



**NVIDIA MGX มอบความสะดวกให้ผู้สร้างระบบ สร้าง  
สถาปัตยกรรมแบบแยกส่วนเพื่อตอบสนองการประมวลผลแบบเร่งความเร็ว  
ที่หลากหลายของในโลกของศูนย์ข้อมูล**

*QCT และ Supermicro  
เป็นผู้ใช้กลุ่มแรกที่ใช้ข้อมูลจำเพาะของเซิร์ฟเวอร์ทำให้สามารถกำหนดค่าระบบได้มากกว่า  
100 รูปแบบ เพื่อเร่งเวิร์คโหลดด้าน AI, HPC, Omniverse*

**TAIPEI, Taiwan—COMPUTEX—May 30, 2023—**

เพื่อตอบสนองความต้องการด้านการประมวลผลแบบเร่งความเร็วที่หลากหลายของในโลกของศูนย์ข้อมูล

NVIDIA ได้เปิดตัวข้อกำหนดเซิร์ฟเวอร์ [NVIDIA MGX™](#)

ซึ่งให้สถาปัตยกรรมอ้างอิงแบบแยกส่วนแก่ผู้ผลิตระบบเพื่อสร้างเซิร์ฟเวอร์มากกว่า 100

รูปแบบอย่างรวดเร็วและคุ้มค่าเพื่อให้เหมาะกับ AI ที่หลากหลาย การประมวลผลประสิทธิภาพสูง

และแอปพลิเคชัน Omniverse

ASRock Rack, ASUS, GIGABYTE, Pegatron, QCT และ Supermicro จะใช้ MGX

ซึ่งสามารถลดต้นทุนการพัฒนาได้ถึงสามในสี่และลดเวลาในการพัฒนางานสองในสามเหลือเพียงหกเดือน

“องค์กรต่าง ๆ กำลังมองหาตัวเลือกการประมวลผลที่เร่งความเร็วมากขึ้น

เมื่อออกแบบดาต้าเซ็นเตอร์ที่ตอบสนองความต้องการทางธุรกิจและแอปพลิเคชันเฉพาะของตน” Kaustubh

Sanghani รองประธานฝ่ายผลิตภัณฑ์ GPU ของ NVIDIA กล่าว “เราสร้าง MGX

ขึ้นมาเพื่อช่วยให้องค์กรสามารถเริ่มต้นระบบ AI ขององค์กรได้

ในขณะเดียวกันก็ประหยัดเวลาและเงินจำนวนมาก”

ด้วย MGX

ผู้ผลิตจะเริ่มต้นด้วยสถาปัตยกรรมระบบพื้นฐานที่ได้รับการปรับให้เหมาะสมสำหรับการเร่งการประมวลผลสำหรับเซสซีเซิร์ฟเวอร์ จากนั้นจึงเลือก GPU, DPU และ CPU

รูปแบบการออกแบบสามารถจัดการกับเวิร์กโหลดที่ไม่ซ้ำกัน เช่น HPC, วิทยาศาสตร์ข้อมูล,

โมเดลภาษาขนาดใหญ่, Edge computing, กราฟิกและวิดีโอ, AI ขององค์กร,

การออกแบบและการจำลอง สามารถจัดการงานหลายอย่าง เช่น การฝึกอบรม AI และ 5G ได้ในเครื่องเดียว

ในขณะที่การอัปเดตเป็นฮาร์ดแวร์รุ่นใหม่ในอนาคตก็สามารถทำได้อย่างรวดเร็ว MGX

ยังสามารถรวมเข้ากับระบบคลาวด์และศูนย์ข้อมูลขององค์กรได้อย่างง่ายดาย

### **ความร่วมมือกับผู้นำในอุตสาหกรรม**

QCT และ Supermicro จะเป็นผลิตภัณฑ์แรกในตลาด โดยการออกแบบ MGX จะพร้อมในเดือนสิงหาคม ระบบ ARS-221GL-NR ของ Supermicro ที่ประกาศในวันนี้ จะรวม NVIDIA Grace™ CPU

