



Unlock the power of AI with secure, scalable infrastructure

Take the risk and complexity out of AI adoption

Enterprises across APAC are eager to explore the potential of AI, but many are hitting the same roadblocks. The cost of validating AI use cases in the public cloud can be up to eight times higher due to token-based API services. At the same time, a lack of in-house expertise, data sovereignty concerns, and long implementation cycles make it difficult to realise the business value of AI. Without a clear path forward, organisations risk falling behind.

Technology Used

- **Red Hat OpenShift** for scaling AI workloads securely and efficiently.
- **Dell PowerEdge R760XA servers with NVIDIA L40S GPUs** – high-performance compute for demanding AI use cases.

Key benefits

- **Faster time to value:** Rapid deployment of AI workloads
- **Lower cost of entry:** Avoid high public cloud expenses
- **Data sovereignty:** Keep sensitive data on-premises
- **Enterprise-grade support:** Backed by trusted vendors
- **Scalable foundation:** Grows with your AI needs

Discover a ready-to-deploy, scalable AI solution

AI-in-a-Box provides integrated, enterprise-ready approach that brings together industry-leading infrastructure, open-source technologies, and trusted implementation support. It allows customers to validate AI functionality in a secure, cost-effective, and controlled on-premises environment—without large upfront investments or cloud costs.

- **Lower barriers to entry** – AI-in-a-box approach allows you to start small, and scale as your AI needs grow, with a competitive starting point.
- **Accelerated deployment** – Get up and running fast with proven components and a streamlined implementation process designed to reduce time to value.
- **Expert support included** – Backed by Dell, Red Hat, and specialist partners, the solution provides the technical expertise needed to support model training, deployment, and development.
- **Secure and compliant** – Deployed on-premises or in a private cloud, AI-in-a-box keeps data within your controlled environment—supporting sovereignty, security, and regulatory compliance.
- **Maximised business impact** – The solution is designed to help enterprises achieve measurable business value from AI, with a focus on real-world use cases and outcomes.



How AI-in-a-Box transforms your journey

Common challenges	AI-in-a-Box benefit
High initial AI investment	A low-risk, competitively priced path to validate AI before making large investments.
Limited internal expertise	Delivered with expert support to accelerate adoption and build internal capabilities.
Time-consuming implementation	Streamlined process reduces complexity and speeds up deployment.
Unrealised business value	Designed to support high-impact use cases and maximise ROI.
Data sovereignty concerns	On-premises deployment and sophisticated OpenShift controls keeps data securely within your control.

