

La mente di Ryzen abbandona AMD per Intel

Intel ha appena annunciato ufficialmente di aver assunto **Jim Keller** come **Senior VP** per guidare il team di ingegneri, occupandosi dello sviluppo **SoC (system-on-chip)** all'integrazione del silicio. **Murthy Renduchintala**, chief engineering di Intel, spiega così la sua assunzione:

«Jim è uno dei più rispettati visionari del settore della microarchitettura e l'ultimo esempio di talento tecnico per entrare a far parte di Intel. Abbiamo intrapreso iniziative entusiasmanti per cambiare radicalmente il modo in cui utilizziamo il silicio, mentre entriamo nel mondo di processi e architetture eterogenei. Jim si unirà a noi per accelerare questa trasformazione»

Jim Keller ha una vasta esperienza su tutti i tipi di processori e SoC. Tra gli appassionati di PC è forse il più noto architetto dell'eccellente microarchitettura **Zen** di **AMD**, che alimenta tutti gli ultimi processori **Ryzen** ed **EPYC** che hanno spinto AMD in una posizione di rilievo nel mercato delle CPU. Ma Keller ha anche lavorato in **Apple**, dove ha guidato il team di progettazione dei processori iniziali **A4** e **A5** e più recentemente, alla **Tesla**, dove è stato vice presidente di **Autopilot** e **Low Voltage Hardware**. Proprio Keller ci parla di questo:

«Ho acquisito una grande esperienza lavorando in Tesla, ho imparato molto e attendo con ansia che tutta la tecnologia proveniente da Tesla in futuro possa essere utile. La mia passione per tutta la vita è stata lo sviluppo dei migliori prodotti al mondo fatti di silicio. Il mondo sarà un posto molto diverso nel prossimo decennio visto la continua evoluzione dell'informatica. Sono entusiasta di unirmi al team Intel per costruire il futuro di CPU, GPU, acceleratori e altri prodotti per l'era dell'informatica basata sui dati»

Keller non è l'unico ex membro di AMD che approda a Intel di recente. [Raja Koudhuri, in precedenza a capo del gruppo Radeon Technologies di AMD](#) è entrato a far parte del team di lavoro sulle GPU.

NVIDIA contro il mining: nuove regole per i distributori

La maggior parte degli appassionati di hardware - anche quelli con un lieve interesse nei PC - è ben consapevole del fatto che i **miner** stanno comprando più **GPU** che possono per alimentare la corsa all'oro della **criptovaluta**. I retailer stanno ora fronteggiando un blocco da **NVIDIA**, che ha [ufficialmente confermato](#) i passi che prenderanno per limitare la fornitura di GPU ai miner e concentrare le proprie **GeForce** nelle piattaforme di gioco. Mentre NVIDIA ha ricevuto recentemente alcuni problemi per i limiti applicati ai server basati su GeForce, questa nuova politica restrittiva sembra applicarsi a entrambe le categorie delle proprie schede video. GeForce è per i giocatori, mentre **Quadro** e **Tesla** sono per professionisti. L'affermazione di NVIDIA è un segno,

anche se piccolo, di resistenza verso tutti quei miner che stanno bloccando il mercato delle schede video. Non è un gesto particolarmente rilevante da parte di NVIDIA, ma solo una “raccomandazione” che già aveva fatto ai distributori.

«Per NVIDIA i giocatori vengono prima di tutto. Tutte le attività relative alla nostra linea di prodotti GeForce sono focalizzate sul nostro pubblico principale. Per garantire che i giocatori con GeForce continuino ad avere una buona disponibilità delle schede grafiche GeForce nella situazione attuale, raccomandiamo che i nostri partner commerciali facciano gli accordi appropriati per soddisfare le esigenze dei giocatori come al solito»

(Boris Böhles, NVIDIA GmbH)

Se la promessa di NVIDIA sarà mantenuta, vedremo presto più schede video per i giocatori. Molti, se non la maggior parte dei rivenditori, applicano già alcune variazioni del limite di una GPU per cliente, con scarso successo. Un sacco di schede grafiche non sembrano mai arrivare ai negozi, in primo luogo perchè i miner acquistano direttamente dai distributori o da partner vari. **AMD**, d'altra parte, apparentemente non può permettersi di reprimere i minatori. Affidarsi all'estrazione di criptovalute per la crescita del business è di per sé un approccio rischioso ma proficuo per AMD stessa poiché molti miner si sono affidati proprio a quest'ultima. Forse questo passaggio da NVIDIA respingerà alcuni cambiamenti positivi, ma sembra senza modifiche dirette alla catena di produzione e di vendita al dettaglio, i giocatori affronteranno delle brutte esperienze riguardo l'aggiornamento della propria scheda grafica per l'immediato futuro, senza contare il prezzo più alto di **SSD** e **DRAM**. In ogni caso, il 2018 è destinato a essere un anno costoso per i PC gamer.

