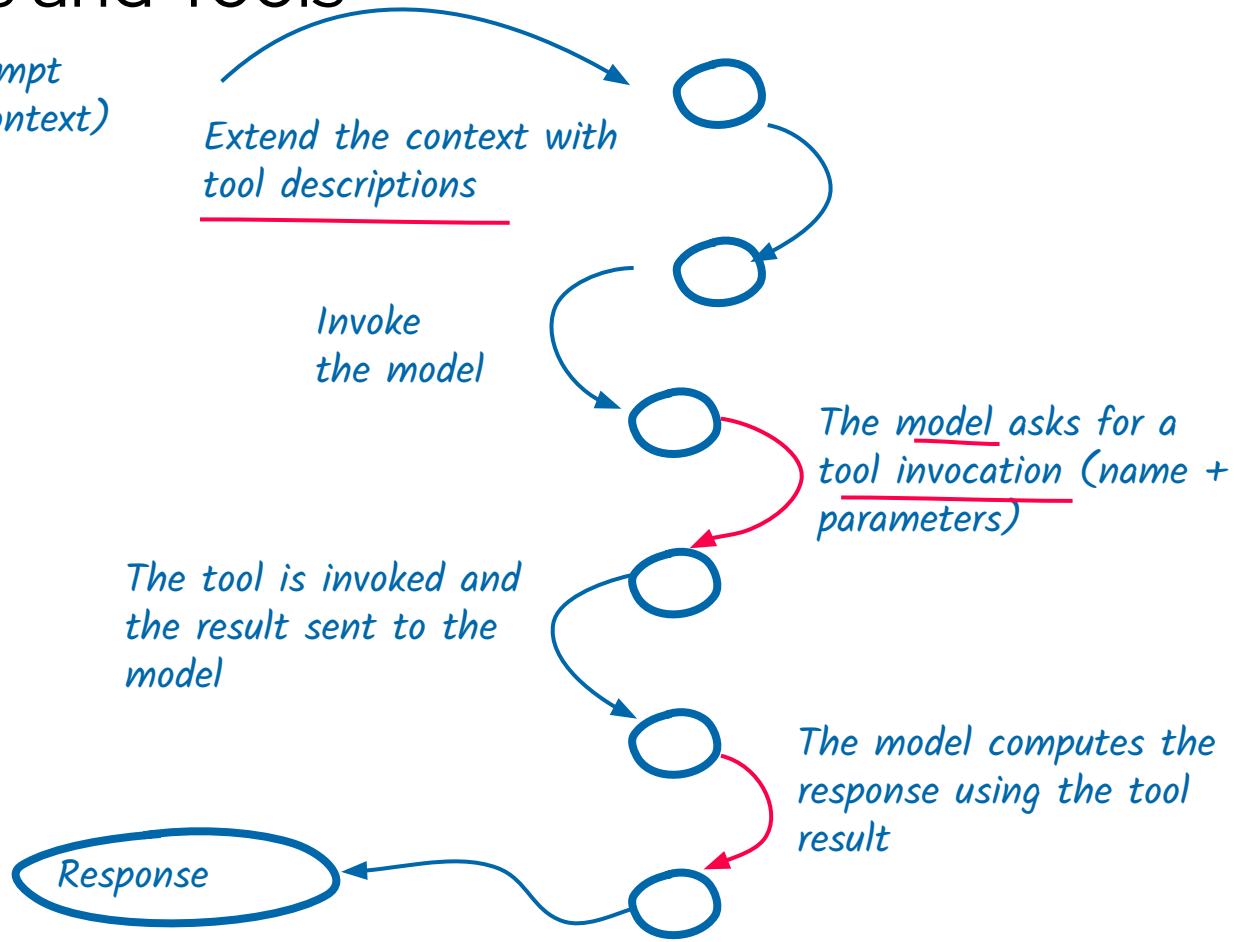
What are Ai Agents?

An **AI agent** is an **autonomous** computational system that perceives its environment, **processes** information, and **takes** actions to achieve specific goals or tasks, often exhibiting traits like **learning**, **reasoning**, and planning.

Agents and Tools



Tools require memory and a reasoning model

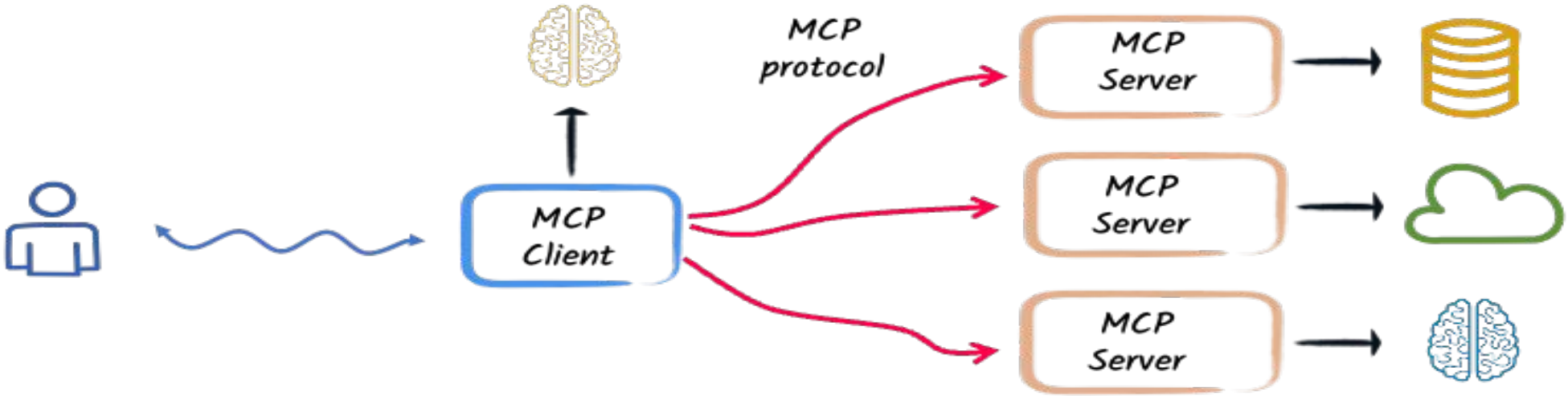


What is A2A?

An Agent-to-Agent (A2A) protocol is a communication standard that allows different Artificial Intelligence (AI) agents to securely interact, collaborate, and coordinate tasks with each other, regardless of the vendor, underlying model, or framework they were built on

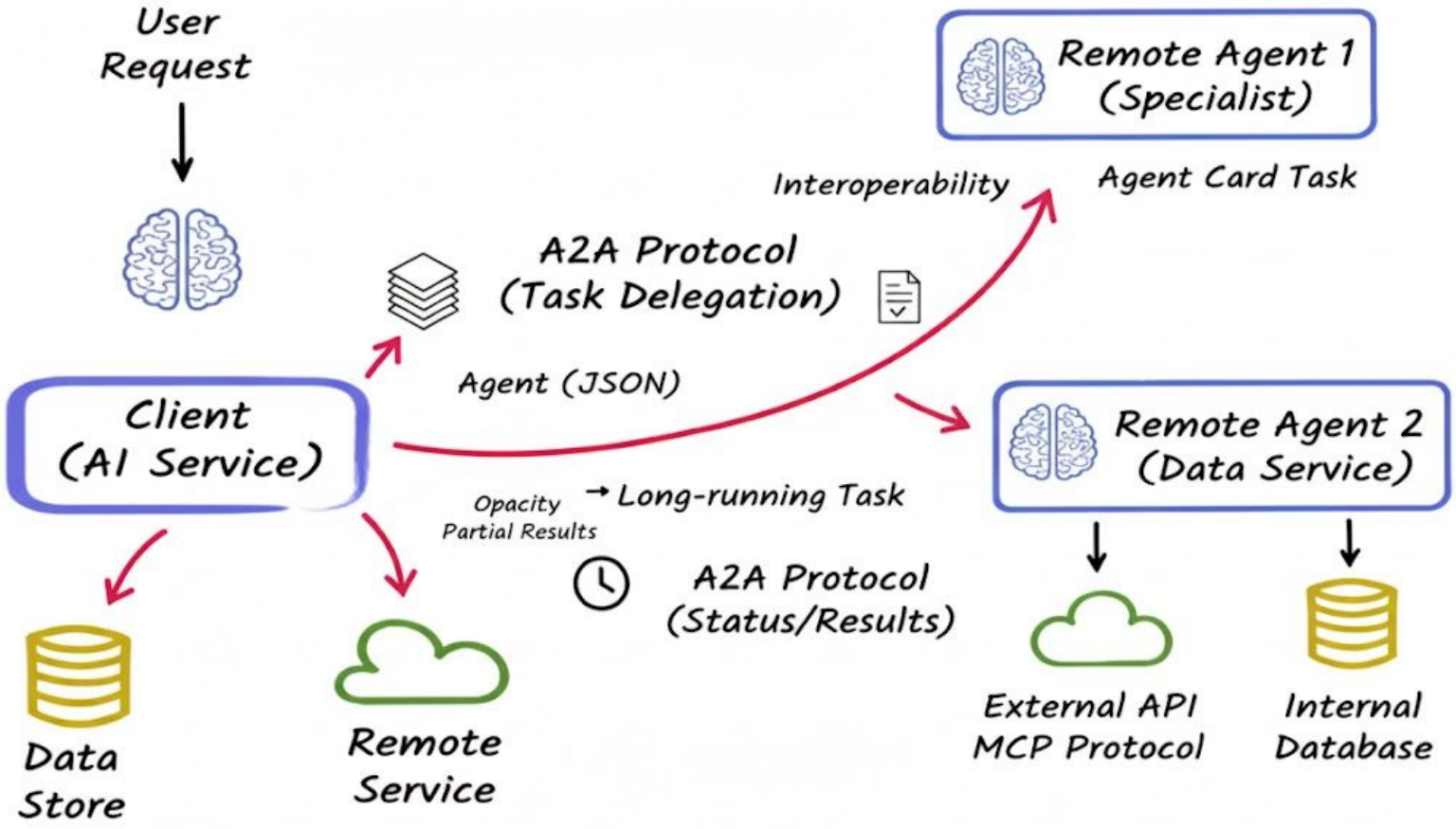
Model Context Protocol (MCP)

