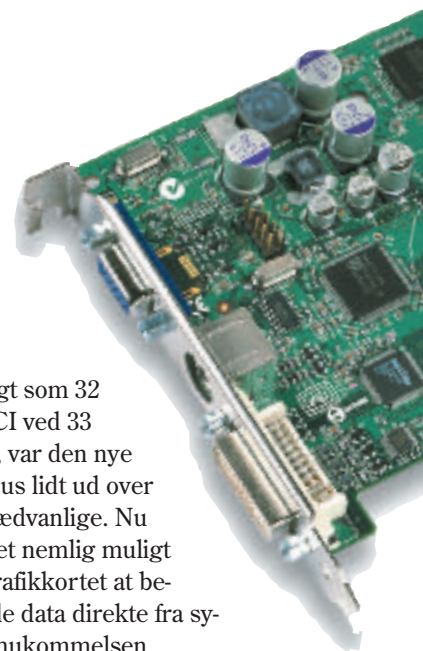


AGP 2, 4 eller 8X:

En hurtigere bus

AGP er det nyeste inden for databusser til grafikkort. Det har afløst de bedagede PCI- og ISA-standarder. Vi ser her nærmere på den reelle hastighedsforskel mellem de forskellige typer af AGP



Af Dan Bohmann

Markedet er på det seneste blevet oversvømmet med nye AGP 8X-grafikkort, men kan det nu også betale sig at opgradere sit grafikkort, blot for at få understøttelse af AGP 8X?

I første omgang skal man jo sikre sig, at det bundkort, man har, understøtter AGP 8X. Gør det ikke det, er der jo ingen grund til at investere i et AGP 8X-grafikkort.

Har man der imod et bundkort, som understøtter AGP 8X, men kun et AGP 4X-grafikkort monteret, er det måske en ide at opgradere. Eller er det?

En række af de nye AGP 8X-grafikkort er nemlig baseret på præcis den samme gpu (f.eks. GF4 M460) som deres brødre med AGP 4X, og da AGP 4X kan levere den båndbredde, der er krævet, for at kortet yder sit bedste, vil der altså ikke være nogen ide i at opgradere i dette tilfælde.

Lidt historie

Helt tilbage i computerens spæde ungdom brugte man ISA-slottet til at montere grafikkortet i. ISA-kortet kan ved en frekvens på 4,77 MHz og en båndbredde på bare 8 bit overføre 2,39 MB/s. Denne

datamængde var rigeligt til datidens applikationer, for som de fleste ved, indeholdt disse ikke noget nævneværdig grafik. I 90'erne kom VESA local bus så til. Det var en databus specielt udviklet til grafikkort, præcis som AGP er det i dag. VL-bussen overfører data med 133,33 MB/s.

Hurtigt efter kom der PCI på banen. Ligesom VL-bussen kunne PCI-bussen overføre 133,33 MB/s, men til forskel fra VL-bussen blev PCI-slottet også brugt til andre indstikskort. I starten var det SCSI-controllere, men senere blev det også brugt til netværkskort, lyd kort og meget mere.

Da grafikkortene begyndte at løbe hovedet mod muren ved 133 MB/s, måtte der selvfølgelig gøres noget. Det var et tidspunkt, hvor der skete mange skift inden for grafikkort, og da hastigheden hele tiden skulle øges, var man nødsaget til at introducere AGP.

AGP

Den første AGP-standard hed i sigens natur AGP 1.0 og understøttede både AGP 1X og AGP 2X. Ud over AGP 1X-hastigheden på 266 MB/s, som jo er dobbelt så

hurtigt som 32 bit PCI ved 33 MHz, var den nye databus lidt ud over det sædvanlige. Nu var det nemlig muligt for grafikkortet at behandle data direkte fra systemhukommelsen.