NVIDIA: presentata TITAN V Volta, un mostro di potenza da 3000\$

NVIDIA ha appena **annunciato** la sua ultima scheda grafica **TITAN** basata sull'architettura **Volta**, la **TITAN V**. **NVIDIA TITAN V** presenta le più recenti tecnologie GPU come l'architettura **Volta 12nm** accoppiata con **12 GB** di memoria **HBM2**.

NVIDIA TITAN V, uno mostro da 3000\$ con 5120 CUDA Cores, 12 GB HBM2 VRAM e tecnologia a 12nm

La scheda grafica **NVIDIA TITAN V** è dotata dell'ultima architettura **Volta** da **12nm** e, come tale, è dotata delle più recenti tecnologie che **NVIDIA** ha da offrire. In primo piano nella famiglia **TITAN**, la GPU mirerà al mercato dei prosumer e in quanto tale, ci si può aspettare un prezzo molto alto in quanto questa bestia avrà un costo di **3000\$**. Annunciata dal fondatore e CEO di **NVIDIA** **Jensen Huang** alla conferenza annuale **NIPS**, **TITAN V** eccelle nell'elaborazione computazionale per la simulazione scientifica. I suoi **21,1 miliardi di transistor** erogano **110 teraflops** di potenza, **9 volte** quella del suo predecessore, con un'estrema efficienza energetica.

«La nostra visione di Volta era quella di spingere i limiti estremi del calcolo ad alte prestazioni e dell'intelligenza artificiale. Abbiamo aperto nuovi orizzonti con la nuova architettura del processore a 12 nm, le istruzioni, i formati numerici, l'architettura della memoria e i collegamenti del processore. Con TITAN V, stiamo mettendo Volta nelle mani di ricercatori e scienziati di tutto il mondo. Non vedo l'ora di vedere le loro scoperte rivoluzionarie»
(Jensen Huang, CEO di NVIDIA)

Non si ottiene solo la straordinaria nuova architettura Volta "GV100", ma gli acquirenti hanno anche 12 GB di memoria HBM2. Questa è la prima scheda grafica TITAN e anche la prima linea di schede grafiche NVIDIA con memoria HBM2.

La NVIDIA TITAN V è basata sull'architettura GV100 e dispone di un totale di **5120 CUDA core e 320 texture unit**. Questa è esattamente la stessa quantità di core presenti su **Tesla V100**. Oltre ai core regolari, la scheda include anche **640 Tensor Core** all'interno della GPU Volta. Questi sono orientati alla massima performance in quanto la scheda può generare fino a **110 TFLOPS** di prestazioni per algoritmi relativi all'IA. La totalità del core è sincronizzata su base **1200 MHz e boost a 1455 MHz**. Anche con caratteristiche così pesanti, la scheda richiede solo un connettore di alimentazione a 8 e 6 pin per l'avvio e arriva a consumare fino a 250 W. Quindi, venendo alla VRAM HBM2, come detto ci sono 12 GB con una velocità dati di **1,7 Gbps** lungo un bus di memoria a **3072 bit**. Questo dà alla scheda una larghezza di banda totale di **652,8 GB/s**, che è molto più veloce della precedente **TITAN Xp**. Rispetto a Tesla V100, stiamo considerando un'interfaccia bus cut down (**4096-bit vs 3072-bit**) e anche una VRAM inferiore di **12 GB** rispetto a **16 GB** su quella scheda.

Nel complesso, questa scheda grafica può essere utilizzata per carichi di lavoro sia professionali che regolari come i giochi, e sarà interessante vedere come si comporterà questo asso di NVIDIA. Mentre il prezzo è decisamente alto, a bordo di TITAN V ci sono molte cose che le normali schede non hanno, elemento che la rende adatta a carichi di lavoro professionali. Queste funzionalità includono:

- **Calculation Cores FP64 dedicati**
- **Tensor Calculation Cores dedicati**
- **12 GB HBM2 con Interfaccia a 3072-bit**
- **Interfaccia NVLINK 2.0**

Foto dettagliate della TITAN V

A parte le specifiche, NVIDIA TITAN V offre lo stesso **Cooler NVTTM**, che abbiamo imparato a conoscere e amare nelle schede della serie **Pascal GeForce 10**. La GPU presenta solo una differenza tra le altre della serie 1000: il nome inciso sulla scocca. Inoltre è dotata di un magnifico corpo in alluminio **pressofuso d'oro** e di un sistema di raffreddamento delle **head pipe** superiore per le migliori performance termiche possibili. Il **PCB è un DrMOS a 16 fasi** con funzionalità di monitoraggio della corrente e di monitoraggio termico in tempo reale.

[NVIDIA presenta al GTC 2017 la nuova Drive PX Pegasus](#)

NVIDIA ha presentato la nuova scheda madre per auto **Drive PX Pegasus** alla **GPU Technology Conference Europe 2017**.

