

หัวเว่ย จับมือ PCL สร้าง Cloud Brain II เพื่อเริ่มต้นการเดินทางบทใหม่สำหรับคลัสเตอร์ AI ที่มีพลังประมวลผลถึง 1000 PFLOPS

- PCL เลือกใช้ Atlas 900 AI Cluster ของหัวเว่ย เตรียมพร้อมสำหรับการส่งเสริมอุตสาหกรรมประมวลผล Kunpeng

เมื่อวันที่ 29 พฤศจิกายน 2562 หัวเว่ย และ Peng Cheng Laboratory (PCL) ได้ร่วมกันเปิดตัว Peng Cheng Cloud Brain II Phase 1 ซึ่งถือเป็นการเริ่มต้นการเดินทางสู่คลัสเตอร์ AI ที่มีพลังประมวลผลสูงถึงระดับ 1000 petaFLOPS (PFLOPS) อย่างเป็นทางการ และนับเป็นหมุดหมายใหม่ในแวดวงการศึกษาทางวิทยาศาสตร์สำหรับอุตสาหกรรมประมวลผล Kunpeng ทั้งนี้ ชุมพลังที่อยู่เบื้องหลัง Cloud Brain II คือ Atlas 900 คลัสเตอร์ AI ซึ่งมีโปรเซสเซอร์ Kunpeng และ Ascend ของหัวเว่ยเป็นตัวขับเคลื่อน Atlas 900 ส่งพลังประมวลผลที่ทรงประสิทธิภาพเข้าสู่ Cloud Brain II เพื่อสนับสนุนการวิจัยพื้นฐานและการสำรวจในสาขา AI อาทิ คอมพิวเตอร์วิทัศน์ ภาษาศาสตร์ การขับเคลื่อนอัตโนมัติ การขนส่งอัจฉริยะ และบริการสุขภาพอัจฉริยะ พลังการประมวลผลของ Peng Cheng Cloud Brain ในปัจจุบันอยู่ที่ 100 PFLOPS และมีแผนที่จะเพิ่มเป็น 1000 PFLOPS และสูงขึ้นในปีหน้า

“ในเดือนกันยายนปีนี้ หัวเว่ยได้เริ่มดำเนินการตามกลยุทธ์การประมวลผลแบบ dual-engine ซึ่งได้แก่ Kunpeng + Ascend ด้วยแรงบันดาลใจจากกลยุทธ์นี้ เรามุ่งมั่นที่จะมอบที่สูงสุดแห่งพลังการประมวลผลให้กับโลก นอกจากนี้ เรายังได้เปิดตัว Atlas 900 คลัสเตอร์เทรนนิ่ง AI ที่เร็วที่สุดในโลกอีกด้วย” โหว จินหลง รองประธานอาวุโสของหัวเว่ย และประธานบริหาร กลุ่มผลิตภัณฑ์และบริการคลาวด์และ AI ของหัวเว่ย กล่าว

“วันนี้ เราภูมิใจที่ได้เห็น Atlas 900 ได้รับเลือกให้เข้าร่วมโครงการ Peng Cheng Cloud Brain II โดยถือเป็นการวางเสาหลักสำหรับ Cloud Brain II ซึ่งเป็นแพลตฟอร์มการวิจัย AI ชั้นนำของอุตสาหกรรม PCL นำกลุ่มนักวิชาการและผู้มีความสามารถในวงการวิจัย AI มารวมกัน และหัวเว่ยจะร่วมกับ PCL เพื่อเป็นแกนนำในการวิจัยทางวิทยาศาสตร์สำหรับโลกอัจฉริยะ” นายโหวกล่าวเสริม

นอกจากนี้ นายโหวยังกล่าวด้วยว่า “ณ เวลานี้ เรากำลังสร้าง Cloud Brain II Phase 1 ผมเชื่อว่า ด้วยความพยายามร่วมกันของเรา โครงการนี้จะปูทางไปสู่ Cloud Brain II ที่มีพลังประมวลผลถึง 1000 PFLOPS ในอนาคตอันใกล้ เราเชื่อมั่นว่า โครงการนี้จะกลายเป็นแพลตฟอร์มการวิจัย AI ชั้นนำของโลก”