The Model Context Protocol,introduced by Anthropic in November 2024, is an **open standard** that enables developers to build secure, two-way connections between their data sources and AI-powered tools.

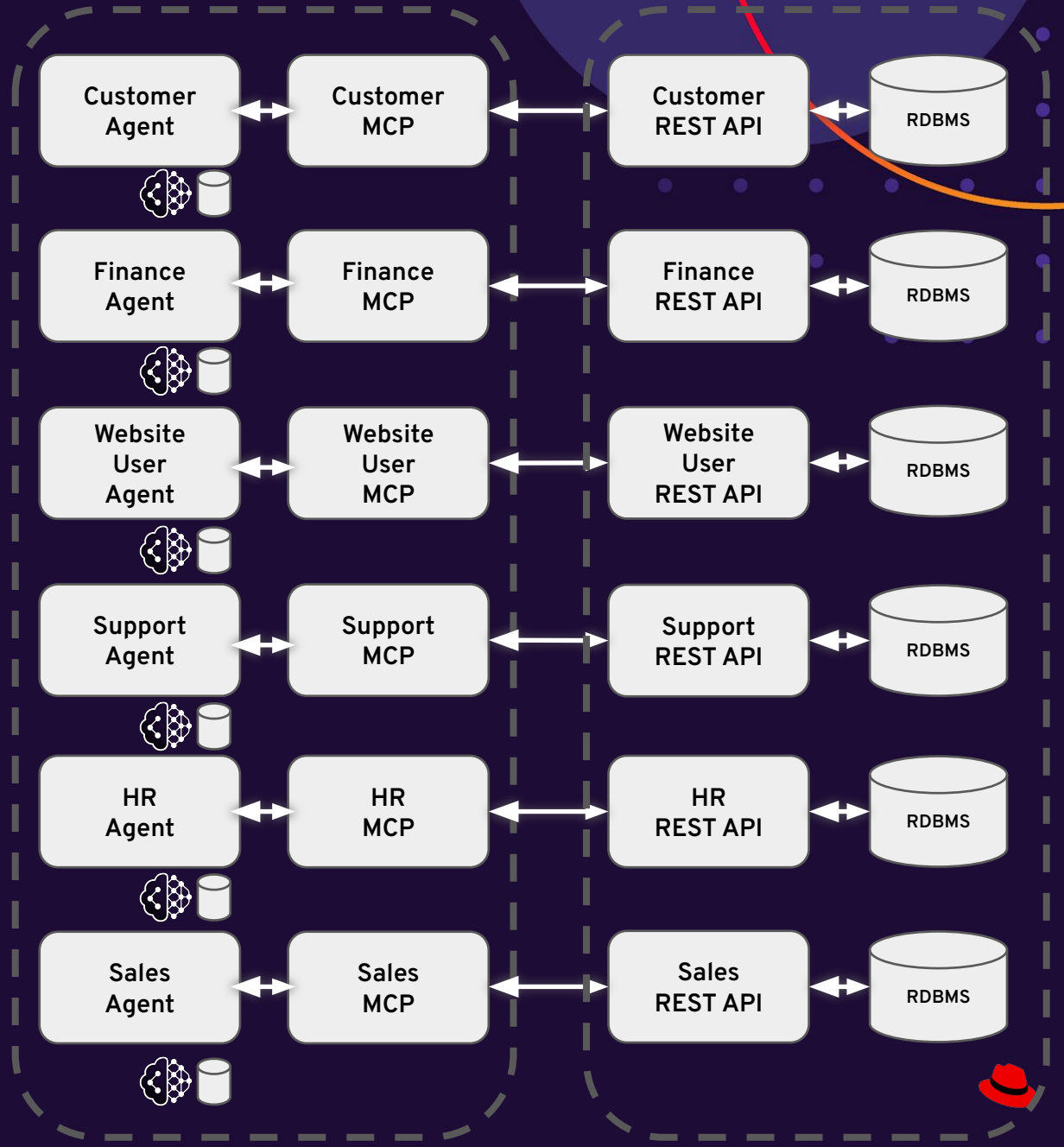


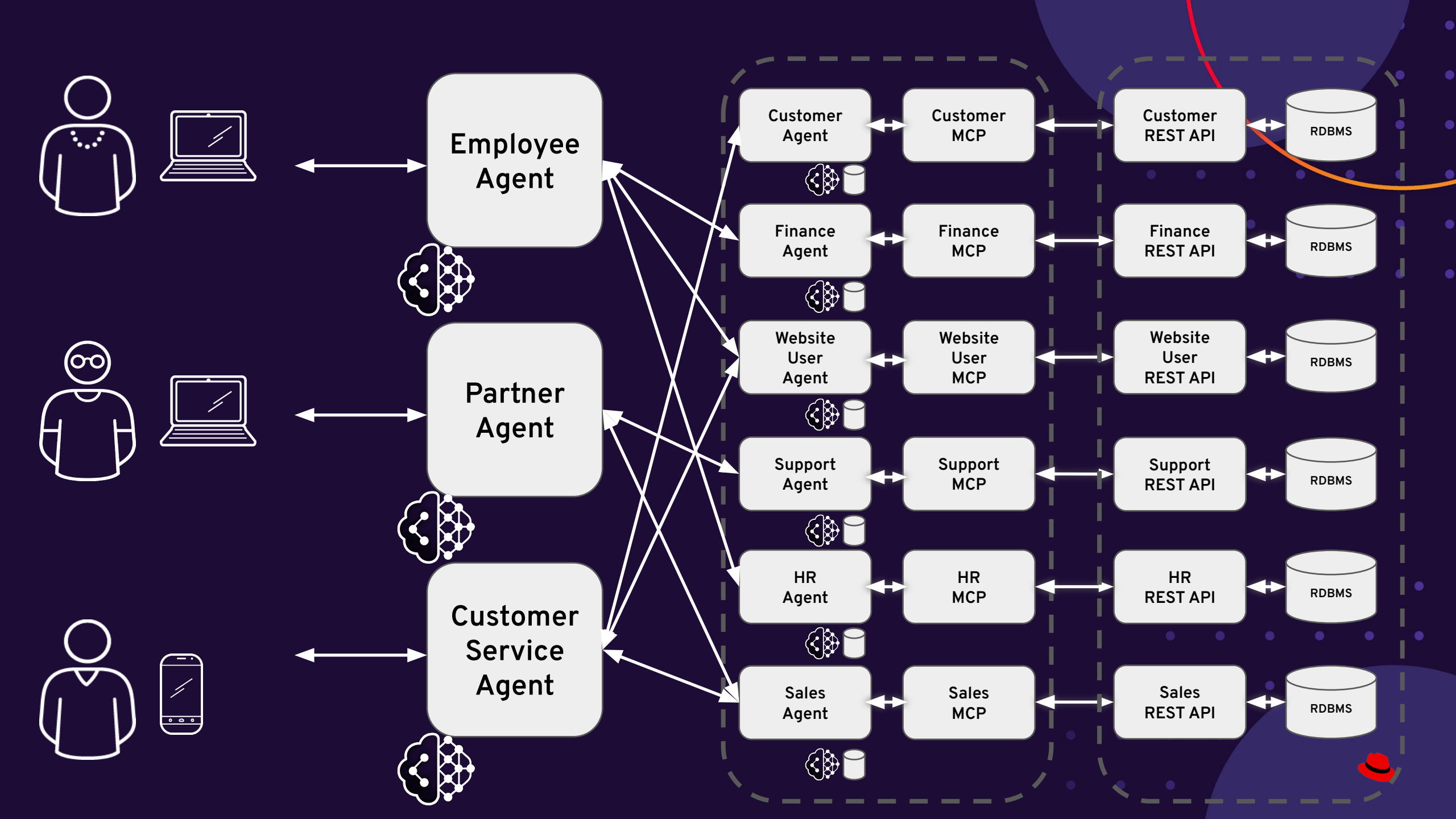
Instead of exposing tools from your code,
discover and use remote services

