

NVIDIA lancia Volta, ma solo per HPC ...



LINK (<https://www.nexthardware.com/news/schede-video/8060/nvidia-lancia-volta-ma-solo-per-hpc-.htm>)

Tesla V100 è una soluzione espressamente pensata per il calcolo e il deep learning.

NVIDIA ha appena lanciato la sua nuova architettura Volta nella forma di Tesla V100, un acceleratore di calcolo ad alte prestazioni per supercomputer.

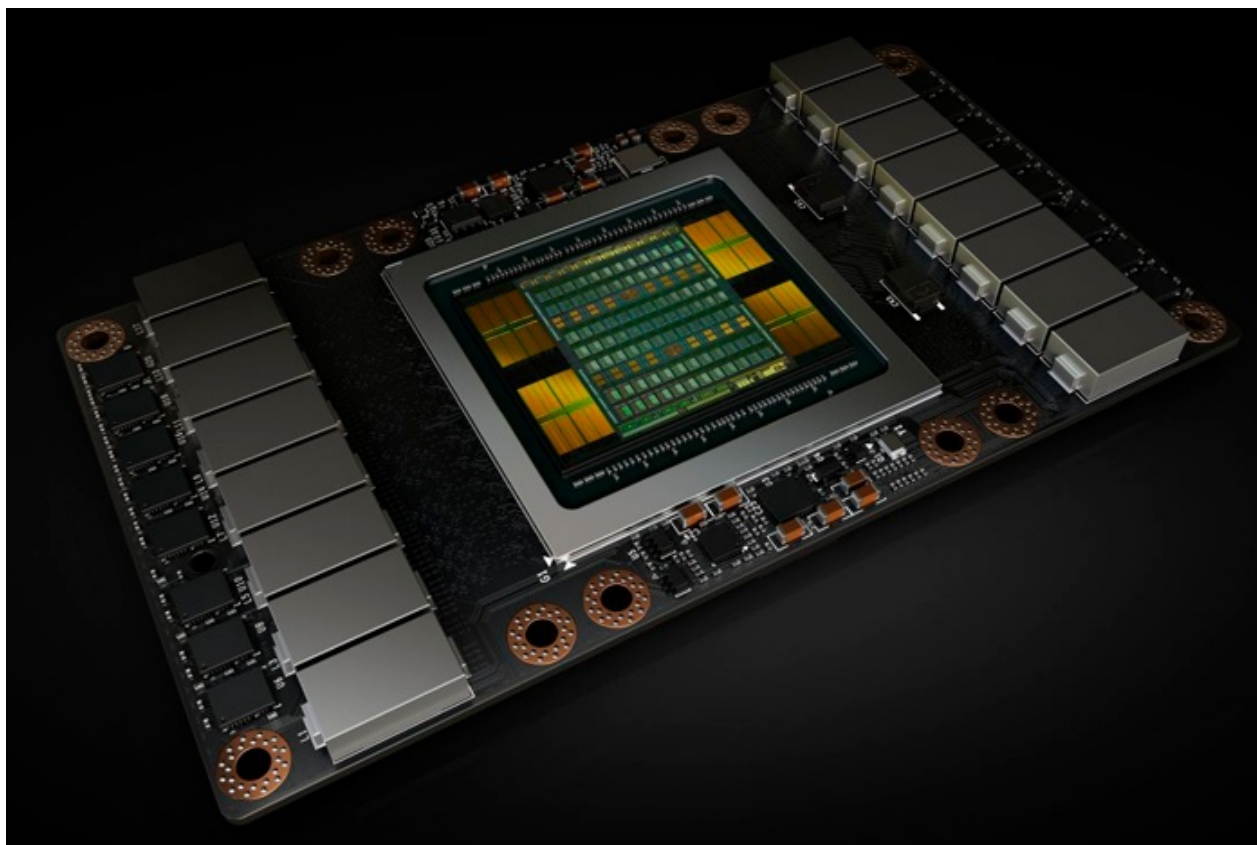
Come la precedente P100, V100 non assomiglia particolarmente alle schede grafiche che conosciamo (risulta mancante anche il connettore video) ed a livello di silicio le differenze sono importantissime.

Questo mostro da 21 miliardi di transistor lancia infatti un nuovo concetto di unità di elaborazione etichettato da NVIDIA come Tensor Core, espressamente progettato per il Deep Learning, e può contare su elevate prestazioni in FP64.

La GPU Volta integra 5120 CUDA Cores a 1455MHz realizzati con la nuova ed inedita tecnologia a 12nm FinFET e collegati ad un quantitativo di memoria HBM2 pari a 16GB con una banda di "appena" 900 GB/s.

Il risultato è un chip di discrete dimensioni, 815 mm² contro i 610 mm² di Tesla P100, con un TDP di 300W che è il risultato di un'ottimizzazione dell'efficienza energetica migliore del 50% rispetto a Pascal, un dato impressionante.

Quello che tutti ci chiediamo, a questo punto, è quando arriverà Volta su una GeForce...



Sulla base dei precedenti, l'acceleratore Tesla dovrebbe precedere le GeForce consumer di qualche mese, ma rispetto allo scorso anno ci sono alcuni presupposti molto differenti.

Per cominciare, la pressione della concorrenza sulla fascia alta, ovvero le soluzioni più redditizie, si sta rivelando più bassa del previsto (tanto per usare un eufemismo) ed inoltre NVIDIA ha lanciato solo da pochissimo tempo la nuova Titan Xp e la GTX 1080 Ti che, in qualche modo, dovranno essere smaltite.