Superchip ในขณะที่ระบบ S74G-2U ของ QCT ที่ประกาศในวันนี้ จะใช้ [NVIDIA GH200 Grace Hopper Superchip](#)

นอกจากนี้ SoftBank Corp. วางแผนที่จะเปิดตัวศูนย์ข้อมูลแบบไฮเปอร์สเกลหลายแห่งทั่วประเทศญี่ปุ่น และใช้ MGX เพื่อจัดสรรทรัพยากร GPU แบบไดนามิกระหว่างแอปพลิเคชัน Generative AI และ 5G

“ในขณะที่ Generative AI แทรกซึมอยู่ในธุรกิจและไลฟ์สไตล์ของผู้บริโภค

การสร้างโครงสร้างพื้นฐานที่เหมาะสมสำหรับต้นทุนที่เหมาะสมจึงเป็นหนึ่งในความท้าทายที่ยิ่งใหญ่ที่สุดของผู้ให้บริการเครือข่าย” Junichi Miyakawa ประธานและซีโอโอของ SoftBank Corp กล่าว “เราคาดหวังว่า NVIDIA MGX จะสามารถรับมือกับความท้าทายดังกล่าวได้ และอนุญาตให้ใช้งาน AI, 5G และอื่น ๆ

อีกมากมายขึ้นอยู่กับความต้องการปริมาณงานตามเวลาจริง”

### **การออกแบบที่แตกต่างกันสำหรับความต้องการที่แตกต่างกัน**

ศูนย์ข้อมูลจำเป็นต้องตอบสนองความต้องการมากขึ้นสำหรับทั้งความสามารถในการประมวลผลที่เพิ่มขึ้นและการลดการปล่อยก๊าซคาร์บอนเพื่อต่อสู้กับการเปลี่ยนแปลงสภาพภูมิอากาศในขณะเดียวกันก็ลดต้นทุนให้ต่ำลง

เซิร์ฟเวอร์การประมวลผลแบบเร่งความเร็วจาก NVIDIA

ได้ให้ประสิทธิภาพการประมวลผลที่ยืดหยุ่นและประสิทธิภาพการใช้พลังงานมาอย่างยาวนาน

ตอนนี้การออกแบบโมดูลาร์ของ MGX

ช่วยให้ผู้ผลิตระบบสามารถตอบสนองงบประมาณเฉพาะของลูกค้าแต่ละรายการส่งพลังงานการออกแบบความร้อนและความต้องการทางกลได้อย่างมีประสิทธิภาพมากขึ้น

### ฟอร์มแฟคเตอร์หลายแบบให้ความยืดหยุ่นสูงสุด

MGX ทำงานร่วมกับฟอร์มแฟคเตอร์ที่แตกต่างกันและเข้ากันได้กับฮาร์ดแวร์ NVIDIA รุ่นปัจจุบันและอนาคต รวมถึง:

- Chassis: 1U, 2U, 4U (ระบายความร้อนด้วยอากาศหรือของเหลว)
- GPU: พอร์ตโพลีโอ NVIDIA GPU เต็มรูปแบบรวมถึง H100, L40, L4 ล่าสุด
- CPU: NVIDIA Grace CPU Superchip, GH200 Grace Hopper Superchip, x86 CPUs
- Networking: NVIDIA BlueField-3<sup>®</sup> DPU, อะแดปเตอร์เครือข่าย ConnectX-7<sup>®</sup>

MGX แตกต่างจาก NVIDIA HGX<sup>™</sup> ตรงที่มอบความยืดหยุ่นและความเข้ากันได้หลายรุ่นกับผลิตภัณฑ์

NVIDIA เพื่อให้แน่ใจว่าผู้สร้างระบบสามารถนำการออกแบบที่มีอยู่กลับมาใช้ซ้ำได้

และปรับใช้ผลิตภัณฑ์รุ่นต่อไปได้อย่างง่ายดายโดยไม่ต้องออกแบบใหม่ราคาแพง ในทางกลับกัน HGX

