



Descrizione tecnica

AGP 8X

Evoluzione dell'interfaccia grafica



Aumento della banda grafica

La crescente dipendenza del valore dell'esperienza multimediale dall'efficienza di elaborazione delle informazioni grafiche è un dato di dominio comune. Siti Web, giochi e quasi ogni tipo di applicazione impiegano la grafica per catturare l'attenzione degli spettatori sulle informazioni presentate in modo più immediato e intuitivo.

Le piattaforme si sono evolute per accogliere nel modo migliore questa espansione impetuosa dell'informazione visiva. Alla sua introduzione, la specifica per porte grafiche accelerate (AGP) aveva definito un percorso più ottimizzato rispetto al PCI per lo spostamento di dati tra il sottosistema grafico, l'unità di elaborazione centrale (CPU) e la memoria. Questa interfaccia continua a recitare un ruolo di primo piano. Ma la quantità e la qualità delle esigenze operative attuali contribuisce a mettere sotto pressione il bus AGP 4X:

- ❑ **Contenuto:** Le scene grafiche contengono quantità sempre maggiori di dati relativi a geometrie complesse e texture.
- ❑ **Precisione:** Gli sviluppatori stanno chiedendo a gran voce dati di maggiore precisione. La prossima generazione di unità di elaborazione grafica (GPU) di NVIDIA offre true-color a 128 bit, lo stesso livello di precisione utilizzato nelle odierne produzioni cinematografiche e la base per conseguire effetti cinematografici spettacolari in tempo reale. I formati di dati più ampi aumentano i requisiti di banda per l'intero sistema.
- ❑ **Interattività:** Le variazioni di scena in tempo reale sono tra le attività più impegnative per il bus AGP, soprattutto in contesti quali il caricamento dinamico degli ambienti di gioco, che richiedono anche lo streaming di effetti speciali attraverso il sottosistema grafico.
- ❑ **Innovazioni di sistema:** Le capacità delle piattaforme e dei sottosistemi grafici continuano a evolversi con costanti incrementi delle velocità dei processori, della capacità di memoria e della banda e funzionalità multiprocessing. La velocità del ponte che permette loro di comunicare deve quindi evolversi per consentire ai programmi software di trarre il massimo vantaggio da queste innovazioni.

La più recente release della specifica AGP — la versione 3.0 — rappresenta un evento davvero fondamentale per il settore dei PC e introduce l'interfaccia AGP 8X, in grado di rispondere in modo efficace a tutte le tendenze precedentemente esposte. L'interfaccia AGP 8X raddoppia l'ampiezza di banda del bus grafico, migliorando nettamente il throughput complessivo delle odierne applicazioni, assetate di risorse grafiche. NVIDIA sta già sfruttando le applicazioni grafiche disponibili o emergenti offrendo la prima e sola famiglia top-to-bottom del settore di GPU e prodotti core logic compatibili con AGP 8X.

Questo documento descrive i vantaggi delle più recenti innovazioni della specifica AGP e illustra nei dettagli i programmi di NVIDIA in merito all'adozione di questo nuovo standard.

La specifica AGP

Lo standard AGP è nato per definire interconnessioni ad elevate prestazioni in grado di ottimizzare la grafica 3D. La porta dedicata ad alta velocità connette il chipset di core-logic e il controller grafico, creando un percorso diretto per il trasferimento di texture grafiche all'interno e all'esterno della memoria quando lo spazio del buffer frame locale viene superato o si deve caricare una nuova scena. Questo design offre diversi vantaggi:

- ❑ L'interfaccia AGP può trasferire dati alla velocità di gigabyte per secondo (GB/s.), che supera quella del bus PCI (2,1 GB/s. invece di 132 MB/s., il massimo del bus PCI), quindi può supportare l'esecuzione di mappe di texture dalla memoria di sistema piuttosto che obbligare al precaricamento di tutti i dati delle texture nella memoria grafica locale.
- ❑ La specifica include una modalità di indirizzamento della banda laterale che permette alla GPU di rispondere a nuovi indirizzi e richieste prima di terminare la richiesta precedente.
- ❑ Il bus PCI viene sollevato da numerose delle richieste che lo congestionavano, massimizzando le prestazioni dei dispositivi limitati a quel bus (controller di disco, schede LAN, sistemi di cattura video, ecc.).

