

Supermicro เปิดตัว AI และจักรกลเรียนรู้ระบบ แรกของอุตสาหกรรมที่ใช้ CPU รุ่นล่าสุดและ NVIDIA Tesla V100 ที่มาพร้อมกับ NVLink GPU เพื่อประสิทธิภาพขั้นสูงสุด



- Supermicro ขยายกลุ่มผลิตภัณฑ์ GPU สุดล้ำของวงการ มุ่งตอบรับความต้องการที่เพิ่มขึ้น 10 เท่าของจักรกลเรียนรู้ ปัญญาประดิษฐ์ และการวิเคราะห์ข้อมูลบิ๊กดาต้า

Super Micro Computer, Inc. (NASDAQ: SMCI) ผู้นำระดับโลกด้านโซลูชันประมวลผล จัดเก็บข้อมูล การวางเครือข่ายระดับองค์กร และเทคโนโลยีกรีนคอมพิวเตอร์ ประกาศเปิดตัวแพลตฟอร์มเซิร์ฟเวอร์ GPU ล่าสุดที่รองรับ NVIDIA(R) Tesla(R) V100 PCI-E and V100 SXM2 Tensor Core GPU accelerator ที่บูธ Platinum Sponsor ภายในงาน GPU Technology Conference (GTC) Taiwan 2018 ซึ่งจัดขึ้นที่โรงแรม Taipei Marriott วันที่ 30-31 พฤษภาคม

เพื่อการขับเคลื่อนการใช้งานคู่ขนานขั้นสูงอย่างปัญญาประดิษฐ์ (AI) จักรกลเรียนรู้ รถยนต์ไร้คนขับ สมาร์ทซิตี บริการสุขภาพ บิ๊กดาต้า การประมวลผลขั้นสูง (HPC) เวอร์ชวลเรียลลิตี (VR) และอีกมากมาย ระบบ 4U ใหม่ของ Supermicro ที่มาพร้อมเทคโนโลยีเชื่อมต่อล้ำสมัย NVIDIA NVLink(TM) ถูกปรับแต่งมาเพื่อประสิทธิภาพขั้นสูงสุด SuperServer 4029GP-TVRT ส่วนหนึ่งของแพลตฟอร์ม NVIDIA HGX-T1 class of GPU-Accelerated Server รองรับตัวเร่ง NVIDIA Tesla V100 32GB SXM2 GPU accelerator จำนวน 8 ตัว พร้อมแบนด์วิธ GPU-to-GPU ขั้นสูงสุดสำหรับแอปพลิเคชันและไฮเพอร์สเกล รวมไปถึงเทคโนโลยี NVIDIA NVLink ล่าสุดและมากกว่าแบนด์วิธของ PCI-E 3.0 ถึงห้าเท่า ระบบนี้มีโซลนความร้อน GPU และ CPU อิสระ เพื่อส่งมอบประสิทธิภาพไร้ที่ติและความเสถียรภายใต้สภาวะการทำงานที่หนักหน่วง

“ในการทดสอบหลักครั้งแรก ระบบ 4029GP-TVRT สามารถรองรับภาพจำนวน 5,188 ภาพ/วินาที บน ResNet-50 และ 3,709 ภาพ/วินาทีบน InceptionV3” ชาร์ลส์ เหลียง ประธานและซีอีโอ Supermicro กล่าว “นอกจากนี้ เรายังมองเห็นประสิทธิภาพที่ดีเยี่ยมและเพิ่มขึ้นอย่างมาก เมื่อทำการสเกลไปยังระบบหลายระบบด้วย GPU Direct RDMA นวัตกรรมล่าสุดของเรารวมเอา NVIDIA V100 32GB PCI-E และ V100 32GB SXM2 เข้าไว้ด้วยกันกับ

หน่วยความจำ 2 เท่าในระบบ 1U และ 4U ทรงประสิทธิภาพที่มาพร้อมกับ NVLink แห่งโลกอนาคต ลูกค้ำของเราสามารถเพิ่มประสิทธิภาพให้กับการใช้งานและนวัตกรรมตนเอง เพื่อแก้ปัญหาที่ซับซ้อนและท้าทายที่สุดในโลก”

เอียน บัค รองประธานและผู้จัดการทั่วไปของแผนก Accelerated Computing จาก NVIDIA กล่าวว่า “ลูกค้ำองค์กรจะได้รับประโยชน์จากความสามารถในการประมวลผลระดับใหม่ด้วยเซิร์ฟเวอร์ความหนาแน่นสูงของ Supermicro ที่ถูกพัฒนามาให้เหมาะกับ NVIDIA Tesla V100 32GB Tensor Core GPUs ซึ่งมาพร้อมกับหน่วยความจำสองเท่าที่ช่วยให้สามารถประมวลผลได้เร็วมากขึ้นถึง 50% แม้ในการเรียนรู้ข้อมูลเชิงลึกที่ซับซ้อนและการทำงานทางวิทยาศาสตร์ อีกทั้งยังช่วยให้นักพัฒนาทำงานได้ดียิ่งขึ้น โดยลดความจำเป็นในการปรับปรุงประสิทธิภาพของหน่วยความจำ”