Agent-to-Agent (A2A) Protocol: Collaborative Workflow



Agents can coordinate other agents

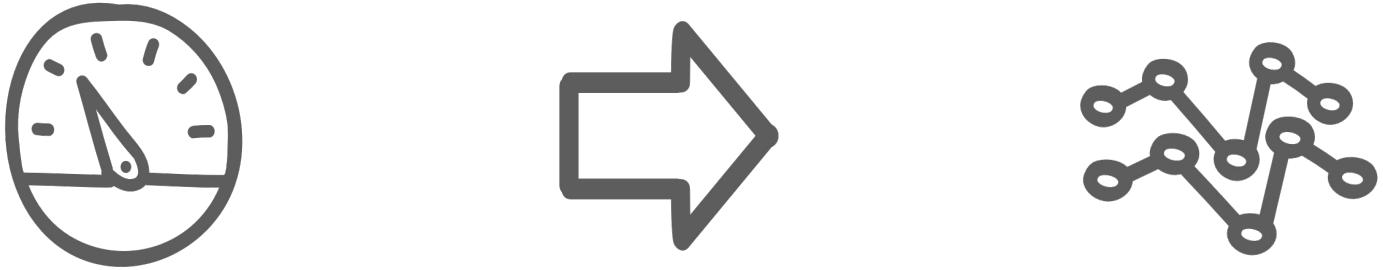



Employee Agent
Partner Agent
Customer Service Agent
Customer Agent
Finance Agent
Website User Agent
Support Agent
HR Agent
Sales Agent
Customer MCP
Finance MCP
Website User MCP
Support MCP
HR MCP
Sales MCP
Customer REST API
Finance REST API
Website User REST API
Support REST API
HR REST API
Sales REST API
RDBMS

EVERY DAY LIFE
- ROUTINE TASKS
INTELLIGENT AGENTS
- - AUTONNOOUS AI

Supercharge your Java application with Agentic AI

Supersonic

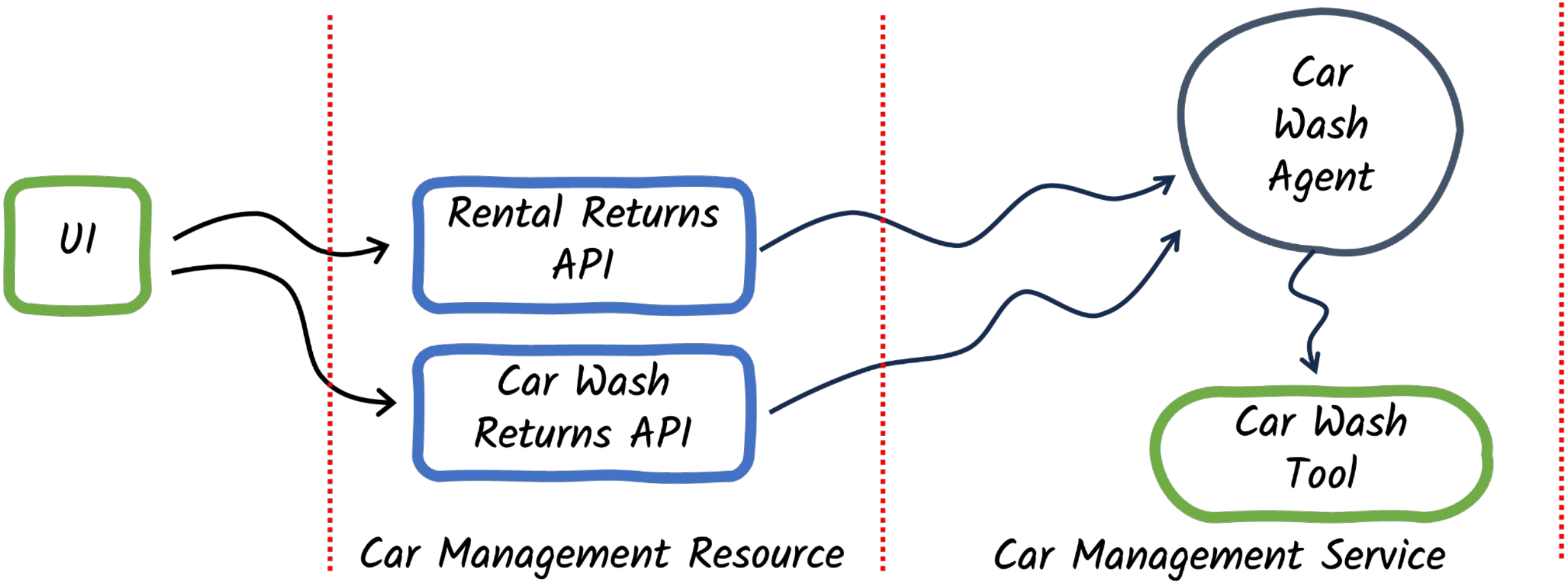


Fast. Blazing fast to start. Millisecond fast!

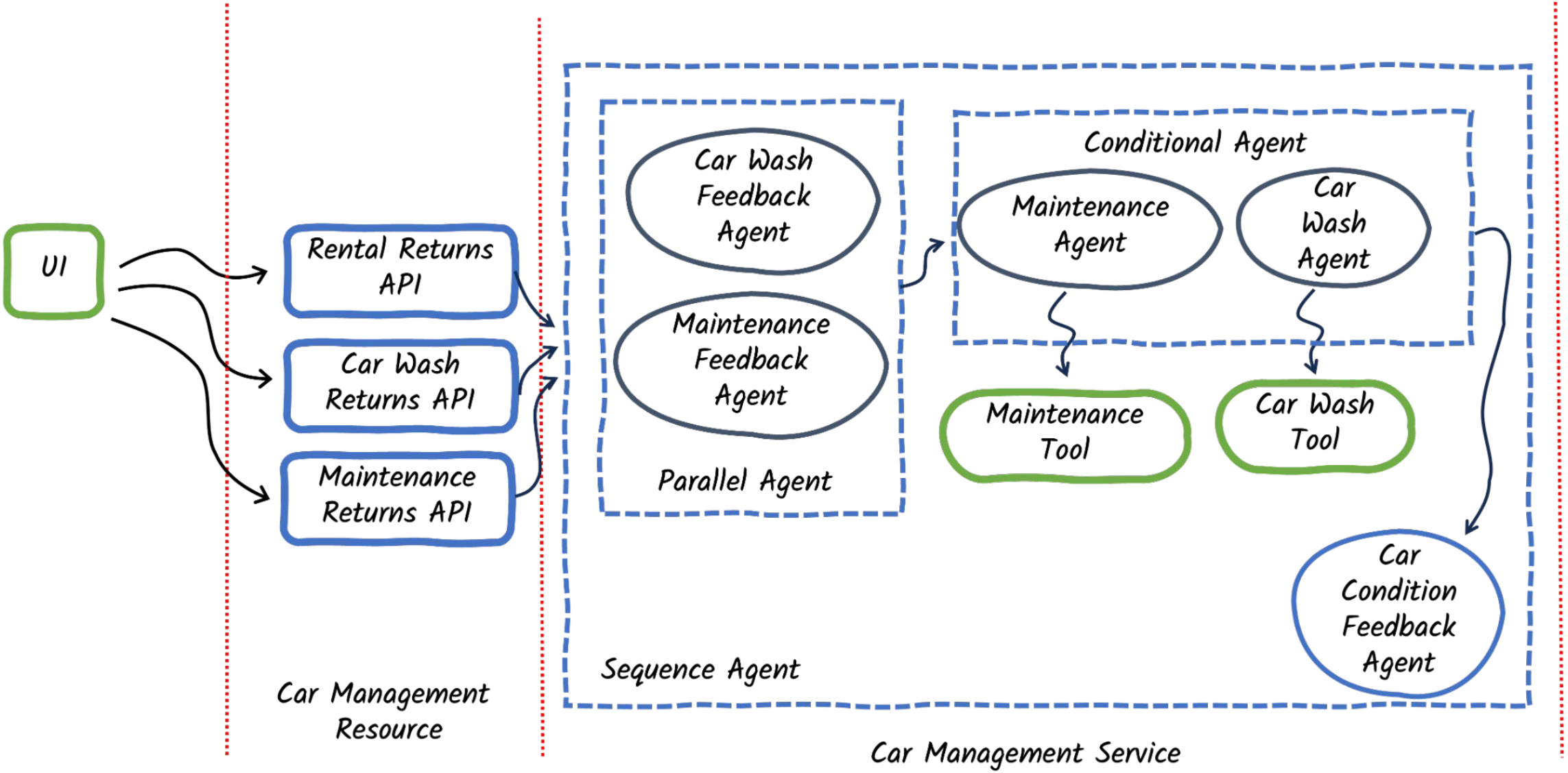
Subatomic

Improve memory consumption. Increase deployment density.