La release AGP 3.0

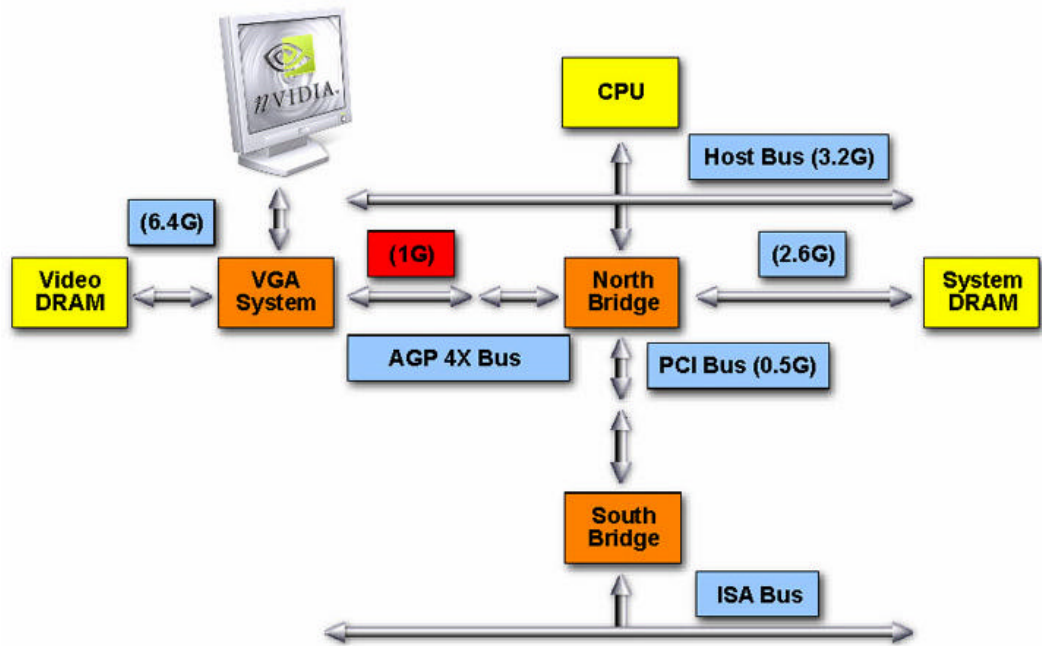
Sin dalla sua introduzione nel 1996, l'interfaccia AGP è stata costantemente aggiornata, in modo evolutivo. Originariamente, la specifica è nata per risolvere i problemi creati dal diffusissimo bus PCI e per definire un'interfaccia ideale per le esigenze delle operazioni grafiche e del movimento di dati. Le revisioni della specifica si sono concentrate sulla scalatura dell'ampiezza di banda:

- ❑ I livelli di banda AGP 1X e AGP 2X sono stati introdotti simultaneamente nella versione 1.0 della specifica di interfaccia AGP 1.0 (o AGP 1.0). La specifica AGP 1.0 consentiva queste due velocità di interfaccia; la AGP 2X era teoricamente due volte più rapida della AGP 1X. (Vedere la tabella 1 nella sezione delle prestazioni per ulteriori dettagli sulle varie velocità di trasferimento).
- ❑ La AGP 4X, definita nella specifica AGP 2.0, è stata introdotta due anni dopo.
- ❑ Oggi, la AGP 3.0 offre al settore l'ampiezza di banda AGP 8X e introduce il funzionamento isocrono e le capacità di texturing AGP (descritte in maggiore dettaglio nelle sezioni seguenti).

La versione 3.0 dell'interfaccia AGP (o AGP 3.0) raddoppia le prestazioni teoriche del bus. Questa ultimissima versione della specifica incorpora anche alcune nuove funzioni e rimuove alcune caratteristiche inutilizzate per semplificare l'interfaccia. NVIDIA supporta AGP 3.0 sin dalle offerte di GPU e core logic programmate per la distribuzione al pubblico ad autunno 2002.

Ottimizzazione delle operazioni grafiche e dell'archiviazione delle texture

L'architettura dei sistemi e dei sottosistemi grafici in particolare è decisamente progredita dal momento della presentazione dell'interfaccia AGP 2.0 nel lontano 1998. Oggi, AGP 2.0 e la sua banda associata AGP 4X costituiscono un collo di bottiglia per il flusso complessivo dei dati correlati alla grafica (vedere la figura 1).



CPU = CPU
 Host bus = Bus host
 Video DRAM = DRAM video
 VGA System = Sistema VGA
 North Bridge = Ponte Nord
 System DRAM = DRAM di sistema
 AGP 4X Bus = Bus AGP 4X
 PCI Bus = Bus PCI
 South Bridge = Ponte Sud

Figura 1. La banda AGP 4X è inadeguata alle prestazioni del resto del sistema. Grazie alla AGP 8X, la banda grafica è all'altezza delle velocità di throughput del sistema e della memoria.

