nexthardware.com

a cura di: **Vittorio Gamba - Vittorio Gamba - 28-01-2015 11:30**

NVIDIA GeForce GTX 970: il "memory-gate"

LINK (https://www.nexthardware.com/news/schede-video/6697/nvidia-geforce-gtx-970-il-memory-gate.htm)

Cerchiamo di fare il punto su uno degli argomenti più discussi degli ultimi giorni ...

Alcuni utenti utilizzando benchmark che facevano uso intensivo della memoria a bordo della scheda, avevano sperimentato dei micro stuttering che li avevano indotti a ritenere che le schede non fossero in grado di indirizzare gli ultimi 500-700 MByte del buffer grafico.

Lo scorso weekend gli ingegneri NVIDIA hanno formalizzato una risposta a questo "j'accuse" degli utenti sostenendo che le GeForce GTX 970 sono perfettamente in grado di indirizzare tutti i 4 GByte di memoria a loro disposizione ma che, effettivamente, e diversamente dalle GeForce GTX 980, lo fanno in modo un po' particolare.

Per poter gestire al meglio questa situazione e il traffico di dati verso la memoria stessa, quindi, gli ingegneri NVDIA hanno suddiviso il buffer grafico in due sezioni, una da 3,5 GByte e una da 0,5 GByte.

Se invece l'occupazione di memoria è superiore a 3,5 GByte, entrambe le "partizioni" vengono utilizzate e quindi il software riporterà correttamente i 4 GByte totali.

La domanda a questo punto è spontanea: quali sono le ragioni di questa suddivisione e quali sono le prestazioni delle due aree di memoria?

Se la risposta fornita dagli ingegneri della casa californiana sembra un po' criptica, quella fornita in seguito e riportata da altre testate lascia sicuramente pochi dubbi sulla correttezza formale delle informazioni ricevute prima del lancio della scheda.

Sembra, infatti, che le specifiche dichiarate e riportate sui documenti confidenziali su cui tutti ci siamo basati per le recensioni, non fossero propriamente corrette.

Come ricorderete, le specifiche che abbiamo riportato, e tratto direttamente dai documenti ufficiali NVIDIA, parlavano di 13 SMM, 64 ROP e 2MB di cache L2.

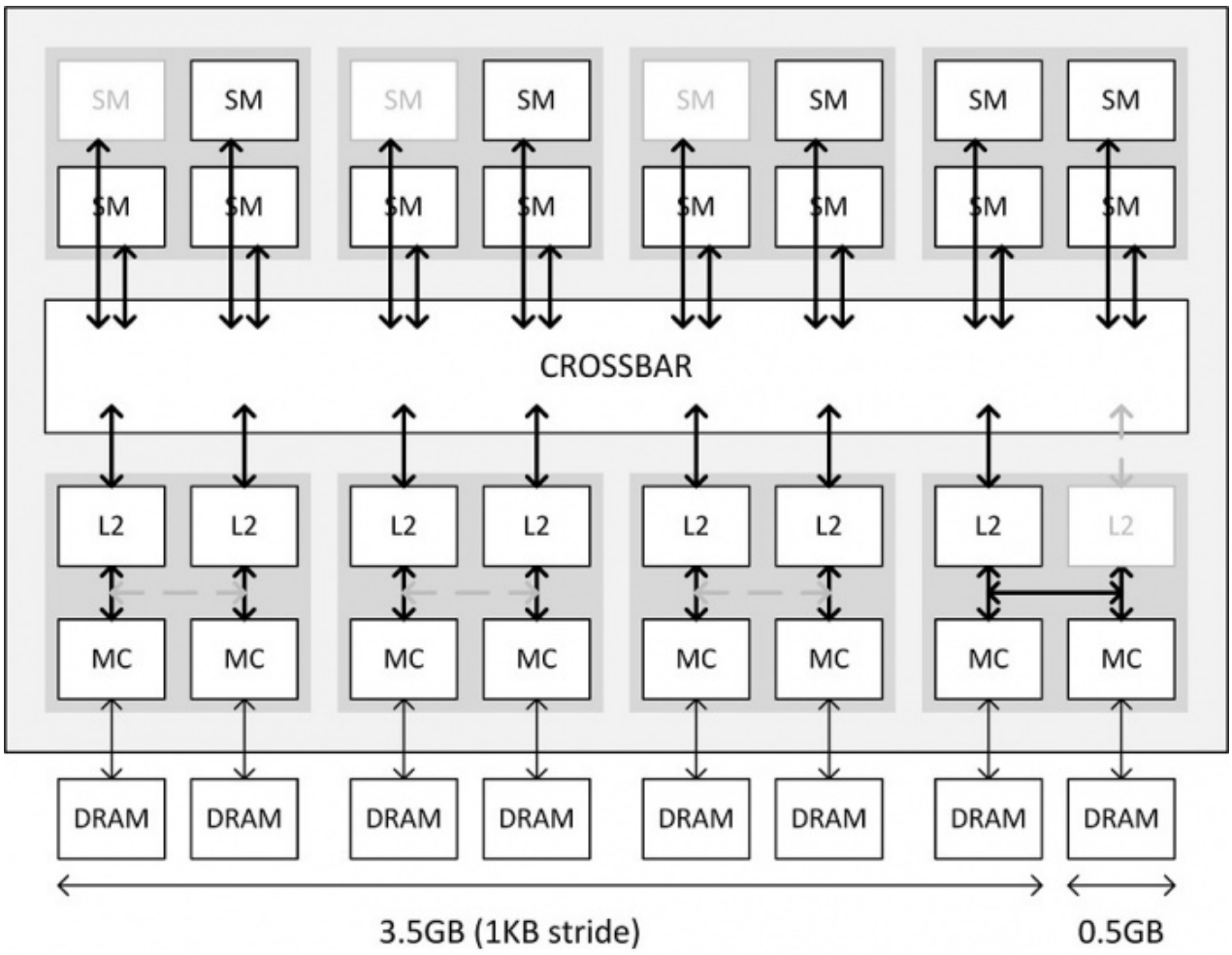
La GXT 970 era quindi una GTX 980 in cui erano state disabilitate tre unità SMM per abbassare il numero di CUDA Core a 1664.

