



Condensé technique

AGP 8X

L'évolution de l'interface
graphique



Une bande passante grandissante

Nul ne le niera, la satisfaction qu'un utilisateur tire de son P.C. dépend toujours plus de la capacité de ce dernier à traiter les informations graphiques. À l'ère de l'information, les sites Internet, les jeux et pratiquement tous les types d'applications emploient des graphiques pour fournir des vues plus intuitives, plus convaincantes au premier coup d'œil.

Les plates-formes ont évolué pour mieux gérer cette explosion d'informations visuelles. À son introduction, la spécification AGP (Accelerated Graphics Port) définissait un bus mieux optimisé que le bus PCI pour le transport des données entre le sous-système graphique, l'unité centrale (UC) et la mémoire. Si cette interface continue à jouer un rôle capital, les contraintes qui pèsent aujourd'hui sur le bus AGP 4X sont toujours plus nombreuses :

- ❑ **Contenu** : les scènes graphiques contiennent des quantités croissantes de géométries complexes et de données de texture.
- ❑ **Précision** : les développeurs réclament à cor et à cris des données plus précises. La prochaine génération de GPU (processeurs graphiques) NVIDIA offre de vraies couleurs 128 bits, soit la même précision que celle utilisée aujourd'hui dans l'industrie cinématographique et la base de la réalisation d'effets cinématiques spectaculaires en temps réel. Les formats de données plus importants augmentent la bande passante requise à travers le système.
- ❑ **Interactivité** : les changements de scène en temps réel pèsent sur le bus AGP tout comme le chargement dynamique d'environnements de jeu ou le streaming d'effets spéciaux par le biais du sous-système graphique.
- ❑ **Avancées des systèmes** : les capacités des plates-formes hôtes et des sous-systèmes graphiques continuent à évoluer, la vitesse des processeurs, la mémoire et la bande passante et les capacités multitraitement augmentent. La vitesse du pont qui les relie doit évoluer pour que les logiciels puissent exploiter ces avancées.

Marquant une étape déterminante dans l'industrie des P.C, la dernière version de la spécification AGP (la 3.0) introduit l'interface AGP 8X afin de répondre à toutes ces exigences. AGP 8X double la bande passante du bus graphique, ce qui se traduit par une augmentation considérable du débit global pour les applications assoiffées de graphiques d'aujourd'hui. En offrant la première et la seule famille complète de GPU et de produits « core logic » compatibles AGP 8X, NVIDIA vous permet d'exploiter pleinement les applications graphiques disponibles et à venir.

Ce document décrit les avantages des dernières améliorations AGP et détaille les plans de NVIDIA pour l'adoption de ce nouveau standard.

La spécification AGP

La norme AGP a été conçue dans l'objectif de définir des interconnexions haute performance pour améliorer la performance graphique 3D. Le port haute vitesse dédié connecte le chipset « core logic » et le contrôleur graphique, en créant un chemin direct pour le transfert de textures graphiques dans et hors de la mémoire en cas de dépassement de l'espace de la mémoire d'écrans locale ou lorsqu'une nouvelle scène doit être chargée. Cette conception offre de nombreux avantages :

- ❑ AGP peut transférer des données de texture à des taux de l'ordre de plusieurs giga-octets par seconde (Go/s) qui laissent loin derrière le bus PCI (2,1 Go/s contre 132 Mo/s pour PCI), et prendre en charge l'exécution de mappes de textures provenant de la mémoire système au lieu d'imposer le préchargement de toutes les données de texture dans la mémoire graphique locale.
- ❑ La spécification inclut un mode d'adressage des bandes latérales qui permet au GPU d'émettre de nouvelles adresses et requêtes sans attendre la fin de la requête précédente.
- ❑ Le bus PCI est du coup bien moins encombré ce qui permet de maximiser la performance des périphériques confinés à ce bus (contrôleurs de disque, cartes LAN, systèmes de capture vidéo, etc.).