Con una banda teorica di 2,1 GB/s., AGP 8X permette inoltre agli sviluppatori di gestire con maggiore efficienza scene con geometrie complesse e di passare in modo dinamico a nuove scene in tempo reale. Riducendo l'attività complessiva associata all'archiviazione e recupero di texture complesse o allo streaming di dati dalla memoria, AGP 8X aumenta il throughput totale del bus AGP. Questo, di conseguenza, offre la possibilità di migliorare nettamente le prestazioni per le visualizzazioni che prevedono texture e geometrie di elevata complessità, e la creazione di esperienze maggiormente realistiche per gli utenti finali.

Funzionamento isocrono per lo streaming

I miglioramenti complessivi della banda offerti dall'interfaccia AGP 3.0 sono disponibili a tutte le operazioni grafiche che prevedano geometrie e texture complesse, oppure streaming. Una nuova funzionalità della AGP 3.0 — la modalità

di funzionamento isocrona — migliora in modo specifico le operazioni grafiche che richiedono un flusso dei dati prevedibile e ininterrotto. Le versioni precedenti dell'interfaccia AGP non possono garantire la banda richiesta per i trasferimenti sensibili alla latenza. Questa architettura “best effort” funziona in modo eccellente per ottenere una bassa latenza media e un elevato throughput medio, ma non protegge da occasionali ritardi di lunghezza abnorme, e pertanto può dare luogo alla perdita di dati. Le applicazioni in streaming — applicazioni che prevedono flussi in tempo reale di informazioni digitali per la trasmissione di video, i download da reti e altre operazioni simili — non possono tollerare la perdita di dati e richiedono trasferimenti prevedibili. Inoltre, implementazioni a basso costo richiedono un metodo per supportare i trasferimenti isocroni senza incrementare le quantità richieste in termini di data buffering, un'attività decisamente costosa. Grazie alla modalità di funzionamento isocrona, l'interfaccia AGP 3.0 risponde a questi requisiti applicativi in modo decisamente conveniente.

Compatibilità

La specifica AGP 3.0 offre un percorso di aggiornamento agevole alla AGP 8X. Le specifiche meccaniche del bus rimangono le stesse. La velocità e le capacità della AGP 8X vengono ottenute sfruttando alcuni pin precedentemente inutilizzati, il tutto in modo da semplificare il supporto delle schede AGP 8X nei sistemi AGP 2X e 4X esistenti, e non solo per i nuovi sistemi che possono sfruttare appieno l'interfaccia 8X. Le soluzioni grafiche AGP 8X di NVIDIA saranno in grado di rilevare il livello di AGP del sistema host e di configurare automaticamente l'interfaccia AGP per l'esecuzione in modalità 3.0 (a velocità 4X o 8X), o in modalità 2.0 (a velocità 2X o 4X). Pertanto, una nuova soluzione grafica NVIDIA sarà pienamente in grado di gestire le velocità 8X e conserverà la massima compatibilità con i sistemi 2X, 4X e 8X. Le schede basate su NVIDIA offrono automaticamente la velocità massima supportata dal sistema host.

Prestazioni AGP 8X

La banda AGP 8X è il doppio della AGP 4X. L'impatto della banda AGP 8X sulle prestazioni complessive delle applicazioni varia con il tipo di applicazioni:

- ❑ **Mondi statici:** Applicazioni eseguite in un ambiente virtuale di dimensioni ridotte — in essenza, l'intero “mondo” viene caricato nella memoria grafica in ogni occasione — vedranno rari, o nulli, miglioramenti nelle prestazioni a causa della nuova velocità di trasferimento AGP 8X.
- ❑ **Mondi complessi:** Gli odierni programmi e giochi “fly-through” ad alta velocità trarranno invece benefici significativi in termini di prestazioni complessive a causa del raddoppio della velocità della scheda AGP 8X. Queste applicazioni e giochi devono prevedere e caricare geometrie e texture nel frame buffer e si avvantaggiano del miglioramento del throughput tra la memoria principale e il sottosistema grafico. Le applicazioni con dati ad alta precisione e texture di grandi dimensioni ne trarranno anch'esse beneficio, dato che richiedono un numero maggiore di trasferimenti dalla memoria principale.