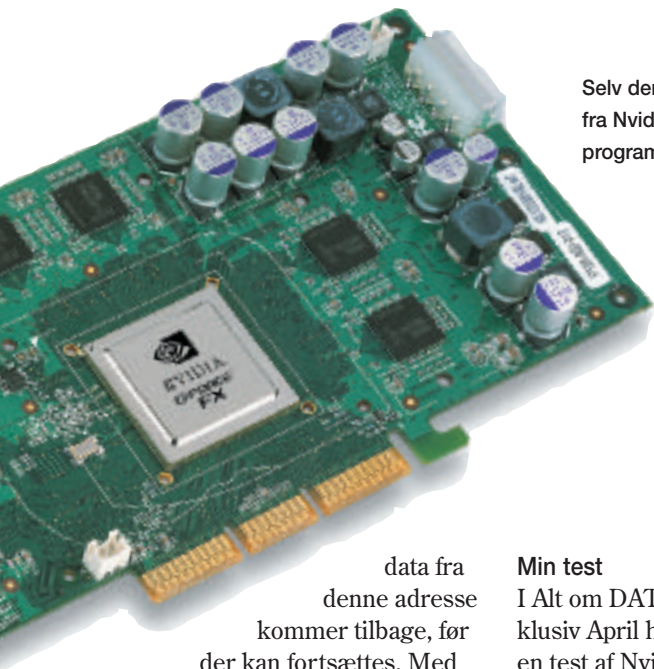
Teknikken hedder direct executing og bør ikke forveksles med DMA, som PCI-kort også understøtter. Direct executing betød, at skulle man køre et program, som benyttede en texture på 16 MB, men var der kun 8 MB ram på ens grafikkort, så kunne grafikkortet »hente« de sidste 8 MB direkte i systemhukommelsen og var altså ikke nødsaget til at hente hele texturen ind i den lokale hukommelse på grafikkortet, før den kunne bearbejdes. Det betød igen, at trafikken til og fra den lokale hukommelse på grafikkortet minimeres, og derfor er båndbredden også til rådighed, når den virkelig skal bruges.

En anden teknik kaldet pipelining betød også, at AGP var PCI overlegent. Nå en adresse sendes over PCI-arkitekturen, er databussen nemlig nødt til at vente på, at

Figur 1. Nvidia viser os, at AGP 8X ingen betydning har for dagens gpu'er. Se f.eks. Ti4200 med og uden AGP 8X, og sammenlign også Ti4600, som har AGP 4X, med Ti4800, som har AGP 8X.

| | Vertex (millioner) | Fill rate AA samples pr. sekund (milliarder) | Beregninger pr. sekund (billioner) | Båndbredde (GB/sek.) | Hukommelse |
|-----------------------------|--------------------|--|------------------------------------|----------------------|------------|
| GeForce4 Ti 4800* | 136 | 4,8 | 1,23 | 10,4 | 128 MB |
| GeForce4 Ti 4800 SE* | 125 | 4,4 | 1,12 | 8,8 | 128 MB |
| GeForce4 Ti 4200 m. AGP 8X* | 113 | 4 | 1,03 | 8 | 128 MB |
| GeForce4 Ti 4600 | 136 | 4,8 | 1,23 | 10,4 | 128 MB |
| GeForce4 Ti 4400 | 125 | 4,4 | 1,12 | 8,8 | 128 MB |
| GeForce4 Ti 4200 | 113 | 4 | 1,03 | op til 8 | 128 MB |

De med * markerede kort er AGP 8X-kompatible



Selv den hurtigste GeForce FX gpu fra Nvidia kan med de nuværende programmer ikke udnytte AGP 8X.

data fra denne adresse kommer tilbage, før der kan fortsættes. Med AGP er det muligt at have 35 adresser kørende på databussen på samme tid, og systemet er derfor ikke nødsaget til at vente på, at data skal modtages, før der fortsættes.

Virkeligheden

Som mange ved, er der siden hen kommet AGP 2X, AGP 4X og sidst, men ikke mindst, AGP 8X. Men kan det betale sig at skifte til AGP 8X? For når det kommer til stykket, handler det om, hvor hurtig gpu'en på grafikkortet er, og hvor mange data der er behov for at hente i systemhukommelsen eller fra processoren.

Et godt eksempel på, at AGP 8X absolut ingen effekt har, hvis gpu'ens krav til båndbredde ikke overstiger det, som AGP 4X kan levere, kommer fra Nvidia selv. Se figur 1.

