

Red Hat OpenShift AI sur Dell APEX Cloud Platform for Red Hat OpenShift

Déploiement et exploitation simplifiés de l'IA

Les possibilités offertes par l'intelligence artificielle (IA) suscitent un véritable engouement : 75 % des organisations déclarent avoir augmenté leur budget pour mener des initiatives d'IA¹. L'IA présente un énorme potentiel, mais les organisations doivent trouver des moyens de l'exploiter afin de créer de la valeur pour leur entreprise. Transformer ces nouvelles capacités en résultats opérationnels requiert un plan complet allant de l'architecture informatique à la gestion des données.

Red Hat® OpenShift® AI sur APEX Cloud Platform for Red Hat OpenShift est un moyen efficace de déployer un ensemble intégré d'outils Open Source et tiers communs pour collaborer sur l'entraînement, le déploiement, la surveillance et la gestion du cycle de vie des modèles d'IA/ML. Grâce à la conception validée conçue conjointement, les organisations peuvent rapidement tirer parti des principaux cas d'utilisation de l'IA.

Dell APEX Cloud Platform for Red Hat OpenShift

La solution Dell APEX Cloud Platform for Red Hat OpenShift a été conçue conjointement avec Red Hat pour optimiser et étendre les déploiements OpenShift® sur site avec une expérience opérationnelle intégrée.

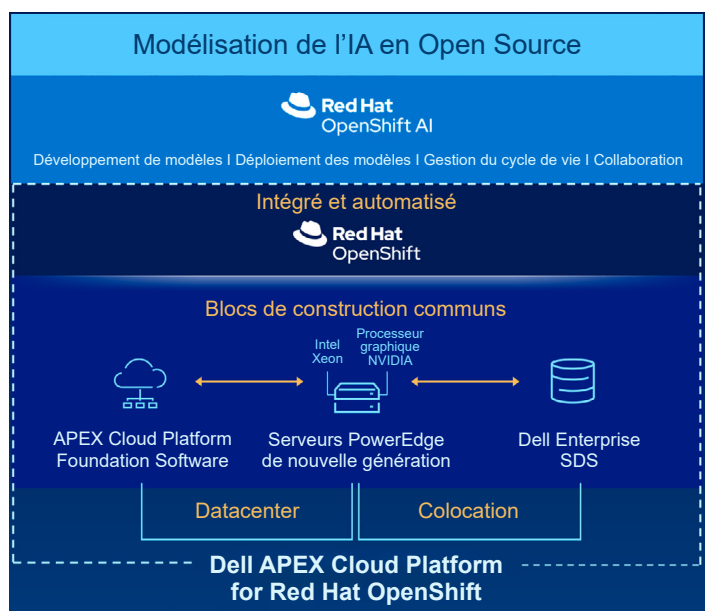
Cette plateforme clé en main offre les avantages suivants :

- Intégrations étroites et automatisation intelligente entre les couches des piles de technologies Dell et OpenShift, ce qui accélère le délai de rentabilisation et élimine la complexité inhérente à la gestion via différents outils sur des portails disparates.
- Architecture sur matériel vierge qui offre les performances, la sécurité et l'évolutivité linéaire nécessaires pour répondre aux contrats de niveau de service (SLA) les plus stricts.
- Sécurité multicouche intrinsèque, disponibilité rapide des correctifs et des mises à jour, et gouvernance OpenShift centralisée pour aider les entreprises à maintenir une posture de sécurité rigoureuse.

Red Hat OpenShift AI

Red Hat OpenShift AI vous aide à profiter de l'accélération matérielle, notamment des processeurs Intel Xeon™ Scalable de 4e génération et de l'infrastructure matérielle prise en charge par les processeurs graphiques NVIDIA, sans avoir à vous préoccuper de la gestion quotidienne de Kubernetes.

La plateforme représente une alternative aux suites d'IA/ML prescriptives proposées par les différents fournisseurs de Cloud. Les utilisateurs disposent d'un ensemble d'outils Open Source collaboratifs et d'une plateforme pour créer des modèles sans se soucier de l'infrastructure ni de la dépendance à l'égard d'outils propres au Cloud public. Il leur est ensuite possible d'étendre cette plateforme de base au moyen d'outils partenaires pour bénéficier de capacités accrues. Les modèles peuvent être déployés et mis à l'échelle de manière cohérente au sein d'environnements de production, que ce soit dans des Clouds hybrides ou à la périphérie.





Donner vie à l'IA générative

Sachant que 82 % des DI préfèrent un modèle sur site ou hybride pour l'IA générative (GenAI)¹, il est plus important que jamais de disposer d'une architecture multicloud flexible capable d'optimiser la valeur de vos données tout en réduisant la complexité.

Dell Validated Design for Red Hat OpenShift AI sur APEX Cloud Platform offre aux organisations un guide simple pour déployer un assistant numérique à l'aide d'un grand modèle de langage (LLM) et du framework de génération augmentée de récupération (RAG). Cette solution multicloud intégrée illustre la manière dont vous pouvez utiliser la GenAI en toute sécurité avec vos propres données sur site. Les assistants numériques constituent l'un des cas d'utilisation les plus courants dans tous les secteurs d'activité. L'implémentation combinée d'un LLM et du framework RAG améliore la qualité des informations et les maintient à jour.



Des services de classe mondiale pour exploiter tout le potentiel de l'IA et accélérer l'adoption des conteneurs

Les organisations qui cherchent à créer une stratégie gagnante pour la GenAI peuvent s'appuyer sur la méthodologie éprouvée des services de conseil afin de définir leur état futur idéal et d'élaborer une feuille de route exploitable alignée sur des cas d'utilisation de l'IA à forte valeur ajoutée. Les services d'infrastructure fournissent des pratiques d'excellence pour le déploiement, le support et l'assistance lors du déploiement et de la configuration d'OpenShift. Conçus pour accélérer l'utilisation de Red Hat OpenShift, les services de conseil offrent une expertise approfondie couvrant le personnel, les processus et les charges applicatives. Les services managés vous aideront à surmonter la complexité liée à la gestion et à la maintenance d'une plateforme conteneurisée nécessaire aux équipes de développement, ce qui permettra aux départements IT de gérer les charges applicatives développées dans APEX Cloud Platform for Red Hat OpenShift ou migrées vers cette plateforme.

| 82 % des DI

préfèrent un modèle sur site ou hybride pour l'IA générative (GenAI)¹



En savoir plus sur Dell APEX Cloud Platform for Red Hat OpenShift ›



Contactez un expert Dell Technologies
1-866-438-3622

¹ Dell Technologies Generative AI Pulse Survey, août et septembre 2023, www.dell.com/GenAIPulse