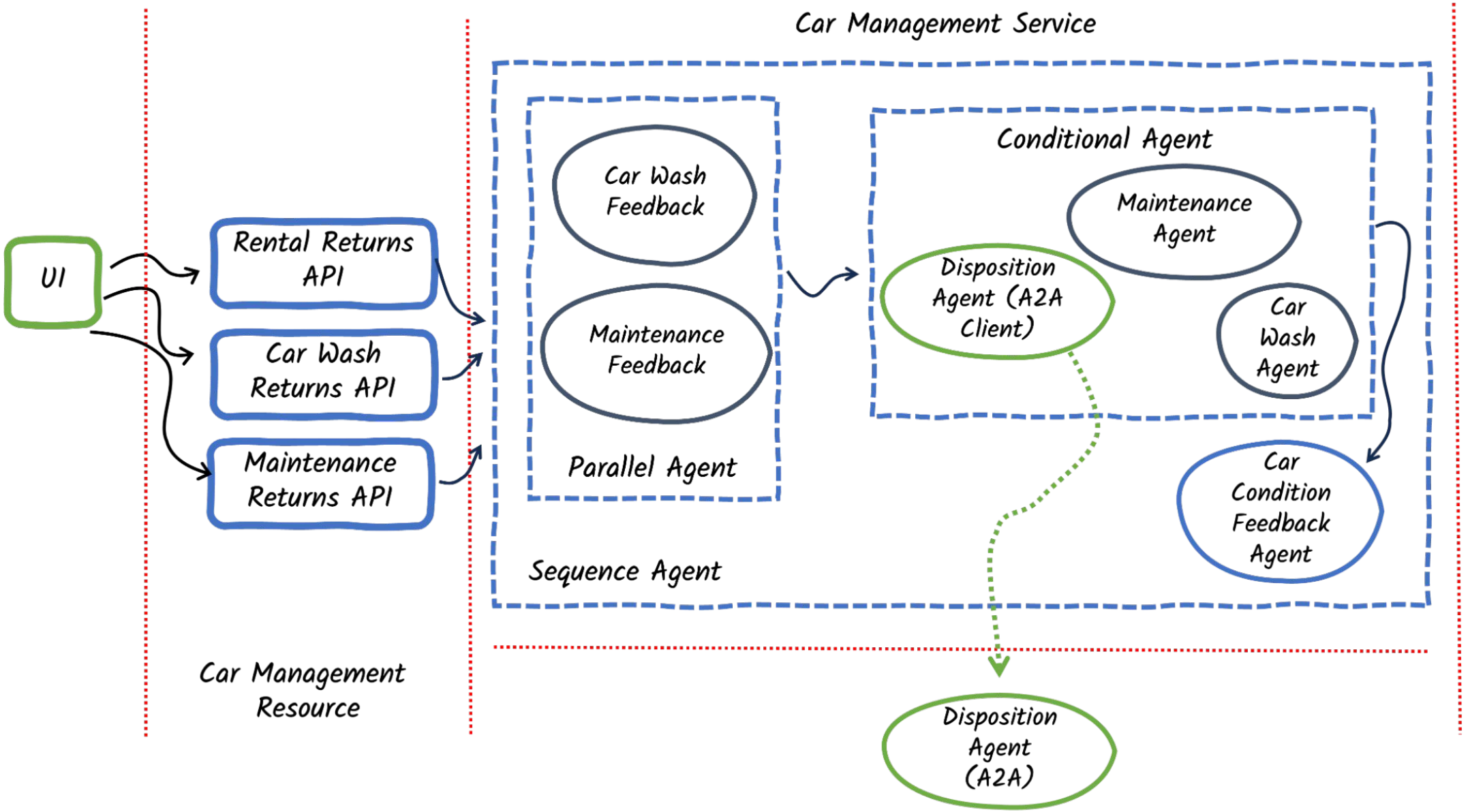
Demo Architecture - Simple Agent



Demo Architecture - Nested Agent Workflows



Demo Architecture - Step 4 - Using Remote Agents (A2A)



Repeat after me....
Before I build an AI agent
I'll check
if I even need one

Red Hat

Red Hat Developer

Join Red Hat Developer.

Build here. Go anywhere.

youtube.com/RedHatDevelopers

linkedin.com/showcase/red-hat-developer

facebook.com/RedHatDeveloper

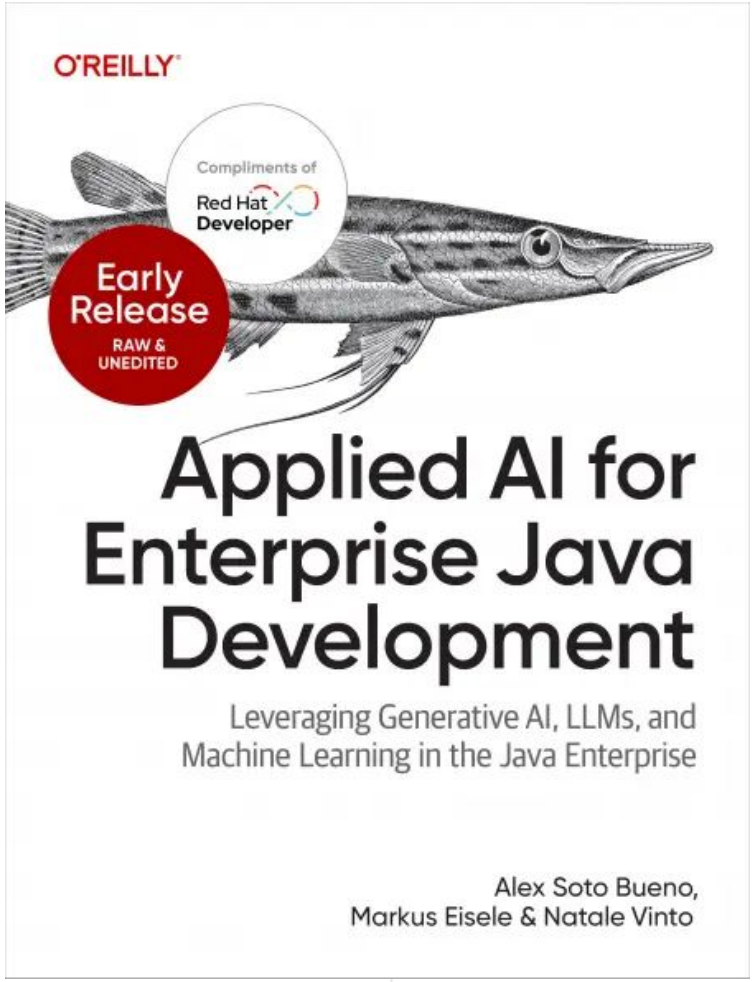
twitter.com/rhdevelopers

Get started

Red Hat Developer

Sign up at developers.redhat.com

Find out more about Red Hat's project and products, and what it offers developers



Red Hat

Red Hat Summit

Connect

Thank you

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat Summit

Connect

AI & Automazione

La nuova era del controllo IT con Red Hat Ansible Automation Platform 2.6

Marcello Panci
Sales Specialist

Giorgio Voltolina
Sales Specialist

Andrea Guarnaccia
Senior Specialist Solution Architect

Marcello Panci

Sales Specialist
Red Hat

Giorgio Voltolina

Sales Specialist
Red Hat

Andrea Guarnaccia

Senior Specialist Solution Architect
Red Hat

What’s new in Ansible Automation Platform 2.6

Red Hat Ansible Automation Platform 2.6

Automate at scale on a proven foundation

Your team isn't just preparing for the future, you're automating for it. And our latest release delivers new features, platform enhancements, and strategic integrations that will help you continue to build a resilient, trusted foundation for the next generation of IT operations.

- ✓ Unlock more value with the new **automation dashboard**
- ✓ Operate more efficiently with the new **Ansible Lightspeed intelligent assistant**
- ✓ Achieve new levels of scale with the new **self-service automation portal**

New!