นั้นใช้พื้นฐาน GPU หลายตัวที่เชื่อมต่อกับ NVLink<sup>®</sup> ซึ่งปรับให้เหมาะกับขนาดเพื่อสร้างสุดยอดระบบ AI

และ HPC

### ซอฟต์แวร์เพื่อขับเคลื่อนการเร่งความเร็ว

นอกจากฮาร์ดแวร์แล้ว MGX ยังได้รับการสนับสนุนโดยชุดซอฟต์แวร์เต็มรูปแบบของ NVIDIA

ซึ่งช่วยให้นักพัฒนาและองค์กรต่างๆ สามารถสร้างและเร่งความเร็ว AI, HPC และแอปพลิเคชันอื่นๆ ได้

ซึ่งรวมถึง [NVIDIA AI Enterprise](#) ซึ่งเป็นเลเยอร์ซอฟต์แวร์ของแพลตฟอร์ม NVIDIA AI

ซึ่งมีเฟรมเวิร์กกว่า 100 รายการ โมเดลที่ได้รับการฝึกอบรมล่วงหน้าและเครื่องมือการพัฒนาเพื่อเร่งความเร็ว

AI และวิทยาศาสตร์ข้อมูลสำหรับการพัฒนาและการปรับใช้ AI ขององค์กรที่รองรับอย่างเต็มที่

MGX เข้ากันได้กับชั้นวางเซิร์ฟเวอร์ Open Compute Project และ Electronic Industries Alliance เพื่อการรวมเข้ากับศูนย์ข้อมูลขององค์กรและระบบคลาวด์อย่างรวดเร็ว

รับชมคำปราศรัยของ Jensen Huang ผู้ก่อตั้งและซีอีโอของ NVIDIA พุดคุยเกี่ยวกับข้อกำหนดเซิร์ฟเวอร์ MGX ได้ที่ [COMPUTEX 2023](#)

### About NVIDIA

Since its founding in 1993, [NVIDIA](#) (NASDAQ: NVDA) has been a pioneer in accelerated computing. The company's invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined computer graphics, ignited the era of modern AI and is fueling the creation of the industrial metaverse. NVIDIA is now a full-stack computing company with data-center-scale offerings that are reshaping industry. More information at <https://nvidianews.nvidia.com/>.

###

### For further information, contact:

Melody Tu  
NVIDIA Asia-Pacific  
(65) 9355 1454  
[metu@nvidia.com](mailto:metu@nvidia.com)

Inez Lim  
CIZA Concept  
(65) 9756 8877  
[inezlimjie@ciza.com](mailto:inezlimjie@ciza.com)

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, performance, features and availability of our products, collaborations, services and technologies, including the NVIDIA MGX server specification, Omniverse, NVIDIA GPUs including H100, L40 and L4, NVIDIA DPUs including BlueField-3, NVIDIA CPUs including x86 CPUs, NVIDIA HPCs, large language models, and edge computing, Grace CPU Superchip, GH200 Grace Hopper Superchip, chassis, ConnectX-7 network adapters, NVIDIA HGX, NVLink, NVIDIA AI Enterprise, and the NVIDIA AI platform; our collaborations with QCT, Supermicro, ASRock Rack, ASUS, GIGABYTE, Pegatron and SoftBank Corp., and the benefits, impact, performance and availability thereof; enterprises seeking more accelerated computing options when architecting data centers to meet their specific business and application needs; and data centers increasingly needing to meet requirements for growing compute capabilities and decreasing carbon emissions, while also keeping costs down are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the most recent reports NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its annual report on Form 10-K and quarterly reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements

are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

© 2023 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, ConnectX, NVIDIA Grace, NVIDIA Grace Hopper, NVIDIA HGX, NVIDIA MGX and NVLink are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.