Tabella 1. Confronto fra AGP 4X e AGP 8X.

	AGP 4X	AGP 8X
Byte per trasferimento	4 (32 bit)	4 (32 bit)
Velocità di clock	266,67MHz	533,33MHz
Banda del bus	1,1GB/s.	2,1GB/s.

Conclusione

Le applicazioni grafiche mettono a dura prova tutte le componenti del sistema e richiedono costanti innovazioni per mantenere un ambiente bilanciato ed evitare colli di bottiglia. AGP 3.0, offre la banda AGP 8X e permette una vera e propria rivoluzione in termini di ampiezza di banda. L'aumento della banda e il perfezionamento della progettazione del bus completano l'hardware grafico emergente e permettono l'impiego di metodi più pratici e creativi per utilizzare le risorse di memoria di sistema al fine di supportare texture e visualizzazioni complesse. Pur conservando la compatibilità con i sistemi AGP 2X e 4X, le nuove soluzioni AGP 8X possono supportare un percorso semplificato a:

- Grafica cinematografica in tempo reale e un migliore uso della memoria principale per geometrie complesse, texture e dati di maggiore precisione.
- Applicazioni a prestazioni più elevate che prevedono caricamento dinamico dei "mondi" o streaming.
- Supporto per i trasferimenti isocroni che caratterizzano numerose applicazioni di media digitali quali lo streaming da dispositivi e reti digitali.
- Prestazioni di sistema bilanciate nelle quali le operazioni grafiche e il flusso dei dati sono scaricati in modo intelligente dalla CPU.

La AGP 8X offre miglioramenti immediati delle prestazioni per numerosi giochi e applicazioni che impiegano texture e scene di elevata complessità. Questa nuovissima release dell'interfaccia AGP alleggerisce la pressione in termini di esigenze di banda cui è sottoposto il sottosistema grafico e offre lo spazio necessario a ottimizzare la gestione delle applicazioni per alcuni anni.

Le nuove GPU e piattaforme PC di NVIDIA incorporeranno capacità AGP 8X e trarranno il massimo vantaggio dalle specifiche AGP 3.0. Come sempre, le soluzioni NVIDIA incorporano nuove tecnologie e avanzamenti senza compromettere la stabilità complessiva e la qualità del sistema, e la NVIDIA Unified Driver Architecture (UDA) semplifica lo sfruttamento delle nuove capacità non appena vengono introdotte.



Nota:

TUTTE LE SPECIFICHE DI PROGETTAZIONE NVIDIA, LE SCHEDE DI RIFERIMENTO, I FILE, I DISEGNI, LA DIAGNOSTICA, LE LISTE E ALTRI DOCUMENTI (UNITAMENTE E SEPARATAMENTE, DEFINITI "MATERIALI") SONO FORNITI NELLO STATO IN CUI SI TROVANO. NVIDIA NON OFFRE GARANZIE, ESPRESSE, IMPLICITE, STATUTARIE O DI ALTRO TIPO IN RELAZIONE AI MATERIALI, E RIFIUTA ESPRESSAMENTE OGNI GARANZIA IMPLICITA DI NON VIOLAZIONE, COMMERCIALIZZABILITÀ E IDONEITÀ A SCOPI SPECIFICI.

Le informazioni fornite sono ritenute accurate e affidabili. Tuttavia, NVIDIA Corporation non si assume alcuna responsabilità per le eventuali conseguenze derivanti dall'uso di tali informazioni o da qualsiasi violazione di brevetti o altri diritti di terze parti che possono conseguire dal loro uso. Non viene concessa alcuna licenza implicita o in altro modo in base a nessun brevetto o diritto di autore di proprietà di NVIDIA Corporation. Le specifiche tecniche menzionate nella presente pubblicazione sono soggette a modifica senza preavviso. Questa pubblicazione rimpiazza e sostituisce tutte le informazioni precedentemente fornite. Non si autorizza l'impiego dei prodotti di NVIDIA Corporation come componenti cruciali di dispositivi per il supporto vitale o per sistemi che non abbiano ricevuto l'espressa approvazione scritta di NVIDIA Corporation.

NVIDIA e il logo NVIDIA sono marchi registrati e GeForce2 Go è un marchio di NVIDIA Corporation.

Altri nomi di società e di prodotti possono essere marchi o marchi registrati dei rispettivi detentori.

Copyright NVIDIA Corporation 2002



NVIDIA.

NVIDIA Corporation
2701 San Tomas Expressway
Santa Clara, CA 95050
www.nvidia.com