Automation dashboard

Overview

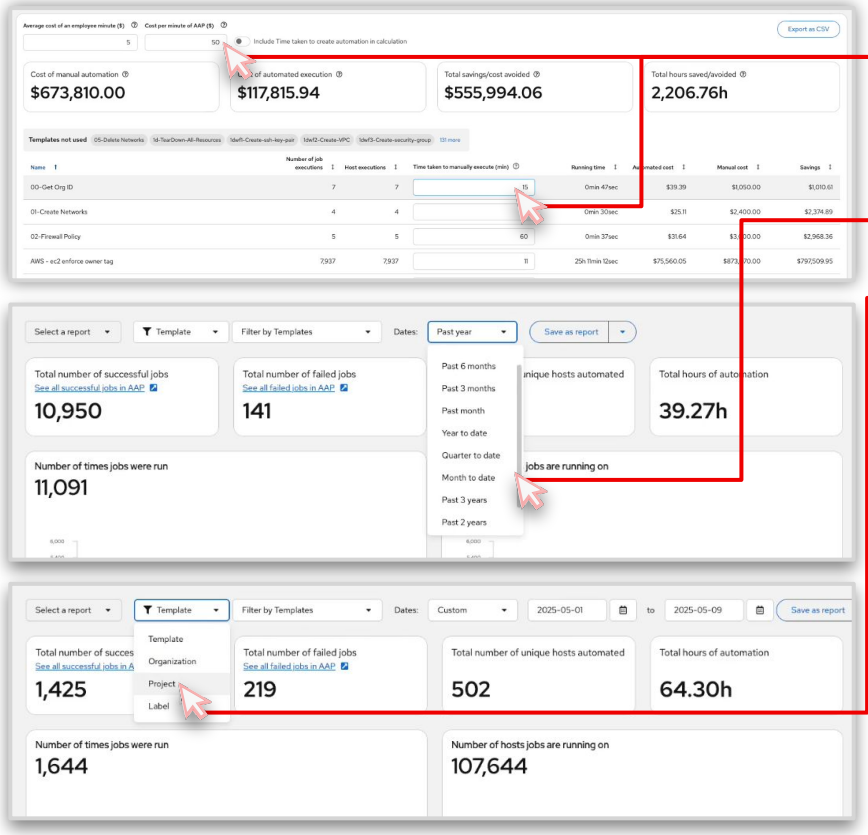
The new **automation dashboard** delivers something that our customers have been waiting for - a streamlined and secure way to monitor, track and report on the value that automation is delivering to the business.

- ✓ Demonstrate value and measure ROI
- ✓ Make smarter, data-driven decisions
- ✓ Optimize and right-size your automation
- ✓ Securely report and share your data

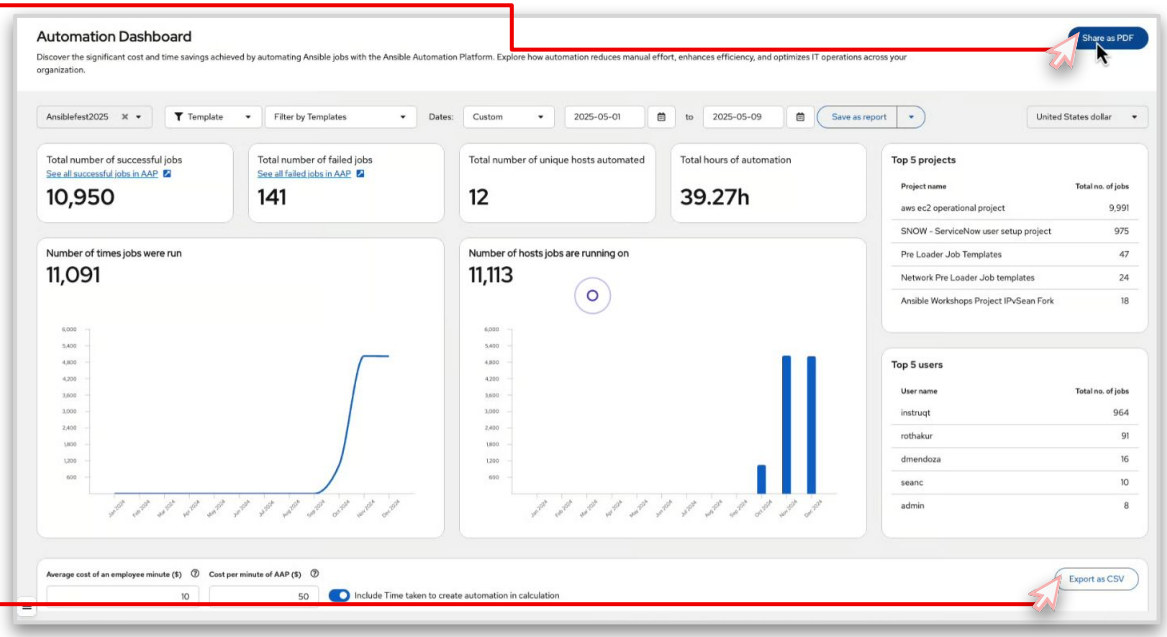
New!

Automation dashboard

Experience



- **Customize** the time taken to manually execute tasks, time taken to create the automation, avg cost of employee per minute, and cost of AAP per minute
- **Select** the desired reporting time period
- **Filter** by Template, Organization, Project, or Label
- **Easily export and share** report via .pdf or .csv



New!

Ansible Lightspeed intelligent assistant

Overview

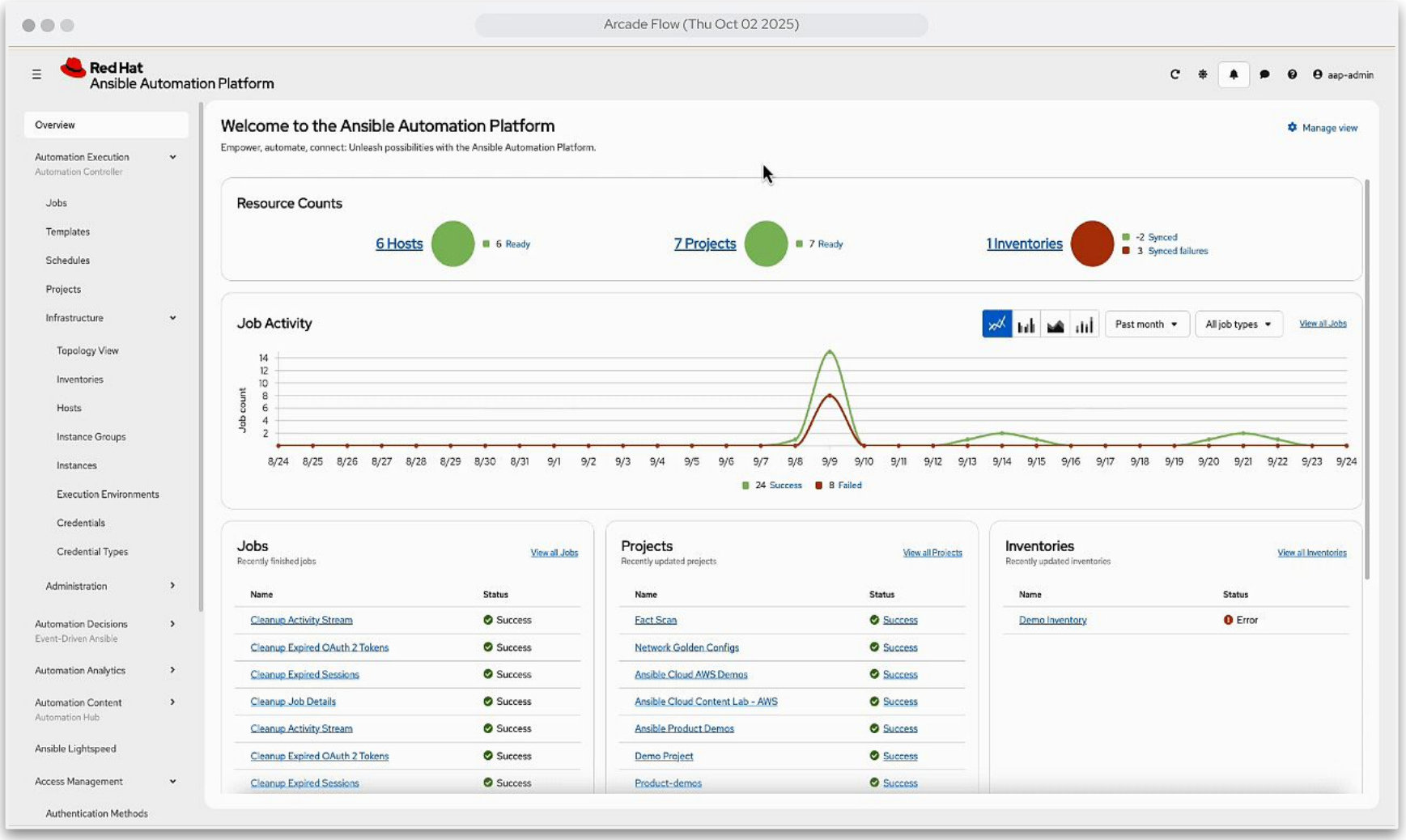
The new **Ansible Lightspeed intelligent assistant** is a chat assistant embedded directly into the Ansible Automation Platform UI. It harnesses generative AI to provide platform admins and users on-the-fly support for day-to-day automation platform management.

- ✓ Provides on-demand platform expertise and insights from Ansible documentation
- ✓ Speeds up troubleshooting, onboarding, and day-to-day management.
- ✓ Includes links to help with source validation and accelerated learning.

New!

