



NVIDIA DGX H100

La Medida Estándar para la Infraestructura de IA

La inteligencia artificial se ha convertido en el enfoque más común para resolver desafíos empresariales difíciles. Ya sea mejorando el servicio al cliente, optimizando las cadenas de suministro, extraendo inteligencia comercial o diseñando productos y servicios de vanguardia en casi todas las industrias, la IA brinda a las organizaciones el mecanismo para lograr la innovación. Y, como pioneros en la infraestructura de IA, los sistemas NVIDIA DGX™ proporcionan la plataforma de IA más potente y completa para hacer realidad estas ideas esenciales.

NVIDIA DGX H100 impulsa la innovación y la optimización empresariales. DGX H100 es la última iteración de los legendarios sistemas DGX de NVIDIA y la base de NVIDIA DGX SuperPOD™. DGX H100 es una potente plataforma de IA que cuenta con la revolucionaria GPU NVIDIA H100 Tensor Core. El sistema está diseñado para maximizar la tasa de transferencia de IA, a fin de proporcionar a las empresas una plataforma altamente refinada, sistematizada y escalable para ayudarlas a lograr avances en el procesamiento de idiomas naturales, los sistemas de recomendación, el análisis de datos y mucho más. DGX H100 está disponible en las instalaciones y a través de una amplia variedad de opciones de acceso e implementación, por lo que ofrece el rendimiento necesario para que las empresas resuelvan los mayores desafíos con IA.

La Piedra Angular de tu Centro de Excelencia de IA

La IA ha acortado la brecha entre la ciencia y los negocios. La IA ya no es el dominio de la experimentación, debido a que, día a día, empresas grandes y pequeñas la utilizan para impulsar su innovación y optimizar su negocio. DGX H100 es la cuarta generación de la primera infraestructura de IA diseñada específicamente del mundo. Se creó para ser la pieza central de un centro de excelencia de IA empresarial. Es una plataforma de hardware y software completamente optimizada que incluye soporte completo para la nueva gama de soluciones de software de IA de NVIDIA, un rico ecosistema de soporte de terceros y acceso a consejos expertos de servicios profesionales de NVIDIA. DGX H100 ofrece una confiabilidad comprobada, ya que los sistemas DGX son utilizados por miles de clientes de todo el mundo en casi todas las industrias.

Derriba las Barreras de la IA a Escala

NVIDIA DGX H100 rompe los límites de la escala y el rendimiento de la IA, dado que es el primer sistema del mundo con la GPU NVIDIA H100 Tensor Core. Ofrece 6 veces más rendimiento, redes 2 veces más rápidas con tarjetas de interfaz de red inteligentes (SmartNIC) NVIDIA ConnectX-7® y unidades de procesamiento de datos (DPU) NVIDIA BlueField-3®, además de escalabilidad de alta velocidad para NVIDIA DGX SuperPOD. La arquitectura de próxima generación está potenciada para los trabajos de IA más grandes y complejos, como el procesamiento de idiomas naturales y los modelos de recomendación de deep learning.

ESPECIFICACIONES

GPU	8 GPU NVIDIA H100 Tensor Core
Memoria GPU	640 GB en total
Rendimiento	32 PetaFLOPS de FP8
NVIDIA® NVSwitch™	4
Uso de energía del sistema	~10.2 kW máx.
CPU	x86 Doble
Memoria del sistema	2 TB
Redes	4 puertos QSFP que ofrecen 8 puertos únicos NVIDIA ConnectX-7 InfiniBand/Ethernet de 400 Gb/s VPI de 2 DPU NVIDIA BlueField-3 de doble puerto 1 InfiniBand/Ethernet de 400 Gb/s 1 InfiniBand/Ethernet de 200 Gb/s
Red de administración	NIC integrada de 10 Gb/s con RJ45 NIC opcional de Ethernet de 50 Gb/s Controlador de administración de placa base de host (BMC) con RJ45 2 BMC de DPU NVIDIA BlueField-3 (con RJ45 cada una)
Almacenamiento	Sistema operativo: 2 NVMe M.2 de 1,9 TB Almacenamiento interno: 8 unidades NVMe U.2 de 3.84 TB
Software del sistema	Los sistemas DGX H100 vienen preinstalados con DGX OS, que se basa en Ubuntu Linux e incluye la pila de software DGX (todos los paquetes y drivers necesarios están optimizados para DGX). Opcionalmente, los clientes pueden instalar Ubuntu Linux o Red Hat Enterprise Linux y la pila de software DGX necesaria por separado.
Rango de temperatura en funcionamiento	5–30 °C (41–86 °F)

Infraestructura Líder según sus Términos

La IA para los negocios no se trata solo de rendimiento y capacidades. También se trata de ajustarse ordenadamente a los límites y las prácticas de TI de una organización. DGX H100 se puede instalar en las instalaciones para la administración directa, colocarse en los data centers NVIDIA DGX-Ready, alquilarse en NVIDIA DGX Foundry y accederse a través de los proveedores de servicios administrados y certificados por NVIDIA. Además, con el programa de Administración del Ciclo de Vida DGX-Ready, las organizaciones obtienen un modelo financiero predecible para mantener su implementación a la vanguardia. Esto hace que DGX H100 sea tan fácil de usar y adquirir como la infraestructura de TI tradicional, sin carga adicional para el personal de TI ocupado, lo que permite a las organizaciones aprovechar la IA para sus negocios hoy en lugar de esperar al mañana.

Información de Pedidos

NVIDIA DGX H100 llegará a finales de 2022. Comunícate con tu representante de ventas de NVIDIA para obtener más detalles.

Más información

Para obtener más información sobre NVIDIA DGX H100, visita www.nvidia.com/es-la/data-center/dgx-h100/