

Nvidia mejora el rendimiento en IA con una Tesla V100s formato PCIe más potente

Nvidia anunció hace un par de años la arquitectura Volta que integraba núcleos de procesamiento tensorial, pero estaba disponible en dos formatos con características distintas.

La versión más potente era el tipo tarjeta con conector NVLink para servidores especiales, mientras que el modelo de tarjeta PCIe para cualquier PC normal era más lenta. Esto cambia con la nueva Tesla V100s.

Este modelo pasa a ser más potente que el modelo anterior NVLink por un estrecho margen, pero pasa su consumo máximo de 300 W a 250 W. La potencia también pasa de 7.8 TFLO/s a 8.2 TFLO/s en doble precisión, y en aprendizaje profundo a través de sus 640 núcleos tensoriales pasa de 125 TFLO/s a 130 TFLO/s.



Incluye 32 GB de HBM2 de más velocidad, que pasa a funcionar a unos 2.2 GHz, lo cual le da un ancho de banda estimado de 1134 GB/s frente a los 900 GB/s de los modelos anteriores. Pero por lo demás no cambia sustancialmente, y eso teniendo en cuenta que los rumores apuntan a que la arquitectura Ampère sería presentada para el sector de inteligencia artificial en torno a marzo o abril de 2020.

Fuente: Guru3D.