Dal nuovo diagramma fornito da NVIDIA per la GPU GM204-200 non è proprio così e sembra che il "problema" sia da attribuire ad una "incomprensione" tra il team ingegneri che ha progettato la GPU e quello del Technical Marketing che ha redatto i documenti con le specifiche ufficiali.

Loro, come del resto noi, non pensavano infatti che con le nuove GPU Maxwell si potessero operare delle

"microsuddivisioni" a livello del silicio sugli elementi base del chip, ovvero su parti precedentemente considerate indivisibili, o che si potessero parzialmente disabilitare alcune componenti.

Il numero di ROP effettivamente attive è infatti 56 e non 64, anche se le 8 mancanti all'appello non sono realmente disabilitate, ma semplicemente non vengono utilizzate in quanto la loro connessione al controller crossbar, di cui la cache L2 è un componente fondamentale, è troppo lenta.



Dallo schema è possibile vedere le 13 unità SMM attive, il controller crossbar, i 7 canali di comunicazione con la cache L2 (che ridotta di un blocco scende quindi a 1792 KByte), gli 8 controller di memoria e gli altrettanti chip di GDDR5, con relativo partizionamento, a cui sono collegati.

Se consideriamo l'ultimo blocco, è facile capire come uno dei memory controller non disponga di una sezione propria di cache L2 a cui attingere, ma debba invece fare sempre ricorso a quella del vicino per poter accedere al controller crossbar.

La porzione di buffer video gestita da questo controller sarà quindi sempre più lenta rispetto alle altre e, per questo motivo, è stato effettuato il partizionamento.

Da qui si può quindi facilmente evincere che eventuali problemi di prestazioni possono insorgere quando un'applicazione utilizza più di 3,5 GByte di memoria; sotto tale soglia, invece, la scheda funziona al massimo delle sue capacità .

Di seguito abbiamo riportato una piccola tabella riassuntiva delle "nuove" specifiche della GTX 970.

Modelli	**GeForce GTX 980**	**GeForce GTX 970**	**GeForce GTX 970 (corrette)**
GPU	GM204-400	GM204-200	GM204-200
Processo Prod.	TSMC 28nm	TSMC 28nm	TSMC 28nm

Stream Processor	2048	1664	1664
TMUs	128	104	104
ROPs	64	64	**56**
Frequenza Base	1126MHz	1050MHz	1050MHz
GPU Boost	1216MHz	1178MHz	1178MHz
Cache L2	2048 KByte	2048 KByte	**1792 KByte**
Memoria	4GB GDDR5	4GB GDDR5	4GB GDDR5
Freq. Memoria	7.0GHz	7.0GHz	7.0GHz
Bus Memoria	256-bit	256-bit	256-bit
Consumo	~165W	~145W	~145W
Alimentazione	2 PCI-E 6pin	2 PCI-E 6pin	2 PCI-E 6pin
Uscite video	1 DVI-D 1 HDMI 1 DP	1 DVI-D 1 HDMI	1 DVI-I 1 HDMI ↔ 1DP

Qualche dubbio ci rimane, dato che l'ultimo controller dovrebbe fornire 28 GB/s per arrivarci, ovvero il massimo teorico possibile, e la stessa NVIDIA ha confermato che questa parte invece è più lenta.

E quindi? Qual'è il succo di tutto ciò?

Fondamentalmente che NVIDIA non è stata totalmente trasparente sulle specifiche della scheda.

Come venga divisa, allocata ed utilizzata la memoria alla fine dei conti non cambia le prestazioni effettive della scheda più di tanto.

Sì, ci sono delle variazioni, quantificate a quanto detto tra il 4-6% quando vengono utilizzati tutti i 4GByte di buffer video, ovvero in 4K con tutti i filtri abilitati, ma nulla di più.

E non pensiate che tutto sia da imputare alla memoria o al ridotto numero di ROP: i 13 SMM della scheda sono infatti in grado di elaborare 52 pixel/clock e le 56 ROP (sette segmenti da otto) "rimaste" sono quindi addirittura sovradimensionate (ogni ROP può trattare un pixel).

Le prestazioni delle GTX 970 quindi sono quelle che sono e, come abbiamo avuto modo di vedere davvero ottime, nonostante il numero di ROP inferiori rispetto a quanto dichiarato inizialmente e, in buona sostanza, limitate solo dal ridotto numero di SMM.

Di questa mattina una comunicazione di NVIDIA circa la possibilità di mettere una pezza alla situazione tramite il rilascio di nuovi driver ma, francamente, ci sembra si stiano arrampicando sugli specchi ...