Oracle Expands Distributed Cloud Capabilities with NVIDIA AI Enterprise

Oracle has announced the integration of NVIDIA AI Enterprise within Oracle Cloud Infrastructure (OCI), significantly enhancing its distributed cloud capabilities and enabling accelerated sovereign AI deployments. This strategic collaboration aims to streamline AI development and deployment for enterprises seeking secure, high-performance, and scalable AI solutions.

Seamless AI Deployment Across OCI's Distributed Cloud

Unlike conventional NVIDIA AI Enterprise offerings available through third-party marketplaces, Oracle has embedded the platform directly within the OCI Console. This move simplifies the user experience by reducing deployment time, enabling direct billing, and integrating Oracle's customer support services.

The integration provides customers access to over 160 AI tools for training and inference, including NVIDIA NIM™ microservices—cloud-native inference services designed to optimize generative AI model deployment. By leveraging OCI's comprehensive suite of AI and cloud solutions, businesses can build and manage applications seamlessly across diverse distributed cloud environments, from public and sovereign clouds to dedicated and edge computing infrastructures.

Accelerating Sovereign AI and Regulatory Compliance

With global organisations increasingly prioritising data sovereignty and security, the availability of NVIDIA AI Enterprise across OCI's distributed cloud infrastructure ensures compliance with strict governance, regulatory, and security requirements. Customers can deploy NVIDIA AI Enterprise on GPU instances and Kubernetes clusters through the OCI Kubernetes Engine, benefiting from robust AI capabilities while maintaining operational control.

Oracle AI Blueprints further enhance AI adoption by providing no-code deployment recipes that allow enterprises to quickly launch AI workloads without complex infrastructure decisions. These blueprints accelerate AI deployment by offering predefined hardware configurations for NVIDIA GPUs, NIM microservices, and observability tools, reducing GPU onboarding time from weeks to mere minutes.

Global Adoption of OCI's AI Innovations

Leading enterprises are already leveraging Oracle's AI-optimized cloud services to drive innovation and meet sovereign AI requirements.

Nomura Research Institute (NRI) in Japan utilises **Oracle Alloy** to deliver customized cloud services while maintaining stringent sovereignty controls. The deployment of NVIDIA Hopper

GPUs and the planned integration of NVIDIA AI Enterprise will support NRI's AI-driven initiatives, including generative AI and large language model development.

e& UAE, a major telecom provider, has integrated NVIDIA Hopper GPU clusters within OCI Dedicated Region to develop AI-powered services tailored for government, enterprise, and smart city applications. The collaboration with Oracle and NVIDIA strengthens its ability to innovate and scale sovereign AI solutions within the UAE.

Zoom Communications has optimized its AI models to run on OCI GPU shapes, ensuring compliance with Saudi Arabia's data regulations while delivering AI-first solutions, such as the Zoom AI Companion.

A Comprehensive Cloud Ecosystem for AI

Oracle's distributed cloud offerings provide unmatched flexibility for AI-driven enterprises. With its public, dedicated, hybrid, and multicloud solutions, OCI delivers a full-stack AI infrastructure that integrates seamlessly with major cloud providers, including AWS, Google Cloud, and Microsoft Azure.

By embedding NVIDIA AI Enterprise within OCI, Oracle reinforces its position as a leader in AI and cloud innovation, empowering businesses worldwide to accelerate AI adoption while maintaining sovereignty, security, and operational control.