Ansible Lightspeed intelligent assistant

Experience



Example prompts:

"What is an execution environment?"

"How do I upgrade to the latest version of Ansible Automation Platform?"

"How do I create a survey within a job template?"

New!

Self-service automation portal

Overview

The new **self-service automation portal** enables IT ops teams to take a platform engineering approach to automation service delivery, helping customers automate at new levels of scale across the enterprise.

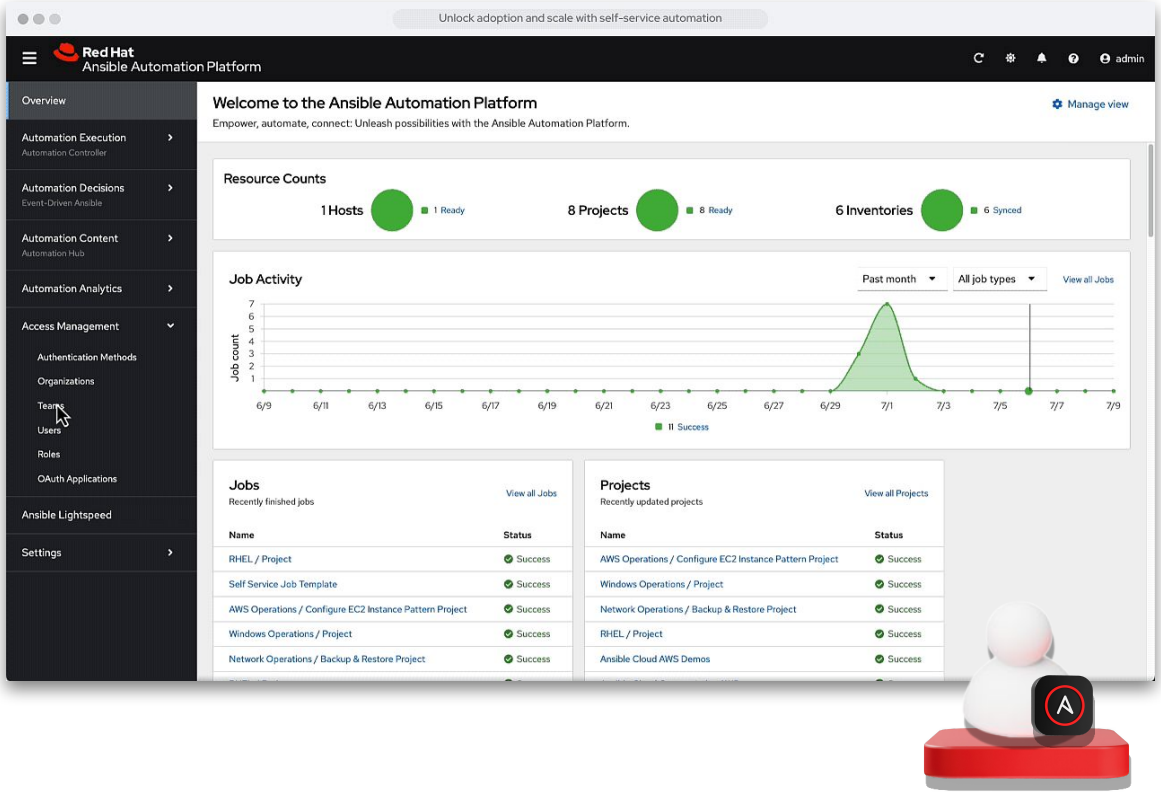
- ✓ Unlock wider automation adoption across your organization.
- ✓ Scale automation without the overhead of building custom solutions.
- ✓ Empower admins to share automation while maintaining full control and consistency.
- ✓ Simplify automation service delivery for non-Ansible experts.

New!

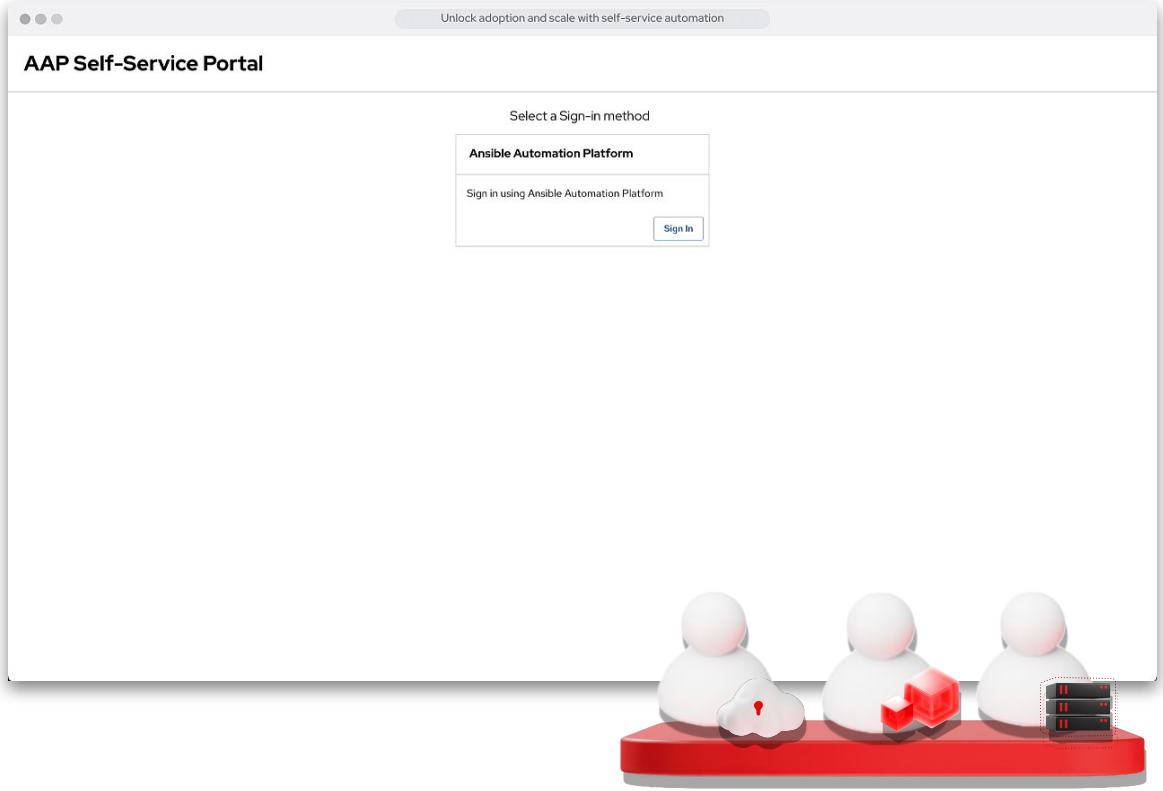
Self-service automation portal

Experience

Platform admin
Defines automation use cases and sets access permissions



Domain subject matter expert
Views and runs automation relevant to their work with an easy-to-use self-service experience.



Additional platform experience enhancements

An upgrade to AAP 2.6 will include 2 releases worth of new features

New!

Event-Driven Ansible (EDA) enhancements
Gain greater flexibility and control over your event-driven automation:

- Support for external secret management systems
- Editable project source control URLs
- EDA job labeling for more efficient tracing + auditability
- Enhanced Kafka support for multiple topics and wildcards.
- Rulebook enhancements for Kubernetes

Refreshed User Interface
Navigate a cleaner, more modern, and more responsive UI that adapts to different devices and includes enhanced accessibility features.

AI-Assisted Inventory Generation
Simplify your installation and onboarding process with this new *developer preview* feature. Describe your desired topology, and the platform will automatically generate a validated inventory file for you, reducing manual effort and potential errors.

ICYMI

AAP 2.5 upgrades you'll also get with 2.6

- Event-Driven Ansible (EDA) improvements
  - Event streams for simplified routing
  - Horizontal scalability
  - Integrated RBAC, SSO, authentication + credential management
- Unified UI
- Containerized installation
- Updated Ansible development tools
- Ansible plugins for Red Hat Developer Hub

New ecosystem integrations and certified collections

Networking

Inventory and Health Check Reports for Arista and Cisco: (*Arista EOS, Cisco IOSXE, Cisco IOSXR, Cisco NXOS*)

Cisco Meraki Ops: Validated content to automate provisioning and Day-2 operations of Meraki devices.

Nautobot: Single source of truth for network configuration and management

Cloud

HashiCorp: Support for initiating Terraform Enterprise and HCP Terraform workflows

AWS: Added SSM connection support, enhanced error handling for Amazon S3, and stricter dependency policies

Azure: Added Azure Arc for hybrid cloud support for Azure and Key Vault support for more secure automation

Google Cloud: Improved reliability and matched functionality with other hyperscaler collections

Infrastructure

Windows Server: Upgraded Microsoft Windows Server management capabilities

OpenShift Virtualization: Extending Use Cases for Day 2 operations and Pre and Post migration Hooks

VMware: Support for Provisioning VMs, advanced setting in vSphere, and Indirect Node Query Counts

AIOps

Red Hat AI: Support for Day-1 infrastructure provisioning (on-prem and cloud) and Day-2 infrastructure config and operations.

