[notasdeprensa.jpg](http://www.mexicopress.com.mx)Publicado en Ciudad de México el 14/04/2021

# [NVIDIA anuncia una CPU para cargas de trabajo gigantes de IA y de computación de alto rendimiento](http://www.notasdeprensa.es)

## La CPU "Grace" ofrece un salto de rendimiento 10 veces mayor para sistemas que entrenan modelos gigantes de IA, gracias a los cores Arm de eficiencia energética. El Centro de Super Computación de Suiza y el Laboratorio Nacional de Los Alamos del Departamento de Energía de EE. UU. serán los primeros en construir super computadoras de IA Potenciadas por las CPUs de NVIDIA

El día de hoy, NVIDIA presentó su primera CPU de data center, un procesador basado en Arm que ofrecerá 10 veces el rendimiento de los servidores más rápidos de la actualidad en las cargas de trabajo de computación de alto rendimiento y de IA más complejas. La CPU “Grace” es el resultado de más de 10,000 años de trabajo de ingeniería. Se diseñó para abordar los requisitos de computación de las aplicaciones más avanzadas del mundo, incluido el procesamiento de idiomas naturales, los sistemas de recomendación y supercomputación de IA, que analizan enormes conjuntos de datos que requieren una gran cantidad de memoria y un rendimiento de computación ultrarrápido. Combina cores de CPUs Arm de bajo consumo energético con un innovador subsistema de memoria de bajo consumo para ofrecer un alto rendimiento con más eficiencia. “La IA y la ciencia de datos de vanguardia impulsan la arquitectura de computación actual más allá de sus límites, a fin de procesar cantidades impensables de datos”, dijo Jensen Huang, fundador y CEO de NVIDIA. “Utilizando las IP con licenciasde Arm, NVIDIA ha diseñado Grace, una CPU para IA y HPC a gran escala. Junto con la GPU y la DPU, Grace nos brinda la tercera tecnología fundamental para la computación y la capacidad de reestructurar el data center para avanzar en la IA. Ahora, NVIDIA es una empresa de tres chips”. Grace es un procesador altamente especializado que se orienta a cargas de trabajo como el entrenamiento de los modelos NLP de próxima generación que tienen más de 1 billón de parámetros. Cuando se combina con las GPUs de NVIDIA, un sistema basado en Grace proporcionará un rendimiento 10 veces más rápido que los sistemas actuales NVIDIA DGX™ de vanguardia, que se ejecutan en CPU x86. Si bien se espera que la gran mayoría de data centers utilizarán CPU existentes, Grace (cuyo nombre es un homenaje a la pionera en programación de computación en EE. UU.), se utilizará en un segmento específico de computación. El Centro Nacional de Supercomputación de Suiza (CSCS) y el Laboratorio Nacional de Los Alamos del Departamento de Energía de EE. UU. serán los primeros en anunciar planes para construir supercomputadoras con la tecnología de Grace para apoyar los esfuerzos nacionales de investigación científica. NVIDIA presenta Grace, a medida que el volumen de datos y el tamaño de los modelos de IA crecen exponencialmente. Los modelos de IA más grandes de la actualidad incluyen miles de millones de parámetros y se duplican cada dos meses y medio. Entrenarlos requiere una nueva CPU que se puede acoplar estrechamente con una GPU para eliminar los cuellos de botella del sistema. NVIDIA desarrolló Grace aprovechando la increíble flexibilidad de la arquitectura de data centers de Arm. Al presentar una nueva CPU de clase de servidor, NVIDIA está avanzando en el objetivo de lograr diversidad tecnológica en las comunidades de IA y HPC, donde la elección es clave para ofrecer la innovación que se necesita para resolver los problemas más urgentes. Dado que Arm cuenta con la arquitectura de procesadores de licencia más utilizada del mundo, impulsa la innovación de formas nuevas e increíbles todos los días”, dijo el CEO de Arm, Simon Segars. “La introducción de la CPU de data centers Grace por parte de NVIDIA ilustra claramente cómo el modelo de licencia abierta de Arm permite una invención importante, una que respaldará aún más el increíble trabajo de los investigadores y científicos de IA en todo el mundo”. Los Primeros Adoptantes de Grace superan los límites de la Ciencia y la Inteligencia ArtificialEl CSCS y el Laboratorio Nacional de Los Alamos planean poner en línea supercomputadoras con la tecnología de Grace, construidas en cooperación con Hewlett Packard Enterprise, en 2023. “La novedosa CPU Grace de NVIDIA nos permite unir tecnologías de IA y supercomputación clásica para resolver algunos de los problemas más difíciles de la ciencia computacional”, dijo el Director de CSCS, el Prof. Thomas Schulthess. “Estamos muy contentos de poner la nueva CPU de NVIDIA a disposición de nuestros usuarios en Suiza y en todo el mundo para procesar y analizar conjuntos de datos científicos enormes y complejos”. “Con un equilibrio innovador de ancho de banda de memoria y capacidad, este sistema de próxima generación dará forma a la estrategia de computación de nuestra institución”, dijo Thom Mason, Director del Laboratorio Nacional de Los Alamos. “Gracias a la nueva CPU Grace de NVIDIA, podremos realizar investigaciones científicas avanzadas utilizando simulaciones 3D de alta fidelidad y análisis con conjuntos de datos que son más grandes de lo que era posible anteriormente”. Ofrecer un rendimiento revolucionarioLa clave para lograr el rendimiento de Grace es la tecnología de interconexión NVIDIA NVLink® de cuarta generación, que proporciona una conexión récord de 900 GB/s entre Grace y las GPUs de NVIDIA para permitir un ancho de banda agregado 30 veces mayor en comparación con los servidores líderes de la actualidad. Grace también utilizará un innovador subsistema de memoria LPDDR5x que ofrecerá el doble de ancho de banda y una eficiencia energética 10 veces mayor en comparación con la memoria DDR4. Además, la nueva arquitectura proporciona coherencia de caché unificada con un único espacio de direcciones de memoria, que combina el sistema y la memoria de la GPU de HBM para simplificar la programación. Grace será compatible con el conjunto de desarrollo de software para HPC de NVIDIA y todo el conjunto de bibliotecas CUDA® y CUDA-X™, que aceleran más de 2,000 aplicaciones de GPU, para potenciar los descubrimientos de científicos e investigadores que trabajan en los desafíos más importantes del mundo. Se espera que esté disponible a principios de 2023. Obtén más información en el evento GTC21, que se realiza en línea del 12 al 16 de abril. La inscripción es gratuita.

**Datos de contacto:**

Carlos Valencia

MKQ PR Agency

55 39 64 96 00

Nota de prensa publicada en: [https://www.mexicopress.com.mx/nvidia-anuncia-una-cpu-para-cargas-de-trabajo](http://www.notasdeprensa.es/educalivecom-disfruta-de-cursos-presenciales-y-clases-particulares-a-traves-de-internet)

Categorías: Nacional Inteligencia Artificial y Robótica Programación Hardware Software

[notasdeprensa.jpg](http://www.mexicopress.com.mx)

[**http://www.mexicopress.com.mx**](http://www.notasdeprensa.es)