นอกจากนี้ ระบบ GPU ของ Supermicro ยังรองรับการทำงานของ Tesla P4 อันทรงประสิทธิภาพ ซึ่งถูกออกแบบมาเพื่อเร่งความเร็วในการอนุมานปริมาณงาน ขณะที่มีการขยายเซิร์ฟเวอร์ออกไปในทุกขนาด โดยที่อุปกรณ์ดังกล่าวจะช่วยเพิ่มความเร็วให้กับตัวกลไกทรานส์โค้ดใน Tesla P4 ให้มีความสามารถในการส่งมอบวิดีโอขนาด 35 HD ได้แบบเรียลไทม์ และช่วยรวบรวมข้อมูลจากการเรียนรู้เชิงลึกมารวมไว้ในไปป์ไลน์ของการทรานส์โค้ดวิดีโอเพื่อมอบประสบการณ์การใช้งานวิดีโออัจฉริยะในระดับใหม่

Supermicro ยังได้สาธิต NVIDIA SCX-E3 class ของ GPU-Accelerated Server Platforms ที่เข้ามาเพิ่มประสิทธิภาพการทำงานให้กับระบบ 4U SuperServer 4029GR-TRT2 ซึ่งรองรับการทำงานของตัวเร่งประสิทธิภาพ 10 PCI-E NVIDIA Tesla V100 ด้วยนวัตกรรมของ Supermicro และการออกแบบ PCI-E แบบ single-root complex ให้เหมาะแก่การใช้งานกับ GPU ซึ่งช่วยเพิ่มประสิทธิภาพการทำงานให้การสื่อสารระหว่างเครือข่ายเป็นไปได้อย่างรวดเร็ว โดยความหนาแน่นที่เพิ่มมากขึ้นในแบบที่ไม่เคยมีมาก่อนนี้ ช่วยให้ SuperServer 1029GQ-TRT สามารถรองรับการทำงานของตัวเร่งความเร็ว NVIDIA Tesla V100 PCI-E GPU ได้มากถึง 4 ตัวในพื้นที่จัดเก็บ 1U แต่เพียงส่วนเดียว นอกจากนี้ SuperServer 1029GQ-TVRT ยังรองรับการทำงานของอุปกรณ์เร่งความเร็ว NVIDIA Tesla V100 SXM2 32GB GPU ที่อยู่ใน 1U ได้มากถึง 4 ตัวเช่นกัน โดยที่เซิร์ฟเวอร์ 1029GQ ทั้งสองตัวต่างก็เป็นส่วนหนึ่งของ SCX-E2 class ใน GPU-Accelerated Server Platforms

การนำเอาการวิเคราะห์บิ๊กดาต้าและจักรกลเรียนรู้มาผสานรวมไว้ด้วยกันนี้ ได้สร้างสถาปัตยกรรมใหม่ล่าสุดให้กับ NVIDIA GPU ตลอดจนปรับปรุงแมชชีนเลิร์นนิ่งอัลกอริทึมให้ดียิ่งขึ้น โดยการใช้งานในรูปแบบของการเรียนรู้ระดับลึกหรือดีปเลิร์นนิ่งนี้ ต้องอาศัยพลังการทำงานของ GPU จำนวนมาก ซึ่งต้องมีประสิทธิภาพในการสื่อสารและความสามารถในการขยายเครือข่าย GPU ระบบ GPU แบบ single-root ของ Supermicro ช่วยให้ NVIDIA GPUs สามารถสื่อสารได้อย่างมีประสิทธิภาพ เพื่อลดเวลาแฝงลงให้เหลือน้อยที่สุด ขณะที่เพิ่มประสิทธิภาพการทำงานให้มากที่สุด ตามที่วัดได้จากการทดสอบ NCCL P2PBandwidthTest

รับชมข้อมูลเพิ่มเติมเกี่ยวกับสายผลิตภัณฑ์ที่เกี่ยวข้องกับระบบ NVIDIA GPU ของ Supermicro ได้ที่

<https://www.supermicro.com/products/nfo/gpu.cfm>

ติดตาม Supermicro บน Facebook และ Twitter เพื่อรับข่าวสารและประกาศล่าสุดจากทางบริษัท

เกี่ยวกับ Super Micro Computer, Inc. (NASDAQ: SMCI)

Supermicro(R) (NASDAQ: SMCI) เป็นผู้นำด้านนวัตกรรมเทคโนโลยีเซิร์ฟเวอร์ประสิทธิภาพสูง ด้วยการจัดการระบบและโซลูชันต่าง ๆ เช่น Server Building Block Solutions(R) for Data Center, Cloud Computing, Enterprise IT, Hadoop/Big Data, HPC และ Embedded Systems ให้กับลูกค้าทั่วโลก บริษัทมุ่งมั่นในการปกป้องสิ่งแวดล้อมผ่านโครงการ “We Keep IT Green(R)” และส่งมอบโซลูชันที่เป็นมิตรต่อสิ่งแวดล้อมและประหยัดพลังงานมากที่สุดในตลาด

Supermicro, Server Building Block Solutions และ We Keep IT Green เป็นเครื่องหมายการค้า และ/หรือ เครื่องหมายการค้าจดทะเบียนของ Super Micro Computer, Inc.

แบรนด์ ชื่อ และเครื่องหมายการค้าอื่น ๆ ทั้งหมดเป็นทรัพย์สินของเจ้าของรายนั้น ๆ

SMCI-F

รูปภาพ -

https://mma.prnewswire.com/media/697266/Super_Micro_Computer_Inc___Supermicro_first_with_GPU_Systems_based_on_Latest_Gen_CPUs.jpg