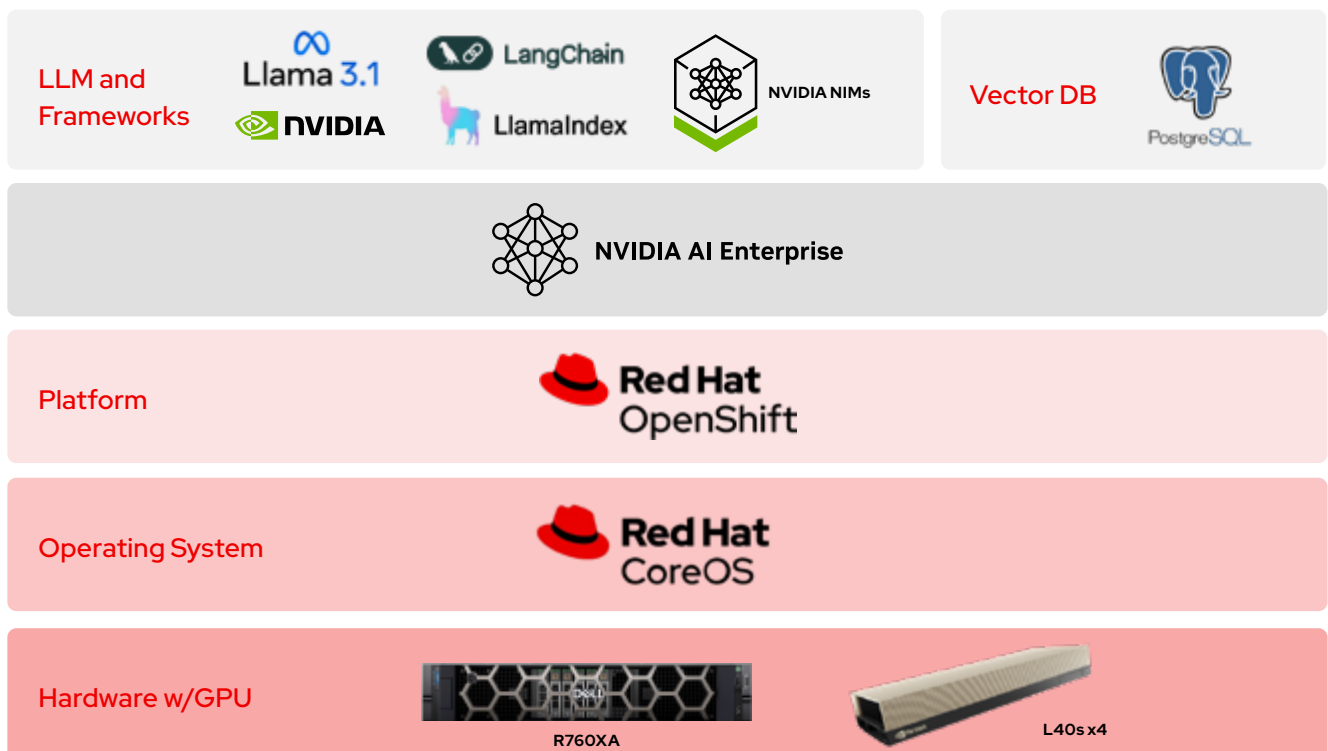


Figure 1: System architecture



Key features

The joint solution provides all the essential components needed to kickstart your AI journey – integrated and ready to deploy:

- **Model, framework & services** – Enterprise-grade support for proven AI frameworks and best practices, delivered by the world's leading vendors.
- **Use case flexibility** – Supports a broad range of AI use cases including natural language processing, enterprise knowledge bases, sales co-pilot assistants, and more.
- **Vector database** – Includes PGVector or other preferred open-source vector database solutions based on customer or partner preferences.
- **Enterprise-grade container platform** – Validated on Red Hat OpenShift, the industry's leading Kubernetes platform for hybrid cloud and AI workloads.
- **High-performance hardware** – Powered by the Dell R760XA platform with four NVIDIA L40S GPUs, built to support mission-critical AI workloads at scale.

Fast track your implementation with our streamlined process

The solution is designed for rapid deployment—so you can start exploring and validating your AI use cases quickly. While timeframes vary based on your environment and requirements, most customers can expect implementation in 4–8 weeks.

1. **Initial consultation** – Gather requirements and assess current infrastructure
2. **Solution planning** – Confirm architecture, scope, and deployment approach
3. **Hardware and platform installation** – Install Dell infrastructure and Red Hat OpenShift
4. **Integration and configuration** – Connect with existing systems and load chosen AI frameworks
5. **Testing and validation** – Ensure all components are working as expected
6. **Go live and support** – Deploy the solution and begin exploring your use cases, with ongoing support GPUs, built to support mission-critical AI workloads at scale

Partner with us

Reach out to your local Dell or Red Hat representative for more details or to schedule a consultation:

Company Name	Country	Name	Email Contact
DELL	APAC	William Tan	William.Tan@dell.com
Red Hat	Taiwan, Hong-Kong	Andrew Li	andli@redhat.com
Red Hat	Singapore, Malaysia, Indonesia, Vietnam, Thailand, Cambodia	Poh Keong Koh	pkoh@redhat.com
Red Hat	Australia, New Zealand	Ben Ros	bro@redhat.com
Red Hat	India	Manish Saran	msaran@redhat.com

About Red Hat

Red Hat helps customers standardise across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with award-winning support, training, and consulting services.

f facebook.com/redhatinc

X @RedHat

in linkedin.com/company/red-hat

North America

1-888-REDHAT1
www.redhat.com

Europe, Middle East, and Africa

00800 7334 2835
europe@redhat.com

Asia Pacific

+65 6490 4200
apac@redhat.com

Latin America

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