ไมเคิล หม่า ประธานฝ่ายธุรกิจการประมวลผลอัจฉริยะของหัวเว่ย กล่าวว่า “หัวเว่ยสร้างแพลตฟอร์มประมวลผล Atlas AI ขึ้นบนโปรเซสเซอร์ Ascend หลายตัว โดยแพลตฟอร์มนี้ครอบคลุมทั้งโมดูล การ์ด เอ็ดจ์สเตชัน เซิร์ฟเวอร์ และคลัสเตอร์ AI โดยโครงสร้างพื้นฐาน AI ที่รองรับการใช้งานทุกรูปแบบของเรานั้นครอบคลุมคลาวด์-เอ็ดจ์-ดีไวซ์ (cloud-edge-device) ซึ่งมอบพลังการประมวลผลที่จำเป็นสำหรับการเทรนและการสรุปอนุมานข้อมูลเพื่อการเรียนรู้เชิงลึก”

“ผลิตภัณฑ์ Atlas รุ่นเรือธงของเราอย่าง Atlas 900 ยืนหยัดเป็นที่สุดในโลกแห่งการประมวลผล AI การรวมกันของ Atlas 900 และ Cloud Brain II จะเป็นการเริ่มต้นการเดินทางบทใหม่สำหรับคลัสเตอร์ AI ที่มีพลังประมวลผลถึงระดับ 1000 PFLOPS และจะปลดล็อกพลังการประมวลผลที่ยอดเยี่ยมเพื่อผลักดันการพลิกโฉมอุตสาหกรรมเป็นอัจฉริยะได้เร็วขึ้นทั่วทุกอุตสาหกรรม” นายหม่ากล่าว

Peng Cheng Cloud Brain เป็นศูนย์รวมเทคโนโลยีที่สำคัญในสาขาของ AI และเป็นแพลตฟอร์มการวิจัยพื้นฐานสำหรับการสำรวจขอบเขตของเทคโนโลยี AI ปัจจุบันแพลตฟอร์มนี้มีพลังประมวลผล AI ในระดับ 100 PFLOPS และคาดว่าจะแตะที่ระดับ 1000 PFLOPS ในโครงการ Cloud Brain II ปีหน้า

Cloud Brain II เป็นโครงการที่ PCL และหัวเว่ยร่วมกันสร้างขึ้น โดยทำงานบนโปรเซสเซอร์ Kunpeng และ Ascend ของหัวเว่ย และมี Atlas 900 คลัสเตอร์ AI เป็นขุมพลังการประมวลผลที่เหนือชั้น โดย PCL กำลังเดินทางพัฒนาเทคโนโลยี Cloud Brain ที่มีพลังประมวลผลสูงถึงระดับ 1000 PFLOPS

Atlas 900 คลัสเตอร์ AI พัฒนาต่อยอดขึ้นจากความรู้ความชำนาญด้านเทคโนโลยีที่หัวเว่ยสั่งสมมาเป็นเวลานานนับทศวรรษ ประกอบด้วยโปรเซสเซอร์ Ascend 910 AI หลายพันตัว ทำให้ Atlas 900 สามารถเทรนแบบจำลองการจำแนกภาพ ResNet ได้เสร็จสมบูรณ์ในเวลา 59.8 วินาที ซึ่งเร็วกว่าสถิติโลกก่อนหน้านี้ถึง 10 วินาที เมื่ออยู่ในรูปแบบ (precision) เดียวกัน ชีตความสามารถในการประมวลผลที่ทรงพลังของ Atlas 900 สร้างความแตกต่างในการวิจัยทางวิทยาศาสตร์และนวัตกรรมเทคโนโลยี อาทิ การสำรวจทางดาราศาสตร์ การพยากรณ์อากาศ รถยนต์ขับเคลื่อนอัตโนมัติ และการสำรวจน้ำมัน เป็นต้น โดย Atlas 900 มีจุดเด่นดังนี้:

- การประมวลผลอันทรงพลัง: Atlas 900 ประกอบด้วยโปรเซสเซอร์ Ascend 910 AI หลายพันตัว Atlas 900 ให้พลังการประมวลผล 256-1024 PFLOPS เมื่ออยู่ในรูปแบบ half precision (FP16) ซึ่งเทียบเท่ากับพลังประมวลผลของคอมพิวเตอร์พีซี 500,000 เครื่อง ดีไซน์ของ SoC ได้รับการออกแบบให้ผนวกรวมการประมวลผล AI การประมวลผลเพื่อวัตถุประสงค์ทั่วไป และฟังก์ชัน I/O เพื่อปรับปรุงประสิทธิภาพการเทรนอย่างได้ผล
- เครือข่ายคลัสเตอร์ความเร็วสูง: รองรับอินเทอร์เฟซเครือข่ายความเร็วสูง 3 ประเภท ได้แก่ Huawei Cache Coherence System (HCCS), PCIe 4.0 และ 100G RoCE ซึ่งลด gradient synchronization latency ลงไปได้ถึง 10% มาอยู่ที่ระดับ 70% เพื่อเพิ่มประสิทธิภาพการเทรนขึ้นเป็นหลายเท่าตัว นอกจากนี้ยังใช้อัลกอริทึม iLossless intelligent switching ที่เป็นนวัตกรรม เพื่อสนับสนุนการเทรนและการเรียนรู้กราฟฟิกทั่วเครือข่ายใน

แบบเรียลไทม์ เพื่อการสร้างเครือข่ายที่มีการสูญเสียเป็นศูนย์ (zero packet loss) และความหน่วง end-to-end ในระดับไมโครวินาที

– ที่สุดของการระบายความร้อน: Atlas 900 ใช้ระบบการถ่ายเทความร้อนแบบ adiabatic ซึ่งมีค่าการระบายความร้อนด้วยของเหลวสูงกว่า 95% และมีค่าประสิทธิภาพการใช้พลังงาน (PUE) ระบบน้อยกว่า 1.1 (ค่า PUE ที่เหมาะสมที่สุดคือ 1.0)

จนถึงขณะนี้ หัวเว่ยได้เปิดตัว Atlas 900 คลัสเตอร์ AI, Atlas 800 เซิร์ฟเวอร์ AI, Atlas 500 เอ็ดจ์สเตชัน AI, Atlas 300 การ์ด AI accelerator และ Atlas 200 โมดูล AI accelerator ผลิตภัณฑ์ทั้งหมดนี้เป็นส่วนหนึ่งของชุดผลิตภัณฑ์ Atlas ครอบคลุมรูปแบบการใช้งานคลาวด์ เอ็ดจ์ และอุปกรณ์ดีไวซ์ และมอบพลังการประมวลผลที่จำเป็นสำหรับการเทรนและการสรุปอนุมานข้อมูล

สำหรับก้าวต่อไป หัวเว่ยจะยังคงเดินหน้าเพิ่มการลงทุนและนวัตกรรมในโครงสร้างพื้นฐาน เช่น โปรเซสเซอร์ ระบบปฏิบัติการ และฐานข้อมูลสำหรับอุตสาหกรรมประมวลผล Kunpeng พร้อมทั้งสนับสนุนการประมวลผล AI ให้กับ Peng Cheng Cloud Brain ซึ่งจะช่วยให้เทคโนโลยี AI เร็วขึ้นได้จริงสำหรับการใช้งานในขอบเขตที่กว้างขึ้น

รูปภาพ - <https://photos.prnasia.com/prnh/20191203/2659308-1-a>

คำบรรยายภาพ: โหว จินหลง (ซ้าย) รองประธานอาวุโสของหัวเว่ยและประธานบริหารกลุ่มผลิตภัณฑ์และบริการคลาวด์และ AI ของหัวเว่ย และ เกา เหวิน (ขวา) ผู้อำนวยการของ Peng Cheng Lab ในพิธีเปิด

รูปภาพ - <https://photos.prnasia.com/prnh/20191203/2659308-1-b>

คำบรรยายภาพ: ไมเคิล หม่า ประธานฝ่ายธุรกิจการประมวลผลอัจฉริยะของหัวเว่ย แนะนำ Atlas 900