Questa rappresenta un sostanziale miglioramento rispetto alla precedente versione **Drive PX 2** che si traduce in un netto miglioramento delle prestazioni di quasi **10x**. Con oltre **320 Deep Learning Tera Operations** al secondo (termine di misurazione delle performance coniato da NVIDIA), questo hardware è sostanzialmente più veloce rispetto alla maggior parte dei rack server dello scorso anno. Per capire di quanto NVIDIA abbia migliorato le prestazioni basti pensare che l'ultima iterazione (DRIVE PX 2) era capace di sole **24-30 DL TOPs**. L'azienda ha anche affermato che questo hardware è capace di garantire il massimo di guida autonoma ovvero di livello 5. Il quinto livello di A.D. (autonomous driving) si riferisce ad un sistema completamente autonomo, che prevede che le prestazioni del veicolo siano uguali a quelle che avrebbe avuto lo stesso con un conducente umano, in ogni scenario di guida, compresi ambienti estremi come ad esempio strade sterrate.

Drive PX 'Pegasus' di NVIDIA: 10x le prestazioni DL TOPs e il primo hardware teoricamente capace di autopilota con "autonomia livello 5"

Il mercato dell'automobile rappresenta un'enorme opportunità per i produttori di chip e NVIDIA sembra essere in grado di gestirne le crescenti esigenze. L'approccio **GPGPU** alla guida di auto sta rapidamente guadagnando popolarità e NVIDIA è una delle poche aziende, in grado grazie alla sua esperienza con le GPU grafiche di sapere come aumentare velocemente le prestazioni.

La scheda Drive PX Pegasus non fa eccezione, e promette un enorme aumento delle prestazioni, oltre ad essere una delle prime che probabilmente ci spingerà oltre l'autonomia di livello 3 (come [l'autopilota prodotto da Tesla](#)). Non sappiamo se le affermazioni della società riguardanti la capacità di autonomia di livello 5 siano accurate, ma una cosa è certa: il Drive Px Pegasus ha molta più potenza (DL TOPs) rispetto a molti rack server dell'anno scorso.

"La creazione di un'auto a guida automatata è uno più grandi progetti della società moderna - e uno dei più impegnativi da raggiungere", ha dichiarato **Jensen Huang**, fondatore ed amministratore delegato di NVIDIA. ***"Le prestazioni e l'efficienza di calcolo AI di Pegasus sono fondamentali per l'industria e per realizzare questa visione"***.

"Le vetture senza conducente consentiranno nuovi servizi di condivisione di auto e di guida. Saranno inventati nuovi tipi di auto, somiglianti a uffici, saloni o camere d'albergo su ruote. I viaggiatori semplicemente ordineranno il tipo di veicolo in base alla destinazione e alle attività pianificate lungo la strada. Il futuro della società sarà ridisegnato".

A differenza della vecchia Drive PX, la scheda Pegasus non sarà raffreddata ad acqua ma esclusivamente dall'aria (il che è un'impresa enorme considerando che ha un TDP di 500 Watts) che ne ridurrà la necessità di manutenzione e ne aumenterà la durata a causa di parti meno mobili e soggette a rottura. La scheda Pegasus è dotata di 2x Xenver SOCs con Volta iGPUs e 2x GPU dedicate. La scheda è certificata **ASIL D**, il che significa che è progettata per avere **failover**.

La DRIVE PX Pegasus è composta da quattro processori AI ad alte prestazioni, una coppia di processori **Xavier** su Soc, con una GPU basata sull'architettura **Volta** incorporata, due GPU discrete di nuova generazione create per accelerare l'apprendimento e gli algoritmi informatici. Il sistema fornirà un'enorme capacità computazionale per veicoli completamente autonomi in un computer della dimensione di una targa, riducendo drasticamente i consumi di energia e il costo complessivo.

Pegasus è stato progettato per la certificazione ASIL D - il livello di sicurezza più elevato del settore - con ingressi/uscite automobilistici, tra cui **CAN** (controller area network), **Flexray**, 16 ingressi ad alta velocità per fotocamere, radar, lidar e ultrasuoni, più connettori Ethernet a 10Gbit. La sua larghezza di banda di memoria combinata supera 1 terabyte al secondo.

Forse l'aspetto più interessante del kit è il fatto che le GPU dedicate utilizzate non sono Volta, ma basate su un'architettura di nuova generazione. Ciò significa che sono maggiormente ottimizzate ed efficienti rispetto alle Volta iGPUs dei SOC precedenti. Considerando che non abbiamo ancora sentito parlare di un'architettura che riesca a reggere il confronto con Volta, è possibile che questa particolare architettura sia stata progettata esclusivamente per applicazioni di apprendimento profondo (Deep Learning). Le aziende come **Mobileye**, **Intel** e **Tesla** hanno intenzione di intraprendere il progetto, ma è chiaro che Nvidia sia in questo momento in testa.