Splunk: Ability to trigger automated responses using Event-Driven Ansible rulebooks, streamlining AIOps workflows from detection to resolution

Learn more: Red Hat Blog, "What's new with Ansible automation content," 8/7/2025

Planning your upgrade

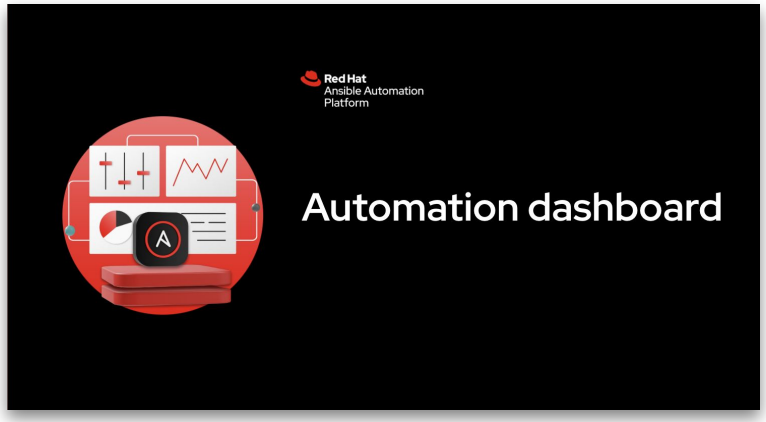
Upgrade with confidence and future-proof your investment

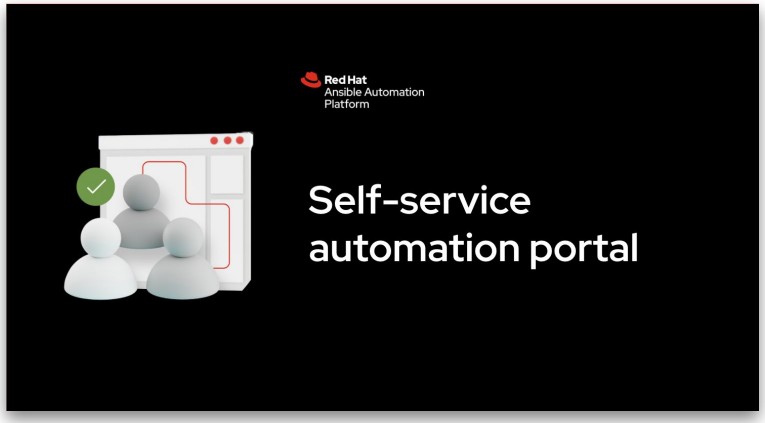
	AAP 2.4 > AAP 2.6*	AAP 2.5 > AAP 2.6
Direct upgrade path available A single, streamlined process to upgrade to AAP 2.6, reducing intermediate steps and complexity.	✓	✓
Supported topology migration A documented, "lift-and-shift" process to move your entire AAP instance—including jobs, data, and configurations—from an RPM install to containerized or OpenShift.	✓	✓
Streamlined Enterprise Auth Migration Enterprise authenticator configurations (SAML, LDAP, OIDC) are seamlessly migrated to the platform gateway, simplifying post-upgrade setup and re-authentication.	✓	AAP 2.5 customers already have the platform gateway installed, and do not need to plan for any transitions or data movement.
Automated IAM Migration Users, teams, roles, and permissions are automatically migrated from the controller to the platform gateway, preserving RBAC model and administrative privileges.	✓	

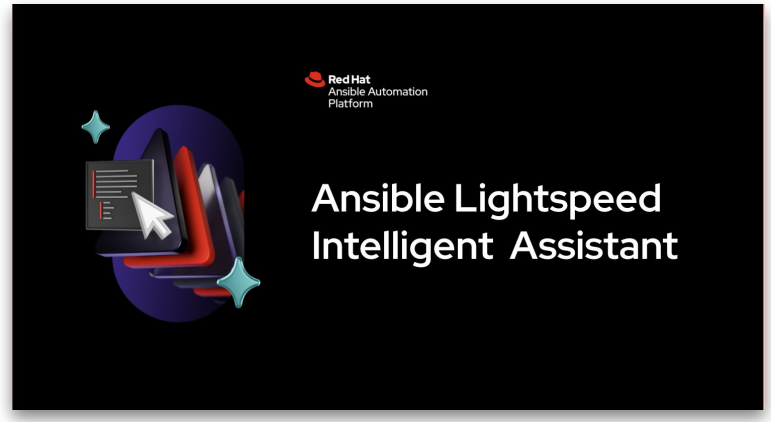
**Customers on RHEL8 will need to migrate to RHEL9 in order to have the direct 2.4 > 2.6 upgrade path*