Nvidia skriver nemlig, at et GeForce4 Ti 4200 med AGP 8X yder præcis det samme som et GeForce4 Ti4200 med AGP 4X. Sjovt er det også at se, at Geforce4 Ti 4600, som er et AGP 4X-kort, yder præcis lige så godt som den nyere model GeForce4 Ti 4800.

Min test

I Alt om DATA Eksklusiv April havde vi en test af Nvidias ny GeForce FX-kort.

Kortet viste sig ikke helt at leve op til forventningerne og kunne ikke engang vippe ATIs Radeon 9700 pro af pinden. Ikke desto mindre har jeg kørt en lille test på GeForce FX-kortet for at se, om jeg kan måle nogen forskel, når kortet kører i AGP 4X eller AGP 8X.

Jeg kørte de samme testprogrammer, som jeg benyttede i sidste grafikkorttest, nemlig 3D Mark 2001 SE, 3DMark 2003, AquaMark og GLMark.

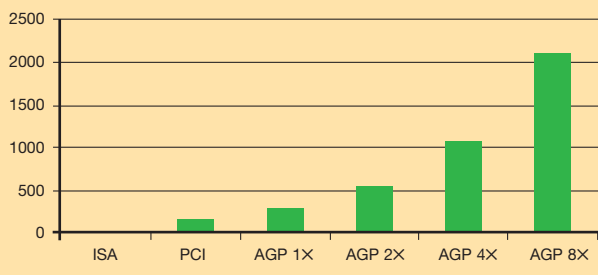
Resultatet? Kortene var så tæt på hinanden, at jeg må afskrive de meget små afvigelser som måleusikkerhed. Det var simpelt hen ikke muligt for mig at se eller måle nogen forskel, uanset om kortet kørte i AGP 4X eller AGP 8X.

Fremtiden

Nu skal vi huske på, at AGP 8X kan overføre 2,1 GB/s mellem chipsættet og grafikkortet, mens AGP 4X kun kan overføre det halve, i hvert fald rent teoretisk.

Det vigtigste man skal se på, når man skal købe et grafikkort, er altså ikke, om det er et AGP 4X- eller et AGP 8X-grafikkort.

Overførselshastighed for databusser (MB/s)



For tiden er det langt vigtigere at se på overførselshastigheden mellem den lokale hukommelse på grafikkortet og selve gpu'en. Og så har gpu'ens egen hastighed naturligvis også stor betydning.

Jo hurtigere en gpu er, jo større behov har den for at kunne hente data hurtigt. Det primære sted, en gpu skal hente data, er i den lokale hukommelse på selve grafikkortet. Derfor skal overførslen mellem gpu og lokal hukommelse være så hurtig som muligt. Her vil vi i den nærmeste fremtid se, at DDRII vil blive standard på grafikkortene, og med forskellige optimeringstricks vil overførslen mellem GPU og lokal grafikhukommelse komme op i nærheden af 20 GB/s.

Der er ingen tvivl om, at AGP 8X vil blive standard på nyere grafikkort, og efterhånden som de forskellige spil bliver programmeret til at udnytte den ekstra båndbredde mellem grafikkortet og systemhukommelsen, vil vi se, at ydelsesforskellen mellem AGP 4X- og AGP 8X-kort vil stige betydeligt.

Det ses tydeligt, hvilken forbedring i hastigheden der er sket gennem tiden.

I denne tabel kan du se lidt flere detaljer om de forskellige standarder.

| | AGP 1.0 | AGP 2.0 | AGP 3.0 |
|------------------|--|-----------------------|--|
| Signal | 3,3 V | 1,5 V | 0,8 V |
| Protokol | Pipelineoverførsler + kildesynkron clock | AGP 1.0 + fast writes | AGP 2.0 + yderligere forbedringer, og nogle ting fjernet |
| Hastighed | 1X, 2X | 1X, 2X, 4X | 4X, 8X |
| Maks. overførsel | 266 MB/s | 1066 MB/s | 2100 MB/s |