La version AGP 3.0

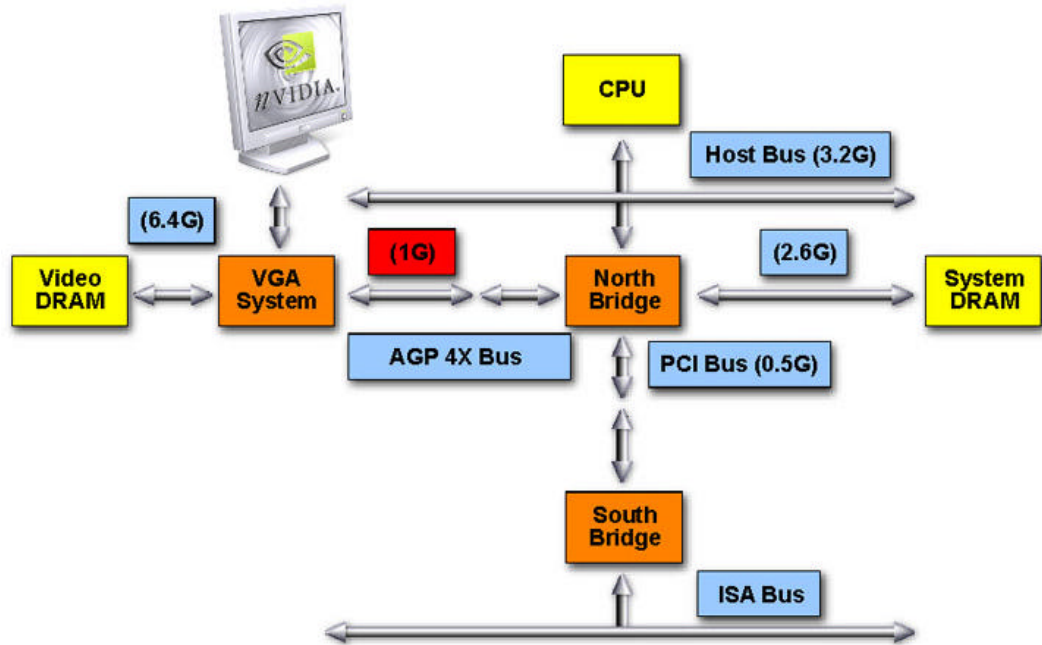
Depuis son introduction en 1996, l'interface AGP a été mise à jour selon une logique évolutive. Si à l'origine, cette norme est née pour compenser les insuffisances du bus PCI et définir une interface sur mesure pour les opérations graphiques et le mouvement de données, les révisions plus récentes ont mis l'accent sur le dimensionnement de la bande passante :

- ❑ Les niveaux de bande passante AGP 1X et AGP 2X ont été introduits simultanément dans la version 1.0 de la spécification AGP 1.0 (ou AGP 1.0). AGP 1.0 autorisait deux vitesses d'interface, AGP 2X étant en théorie deux fois plus rapide qu'AGP 1X (voir le Tableau 1 de la partie consacrée aux performances pour le détail des divers modes de transfert).
- ❑ AGP 4X, défini dans la spécification AGP 2.0, a été introduit deux ans plus tard.
- ❑ Aujourd'hui, AGP 3.0 met la bande passante AGP 8X au service de l'industrie et introduit le fonctionnement isochrone et des capacités de texturation AGP (décrites en détail dans les parties qui suivent).

La version 3.0 de l'interface AGP (ou AGP 3.0) double la performance théorique au travers du bus. Cette dernière version de la spécification, outre incorporer de nouvelles fonctionnalités, en a éliminé d'autres jugées inutiles en vue de simplifier l'interface. Chez NVIDIA, la prise en charge d'AGP 3.0 deviendra réalité dans les produits GPU et « core logic » dont la sortie est programmée pour l'automne 2002.

Des opérations graphiques et un stockage des textures optimisés

La conception des systèmes et des sous-systèmes graphiques en particulier a avancé à un rythme soutenu depuis l'introduction d'AGP 2.0 en 1998. Aujourd'hui AGP 2.0 et sa bande passante AGP 4X sont devenus à leur tour un goulot d'étranglement dans le flux général des données relatives à la carte graphique (voir Figure 1).



CPU = UC
 Host Bus = Bus hôte
 Video DRAM = DRAM vidéo
 VGA System = Système VGA
 AGP 4X Bus = Bus AGP 4X
 PCI Bus = Bus PCI
 System DRAM = DRAM système
 ISA Bus = Bus ISA

Figure 1. La bande passante AGP 4X ne répond plus aux besoins du reste du système. Avec AGP 8X, la bande passante graphique sera équilibrée avec le débit du reste du système et celui de la mémoire.

Avec une bande passante théorique de 2,1 Go/s, AGP 8X permet aux développeurs de gérer efficacement les scènes aux géométries complexes et de passer de façon dynamique et en temps réel à de nouvelles scènes. En abaissant la charge associée au stockage et à l'extraction de textures complexes ou au streaming de données provenant de la mémoire, AGP 8X augmente le débit global pour le bus AGP. Cette hausse se traduit à son tour par de nettes améliorations de performance pour les visualisations qui impliquent des textures et une géométrie complexe, et ne pourra qu'inspirer des expériences plus réalistes pour l'utilisateur.

Un fonctionnement isochrone pour le streaming

Les améliorations de bande passante générales d'AGP 3.0 bénéficient à toutes les opérations graphiques comportant des géométries complexes, des textures ou du streaming. L'une des nouvelles fonctions d'AGP 3.0, le mode isochrone, améliore spécifiquement les opérations graphiques qui requièrent un flux de données ininterrompu et prévisible. Les versions précédentes de l'interface AGP ne

pouvaient pas garantir la bande passante requise pour les transferts dépendant du temps d'attente. Cette solution « au mieux » fonctionne bien pour obtenir un temps d'attente moyen réduit et un débit moyen élevé, mais ne protège pas contre d'occasionnels et arbitraires délais longs et peut donc se traduire par des pertes de données. Or, les applications de streaming (c'est à-dire les applications qui impliquent des flux en temps réel d'informations numériques pour la diffusion de vidéos, les téléchargements ou d'autres tâches similaires) ne tolèrent pas les pertes de données et exigent des transferts prévisibles. En sus, les configurations bon marché doivent pouvoir prendre en charge les transferts isochrones sans augmenter le coûteux volume des données mises en tampon. En fonctionnant en mode isochrone, AGP 3.0 satisfait les exigences de ces applications de façon économique.

La compatibilité

La spécification AGP 3.0 fait de la mise à jour vers AGP 8X une promenade. La spécification mécanique du bus reste la même. Les vitesses et les capacités d'AGP sont, en effet, atteintes en exploitant certaines broches autrefois inutilisées et ce, d'une manière qui facilite la prise en charge des cartes AGP 8X dans les systèmes AGP 2X et 4X existants, comme dans les nouveaux systèmes qui reposent entièrement sur l'interface 8X. Les solutions graphiques NVIDIA AGP 8X seront en mesure de détecter le niveau AGP du système hôte, et de configurer automatiquement l'interface AGP pour fonctionner en mode 3.0 (vitesse 4X ou 8X), ou en mode 2.0 (vitesse 2X ou 4X). Ainsi, une nouvelle solution graphique NVIDIA pourra atteindre pleinement des vitesses 8X tout en étant entièrement compatible avec des systèmes 2X, 4X et 8X. Les cartes basées sur NVIDIA fourniront automatiquement la vitesse maximale prise en charge par le système hôte.

Les performances d'AGP 8X

La bande passante AGP 8X est le double de la bande AGP 4X. L'impact de la bande passante AGP 8X sur la performance applicative globale dépendra du type des applications :

- ❑ **Mondes statiques** : les applications qui tournent au sein d'un environnement virtuel restreint — autrement dit, celles dont le « monde » entier est chargé à tout moment dans la mémoire graphique — verront peu, voire pas, d'améliorations de performance dues au taux de transfert AGP 8X.
- ❑ **Mondes complexes** : les applications et jeux « fly-through » aux cadences soutenues d'aujourd'hui enregistreront de nettes améliorations de performance grâce à la vitesse doublée d'AGP 8X. Ces applications et jeux doivent en effet prévoir et charger géométries et textures dans la mémoire d'écrans et bénéficient donc du débit amélioré entre la mémoire principale et le sous-système graphique. Les applications comportant des données haute précision et des textures importantes en tireront aussi avantage puisqu'elles requièrent aussi davantage de transfert de et vers la mémoire principale.

Tableau 1. Comparaison d'AGP 4X et AGP 8X

	AGP 4X	AGP 8X
Octets par transfert	4 (32 bits)	4 (32 bits)
Fréquence d'horloge	266,67 MHz	533,33 MHz
Bande passante du bus	1,1 Go/s	2,1 Go/s

Conclusion

Les applications graphiques pèsent sur tous les éléments d'un système et des avancées constantes s'imposent pour assurer un environnement équilibré et éviter les goulots d'étranglement. AGP 3.0, en offrant une bande passante AGP 8X, fait un pas révolutionnaire vers la bande passante continue. Cette bande passante accrue et la nouvelle conception du bus complètent le matériel graphique émergent et ne pourront qu'inspirer des façons plus créatives et plus pratiques d'utiliser les ressources de mémoire des systèmes pour prendre en charge des textures et de visualisations complexes. Tout en maintenant la compatibilité avec les systèmes AGP 2X et 4X existants, les nouvelles solutions AGP 8X sauront tranquillement évoluer vers :

- ❑ Les graphiques cinématiques en temps réel et une meilleure utilisation de la mémoire principale pour les géométries complexes, les textures et les données de plus haute précision .
- ❑ Les applications aux performances accrues qui impliquent le chargement de « mondes » dynamiques ou du streaming.
- ❑ La prise en charge de transferts isochrones qui caractérisent de nombreuses applications média numériques telles que le streaming de vidéos provenant d'appareils numériques et de réseaux.
- ❑ Des performances système équilibrées où l'UC se verra libérée des opérations et des flux de données graphiques.

AGP 8X fournit des améliorations de performance immédiates pour les nombreux jeux et applications qui emploient des textures et des scènes complexes. Cette nouvelle version de l'interface AGP élimine les contraintes liées à la bande passante du sous-système graphique et laisse une marge de manœuvre qui sera nécessaire pour gérer les applications des années à venir.

Les nouveaux GPU et plates-formes P.C. de NVIDIA intégreront les capacités AGP 8X et tireront pleinement parti de la spécification AGP 3.0. Comme toujours, les solutions NVIDIA incorporent nouvelles technologies et éléments de pointe sans compromettre la stabilité et la qualité générales du système, et l'architecture de pilote unifiée de NVIDIA (UDA) simplifie l'exploitation de ces nouvelles capacités dès leur apparition.



Avis

L'ENSEMBLE DES SPÉCIFICATIONS DE CONCEPTION, CARTES DE RÉFÉRENCE, FICHIERS, DESSINS, DIAGNOSTICS, LISTES ET AUTRES DOCUMENTS DE NVIDIA (ENSEMBLE ET SÉPARÉMENT, LES "MATÉRIAUX") SONT FOURNIS "TELS QUELS". NVIDIA N'ACCORDE AUCUNE GARANTIE EXPRESSE, IMPLICITE, STATUTAIRE OU AUTRE EN CE QUI CONCERNE LES MATÉRIAUX ET DÉCLINE EXPRESSÉMENT TOUTE GARANTIE IMPLICITE DE NON-INFRACTION AUX DROITS DE PROPRIÉTÉ INTELLECTUELLE, DE VALEUR MARCHANDE ET D'ADAPTABILITÉ À TOUT OBJECTIF PARTICULIER.

Les informations fournies sont réputées précises et fiables. Toutefois, NVIDIA Corporation décline toute responsabilité quant aux conséquences de l'utilisation qui pourrait en être faite ou de la contrefaçon de brevets ou autres droits de tierces parties pouvant résulter de leur utilisation. Aucune licence n'est octroyée implicitement ou de quelque autre manière sous quelque brevet ou droit de brevet de NVIDIA Corporation. Les caractéristiques techniques mentionnées dans ce document peuvent être modifiées sans préavis. Cette publication annule et remplace toute information diffusée antérieurement. Les produits de NVIDIA Corporation ne peuvent en aucun cas être utilisés en tant que composants critiques pour des systèmes de survie sans l'accord préalable écrit de NVIDIA Corporation.

NVIDIA et le logo NVIDIA sont des marques déposées et GeForce2 Go est une marque commerciale de NVIDIA Corporation.

Les autres noms de sociétés ou produits peuvent être des marques commerciales des sociétés auxquelles ils sont associés.

Copyright NVIDIA Corporation 2002



NVIDIA.

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