New resource available: Interactive experiences

Step-by-step walkthroughs and demos of new features







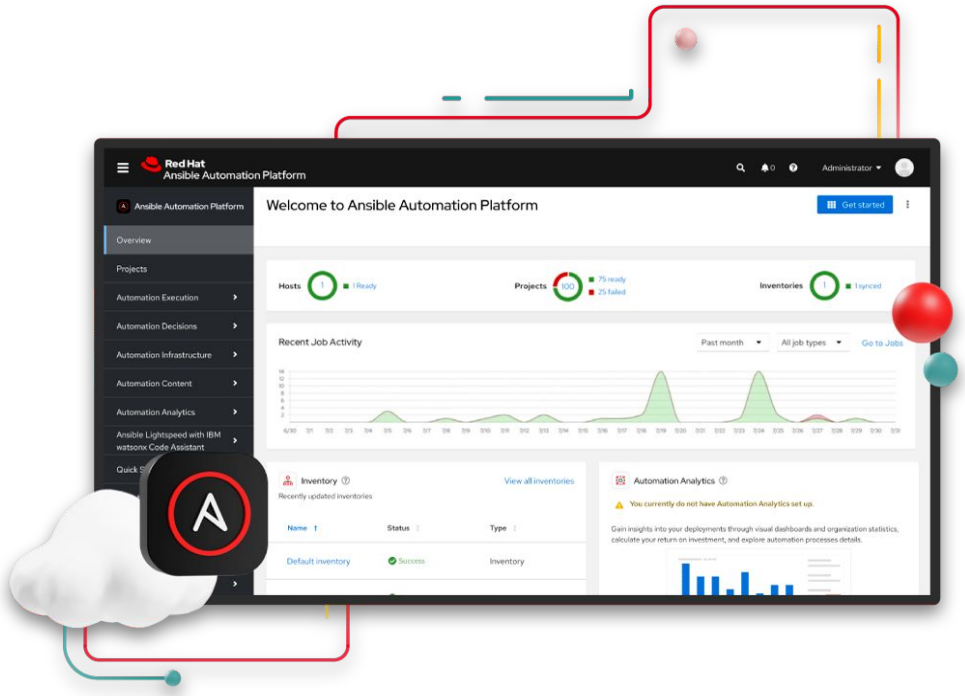
Automation dashboard

Self-service automation portal

Ansible Lightspeed Intelligent Assistant

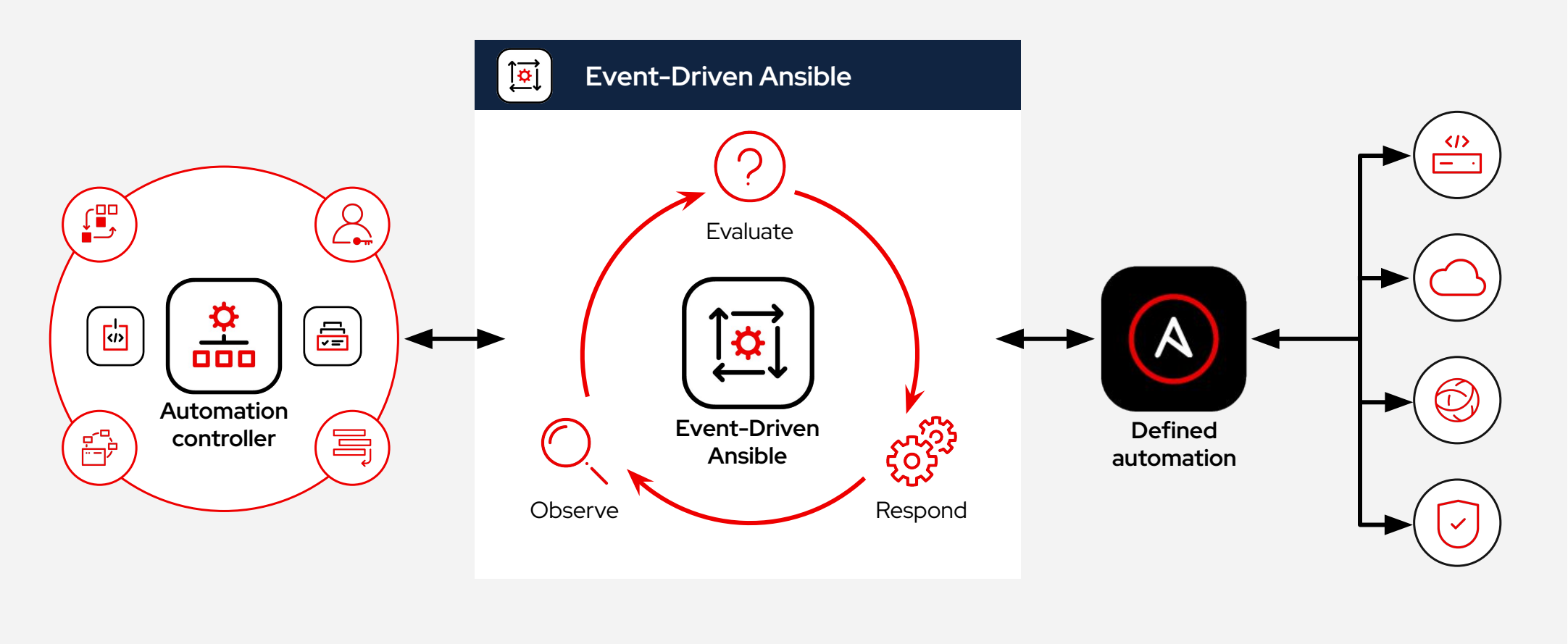
Next steps + resources

- **Read the "What's new in AAP 2.6" blog**, and be on the lookout for several more upcoming blogs that dive deeper on key features.
- **Explore the updated documentation and release notes**
- **Check out the "What's new in AAP 2.6" overview video**
- **Register for the webinar**
- **Visit www.redhat.com/whats-new-aap for all the latest**



AIOps in Ansible Automation Platform

Event-Driven Ansible. Observe, evaluate, respond.



Event-Driven Ansible | Integrations

NEW splunk> a CISCO company

NEW nautobot

INSTANA an IBM Company

dynatrace

kentik

paloalto NETWORKS

servicenow

IBM Turbonomic

LogicMonitor

CROWDSTRIKE

CYBERARK

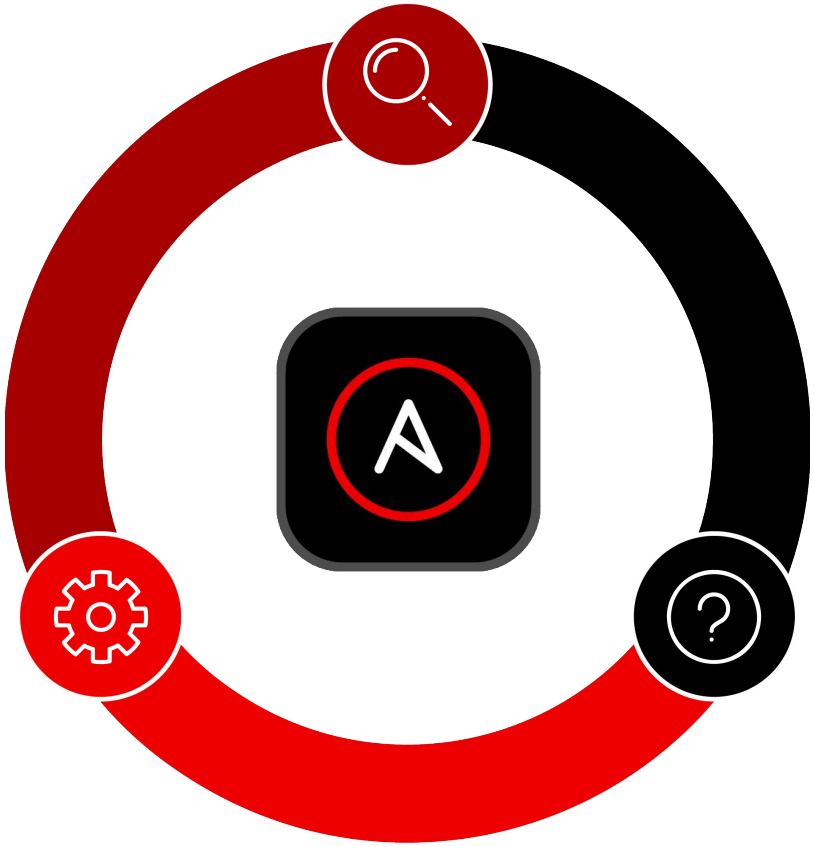
BigPanda

ZABBIX

f5

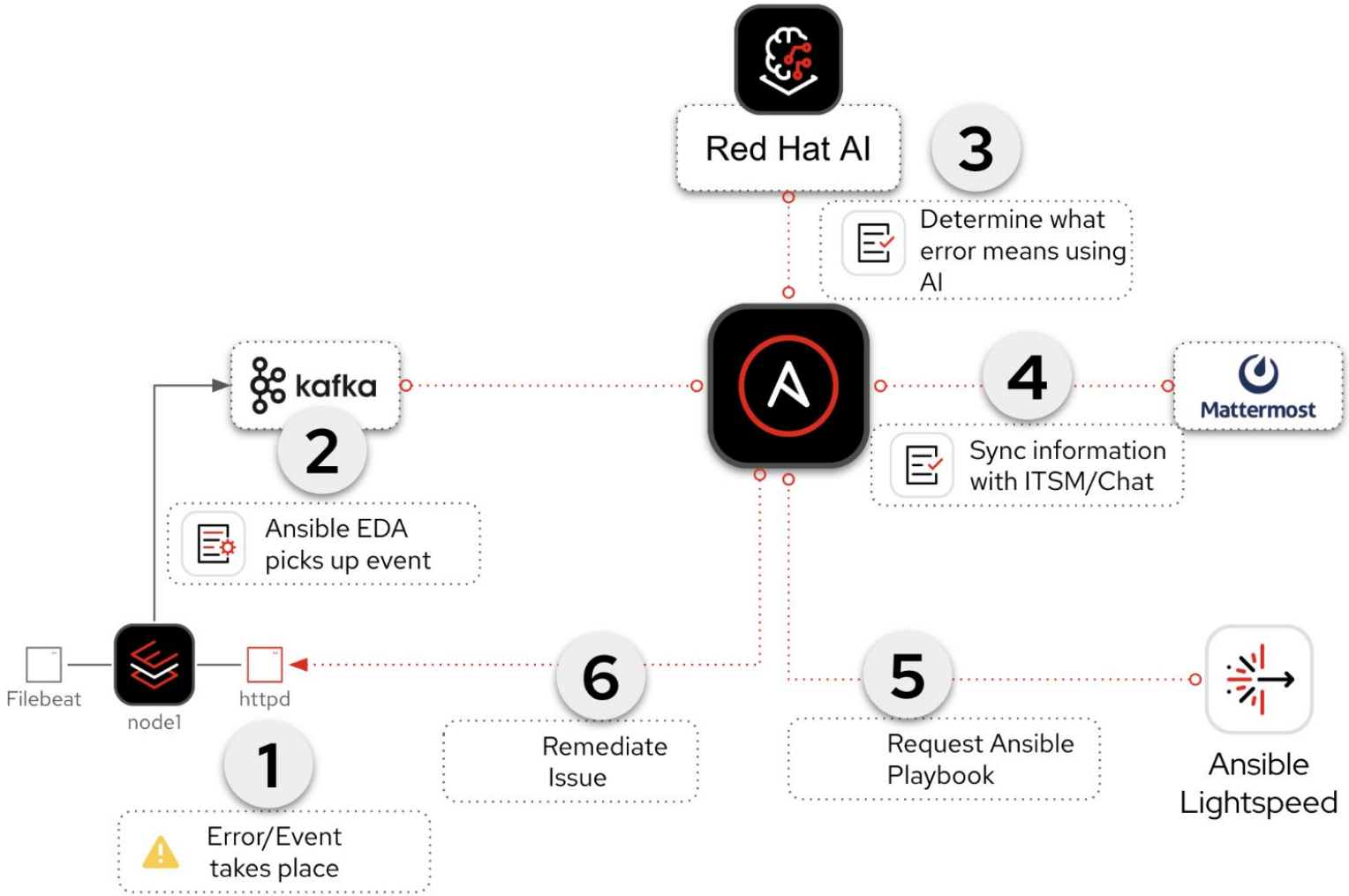
Red Hat Insights

ansible.eda



Blog: custom plugins

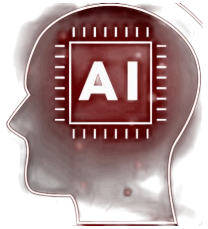
Self-healing Infrastructure | Path to zero faults



AIOps & Policy Enforcement

Monitoring Tools

Predictive **AI**



Ansible Automation Platform

Generative **AI**

Event Driven Ansible

Ansible Lightspeed

Automatic Remediation

Policy Evaluation

Red Hat Summit

Connect

Thank you

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat