

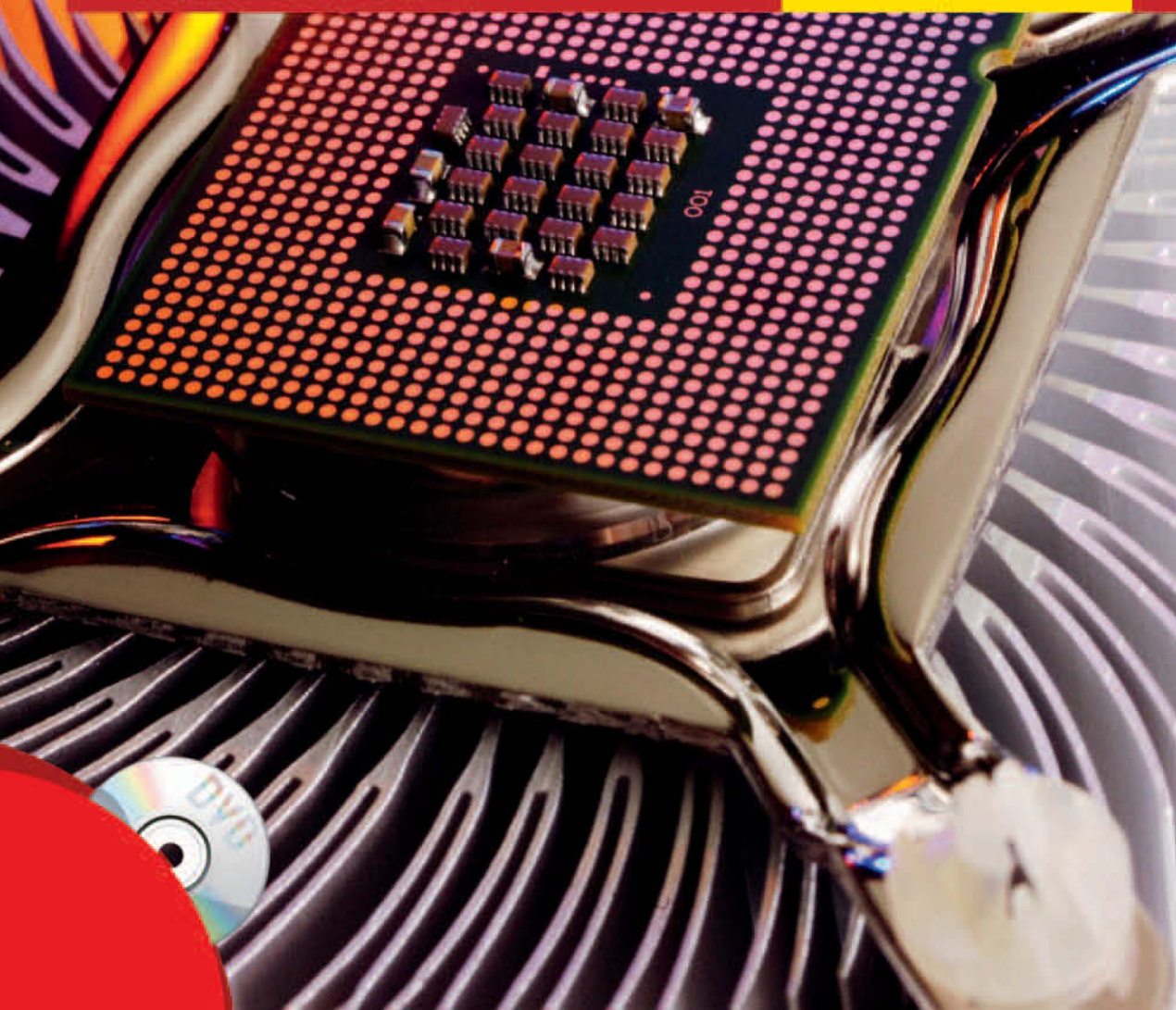
100%
Markt + Technik

PC-Werkstatt

Hardware konfigurieren,
optimieren und reparieren

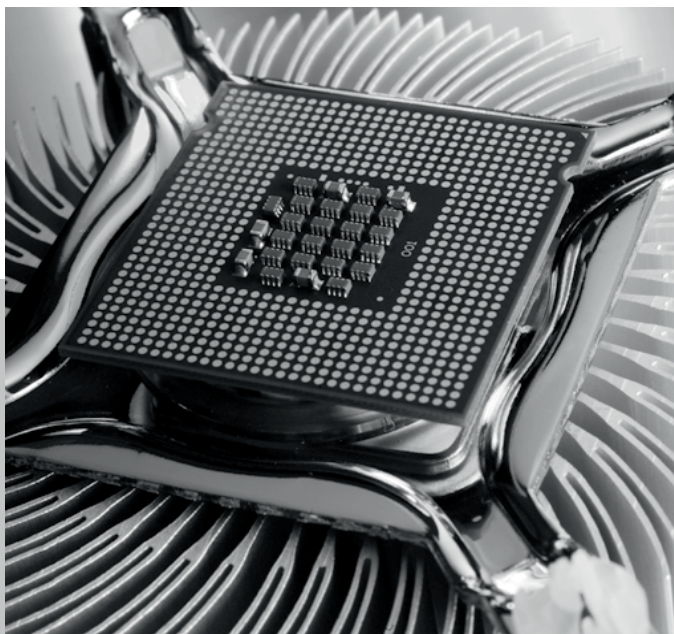
KLAUS DEMBOWSKI


Markt + Technik



3

Das Grafiksystem



KAPITEL 3

Das Grafiksystem

Kapitelübersicht

- >> **Grafikkarten**
- >> **3D-Grafik**
- >> **3D-Erweiterungsmöglichkeiten**
- >> **Bussysteme für Grafikkarten**
- >> **Anschlüsse**
- >> **Monitore**
- >> **Flachbildschirme**



Als *Grafiksystem* wird hier die Grafikkarte plus dem dazugehörigen Monitor verstanden. Diese beiden Komponenten müssen zum Erlangen der bestmöglichen Darstellungsqualität aufeinander abgestimmt sein. Das lässt sich jedoch nicht immer einfach erreichen, da es sich nun einmal um zwei unterschiedliche Einheiten handelt, welche von zahlreichen Herstellern angeboten werden, die nicht immer unmissverständliche Daten angeben und häufig unterschiedliche Vorgehensweisen für die Konfigurierung vorsehen. Nach einer ausführlichen Behandlung der Grafikkarten werden ab dem Kapitel 3.6 die Monitore näher betrachtet.

3.1 Grafikkarten

Ein Monitor benötigt stets eine Grafikkarte, um die vom Mainboard gesendeten Signale – je nach Typ – in digitale oder analoge Informationen umzusetzen, welche der Monitor weiterverarbeiten kann. Bei einigen PC-Typen befindet sich die Grafikkartenschaltung auch gleich mit auf dem Mainboard, wie es oftmals bei den preiswerteren Standardmodellen der Fall ist. Vom Standpunkt der universellen Verwendbarkeit, der Erweiterbarkeit und dem Treibersupport her ist von diesen Modellen jedoch eher abzuraten, und für aktuelle Spiele sind sie meist ebenfalls ungeeignet.

Ein Bildspeicher, auch als Grafikspeicher bezeichnet, wird für jede heute übliche Grafikkarte benötigt, denn er sorgt, vereinfacht dargestellt, dafür, dass das Bild überhaupt als stehend empfunden werden kann. Auch wenn man es nicht bemerkt, wird das Bild laufend auf den Bildschirm geschrieben, mindestens 50 mal in der Sekunde. Ein ruhiges Bild ergibt sich erst mit 75 Hz, was mit den VGA-Karten und den entsprechenden Monitoren erreicht wird.

Der Bildspeicher wird wie jeder andere Speicher des PC adressiert und enthält das aktuelle Monitorbild als digitale Informationsabbildung. Auf den Grafikkarten befindet sich ein spezieller Chip – der Grafikcontroller –, der über I/O-Adressen angesprochen wird und spezielle Register für die Bildsteuerung besitzt. Dem Grafikcontroller ist bei allen Karten ab dem VGA-Typ aufwärts ein DAC (Digital Analog Converter) nachgeschaltet, der die digitale Information in analoge für die Ansteuerung des Monitors umsetzt.

3.1.1 Traditionelle Grafikkarten im Überblick

Im Laufe der Zeit sind für den PC eine ganze Reihe von verschiedenen Einsteckkarten zur Bildwiedergabe entwickelt worden, die in der folgenden Tabelle angegeben sind. Obwohl sie gar nicht alle grafikfähig sind, wird für die folgenden Erläuterungen dennoch immer von einer Grafikkarte die Rede sein, um zu kennzeichnen, dass diese Einheit den Monitor ansteuert. Eine Videokarte ist demgegenüber eine Einheit für die Videowiedergabe bzw. für die Aufnahme von Videos.

Heute übliche Grafikkarten werden oftmals pauschal als VGA-Grafikkarten bezeichnet, was für sich allein genommen kaum etwas über die Leistung besagt und noch genau erläutert wird.

Typ	Bezeichnung	maximale Auflösung	Bussysteme	typischer max. Bildspeicher	Anschluss
MDA	Monochrome Display Adapter	25 Zeilen x 80 Zeichen	8-Bit-PC	4 kByte	DSUB-9
CGA	Color Graphic Adapter	640 x 200 Pixel (2 Farben) 320 x 200 Pixel (4 Farben)	8-Bit-PC	16 kByte	DSUB-9, BAS
HGC	Hercules Graphic Card	720 x 348 Pixel (2 Farben)	8-Bit-PC	64 kByte	DSUB-9
AGA	Advanced Graphics Adapter	640 x 200 Pixel (16 Farben)	8-Bit-PC	64 kByte	DSUB-9
PGA	Professional Graphics Adapter	640 x 480 Pixel (256 Farben)	16-Bit-ISA, MCA	512 kByte	DSUB-15 (analog)
MCGA	Multi Color Graphics Adapter	320 x 200 Pixel (256 Farben)	MCA	256 kByte	DSUB-15 (analog)
EGA	Enhanced Graphic Adapter	640 x 350 Pixel (16 Farben)	8-Bit-PC, 16-Bit-ISA	256 kByte	DSUB-9, Feature-Connector
XGA	Extended Graphic Array	1024 x 768 Pixel (256 Farben)	MCA, 16-Bit-ISA	1 MByte	DSUB-15 (analog)
8514	IBM-Standard	1024 x 768 Pixel (256 Farben)	MCA, 16-Bit-ISA	1 MByte	DSUB-15 (analog)
VGA	Video Graphics Array	640 x 480 Pixel (256 Farben)	8-Bit-PC, 16-Bit-ISA, MCA, EISA, VLB	1 MByte	DSUB-15 (analog), BNC, Feature-Connector

Tabelle 3.1: Die Daten der verschiedenen traditionellen Grafikkarten

Für die *IBM Personal System 2-Computer (PS/2)* nach der *Micro Channel Architecture (MCA)* wurde eine Grafikkarte eingeführt, die alle anderen vorherigen Grafikauflösungen emulieren kann. VGA steht dabei für Video Graphic Array und bezeichnet ein Gate-Array – einen kundenspezifischen Baustein –, der speziell für die Grafikanwendung entwickelt wurde und eine Reihe einzelner Bausteine ersetzt. VGA-Karten wurden alsbald auch mit ISA-Bus-Anschluss entwickelt, was sich dann jahrelang als Standard erwiesen hat.

In der obigen Tabelle sind die Daten für den ursprünglichen VGA-Standard angegeben, und tatsächlich sind die für heutige Verhältnisse geradezu schwachen Daten diejenigen des vielfach zitierten VGA-Standards, der somit für sich allein kaum mehr aussagefähig ist und keine 2D- und erst recht keine 3D-Beschleunigerfunktionen besitzt, wie sie heute allgemein üblich sind.

Alles, was darüber hinaus geht – wie höhere Auflösungen und mehr Farben –, entspricht im Grunde genommen nicht mehr VGA. Es hat hier Modelle laut XGA, 8514, Super-VGA oder VGA-Deluxe gegeben, wobei es sich um zahlreiche verschiedene, herstellerspezifische Realisierungen handelt, die dementsprechend mit unterschiedlicher Hardware, BIOS-Unterstützungen, Performance-Daten, Treibern und Programmierschnittstellen einhergehen.

Info

Aktuelle Monitore erkennen unterschiedliche, von einer Grafikkarte gesendete Signale und Frequenzen und können hieraus auch ein stabiles Bild produzieren. Die gemeinsame Grundlage bildet dabei vielfach ein mit dem VESA-Standard definierter Mode, der aber keineswegs die leistungsfähigste, sondern lediglich eine automatisch funktionierende Konfiguration zur Verfügung stellt, so dass oft noch manuelle Einstellungen notwendig sind, wobei im Zusammenhang mit Flachbildschirmen und (etwas) älteren Grafikkarten immer wieder Einstellungsprobleme einhergehen.

Das VESA-Konsortium (Video Electronics Standard Association) hat erst Jahre später die verbindlichen VESA-Modi definiert, die einen allgemeinen Standard für alles, was eben über das *alte* VGA hinausgeht, darstellen. Hier sind – je nach Modus – die Anzahl der Farben, die Zeichen- und Pixelgröße sowie auch die Horizontal- und Vertikal-Frequenzen (Bildwiederholfrequenz) festgelegt. Diese Modi haben prinzipiell aber nichts mit dem VESA-Local-Bus (VLB) zu tun, der eine 32-Bit-Buserweiterung des ISA-Bus darstellt und praktisch der Vorläufer von PCI ist. Es ist ratsam, bei Auflösungen größer Standard-VGA generell darauf zu achten, dass der Monitor und die Grafikkarte über möglichst viele gemeinsam nutzbare VESA-Modi verfügen. Mitunter stellt einer der VESA-Modi die einzige Kombination dar, die automatisch von beiden Komponenten des Grafiksystems (x-beliebige Grafikkarte plus Monitor) erkannt und verwendet werden kann. Je nach Grafikkartentyp werden die VESA-Modi direkt durch das Grafik-BIOS zur Verfügung gestellt, oder es muss vor der Ausführung eines Programms ein speicherresidenter Treiber, welcher dementsprechend VESA-Betriebsarten verwendet, geladen werden, was aber nur noch für Nicht-Windows-Systeme von Bedeutung ist.

Das Besondere der VGA-Karte ist, dass sie den Monitor nicht wie die Vorläufer mit digitalen Signalen ansteuert, sondern mit analogen. In der Tabelle 3.1 sind weitere Typen angegeben, die ebenfalls mit analogen Ausgangssignalen arbeiten, die nach Standard-VGA auf dem Markt kamen und heute im Grunde genommen ausgestorben sind.

Der Grund für die analoge Signalübergabe ist die erheblich verbesserte Darstellung der Farbinformation. Ein Monitor mit digitalem Eingang, wie es bis zum EGA-Standard üblich ist, kann jedoch nichts mit den analogen Signalen von VGA anfangen, was dementsprechend auch für die umgekehrte Konstellation gilt. Sowohl der Monitor als auch die Grafikkarte können Schaden nehmen, wenn hier falsche Kombinationen hergestellt werden.

Die Farben auf einer traditionellen VGA-Karte werden von einem Digital/ Analog-Wandler – auch als *Digital Analog Converter* bezeichnet, DAC – in 64 verschiedene Stufen für jeweils rot, grün und blau umgesetzt. Damit sind theoretisch $64^3 = 262.144$ Farben möglich. Selbstverständlich muss der Monitor für die analogen Eingangssignale vorgesehen sein. Wird kein VGA-Farbmonitor, sondern ein VGA-Schwarzweiß-Monitor verwendet, werden die Farben über eine Formel im BIOS so verändert, dass auf dem Monitor 64 verschiedene Graustufen entstehen. Der Monitor darf also nicht vom einfachen Typ *monochrom* sein, sondern muss explizit *Graustufen* unterstützen.

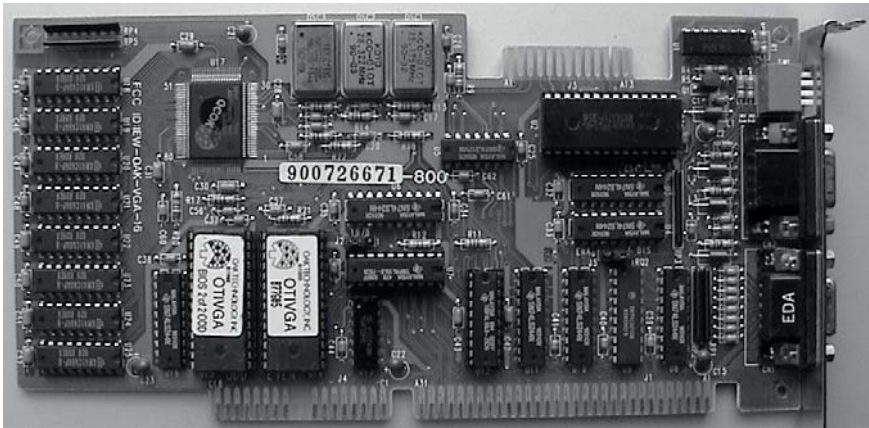


Abbildung 3.1: Eine traditionelle VGA-Karte mit 512 kByte Speicher, der in zwei Bänken mit jeweils 256 kByte ausgeführt ist

Auf den VGA-Karten befindet sich wie beim Vorläufer – der EGA-Karte – ein Speicherbaustein, der das BIOS der VGA-Karte enthält, welches die Registerkompatibilität zum Vorgänger EGA wahrt. Das BIOS beginnt ebenfalls bei C0000h, reicht jedoch bis zur Adresse C7FFFh.

Parameter	Daten
Zeichenmatrix	9 x 16 Pixel im VGA-Modus
Speichergöße	typisch 256 kByte bis zu 1 MByte
Adressbereich	A0000h-BFFFFh
Grafik-BIOS	C0000h-C7FFFh

Tabelle 3.2: Die Daten der VGA-Karte

Der Bildspeicher beginnt ebenfalls ab A0000h. Der kleinste VGA-Bildspeicher ist 256 kByte groß und belegt damit den gleichen Bereich wie eine EGA-Karte. Es gibt darüber hinaus auch Standard-VGA-Karten mit 512 kByte und 1 MByte Speicher. Da für den Speicher der Grafikkarten im PC aber lediglich der Bereich von A0000h bis BFFFFh (128 kByte) zur Verfügung steht, wird der Speicher der Grafikkarte in diesen Bereich eingeblendet. Dies geht nach einem ähnlichen Verfahren vonstatten wie die Einblendung des Expanded-Memory in den Adressbereich des PC. Der Grafikspeicher wird »scheibchenweise« in das 128 kByte große Fenster eingeblendet.

Text-Modi:			
Punktauflösung	Textauflösung	Farben	Modus
320 x 200	40 x 25	2	CGA
320 x 200	40 x 25	16	CGA
640 x 200	80 x 25	2	CGA
640 x 200	80 x 25	16	CGA
720 x 348	80 x 25	2	HGC, MDA
Grafikmodi:			
Punktauflösung	Textauflösung	Farben	Modus
320 x 200	40 x 25	2	CGA
320 x 200	40 x 25	4	CGA
640 x 200	80 x 25	2	CGA
320 x 200	40 x 25	16	EGA
640 x 200	80 x 25	16	EGA
640 x 350	80 x 25	2	EGA
640 x 350	80 x 25	16 aus 64	EGA
320 x 200	80 x 25	256	VGA
640 x 480	80 x 25	2 aus 16	VGA
640 x 480	80 x 25	16 aus 256	VGA
640 x 480	80 x 25	256 aus 64k	VGA

Tabelle 3.3: Die Auflösungen der ursprünglichen VGA-Karte

Im Grafikmodus sind 256 Farben mit einer Auflösung von 640 x 480 Bildpunkten gleichzeitig darstellbar. Der Speicher der VGA-Karte muss dann dafür mindestens 512 kByte groß sein.

Mit der Mindestausstattung von 256 kByte sind 640 x 480 Bildpunkte in 16 Farben möglich. Im Textmodus besteht ein Zeichen aus einer Matrix von 9 x 16 Bildpunkten. Selbst die meisten VGA-Karten können jedoch noch weitaus mehr. Höhere Auflösungen wie 800 x 600 Punkte oder 1024 x 786 werden auch als *Super-VGA*, *VGA-Deluxe* oder ähnlich bezeichnet, wie es auch weiter oben im Text erläutert ist, und erst mit den VESA-Modi ist hier eine allgemein gültige Erweiterung geschaffen worden.

Dem simplen Grafikcontroller von einst, dessen bevorzugtes Einsatzgebiet der Textmodus (DOS) war, sind mit der Verbreitung von Windows spezielle Funktionen für den Grafikmodus hinzugefügt worden. Bei jedem Öffnen eines Windows-Fensters oder der Anwahl eines Pull-Down-Menüs muss die Grafikinformation von der CPU des PC neu generiert werden. Wird beispielsweise ein Bildausschnitt, bestehend aus 100 x 100 Pixeln – und dies ist noch wenig – aufgebaut, sind 10.000 Bildpunkte zu manipulieren. Wird das Bild auf der Windows-Oberfläche verschoben, sind dementsprechend 20.000 Pixel zu bearbeiten, d.h. zu löschen und in die neue Position zu bringen, was natürlich möglichst schnell und flüssig erfolgen soll. Der Grafikmodus einer üblichen VGA-Grafikkarte ist bei derartigen Pixel-Operationen aber bereits überfordert.



Abbildung 3.2: Eine der ersten Grafikkarten mit 2D-Beschleuniger der Firma S3

Ein wichtiger Schritt für die Beschleunigung von Grafikfunktionen wurde im Jahre 1992 von der bis dahin eher unbekannten Firma *S3 Incorporated* vollzogen, die damals sogenannte *Windows-Acceleratorkarten* (Windows-Beschleuniger) auf den Markt brachte. Das Prinzip, welches quasi auch die Grundlage der heute bekannten 3D-Grafikkarten bildet, ist hier, dass häufig verwendete Grafikfunktionen (für Windows) quasi im Grafikchip in Hardware implementiert sind und nicht erst mühsam vom PC-Prozessor berechnet werden müssen.

Typische Funktionen für 2D-Acceleratorchips sind:

>> BitBlt

Verschieben von Fensterinhalten (Pixelblöcken). Die BitBlt-Funktion nimmt ein Rechteck auf, verschiebt es und füllt es selbstständig wieder mit dem ursprünglichen Pixelinhalt.

>> Hardware Cursor

Die Cursorbewegung und die Darstellung des Mauszeigers werden mit dem Beschleunigerchip ausgeführt. Lediglich die Mauskoordinaten werden von der CPU übergeben.

>> Line Drawing

Zeichnen von Linien. Es werden lediglich die Anfangs- und die Endkoordinate von der CPU geliefert.

>> Circle Drawing

Zeichnen von Kreisen. Es werden der Kreismittelpunkt und der Radius von der CPU geliefert, den Rest erledigt der Beschleunigerchip.

>> Polygon Fill

Füllen von Polygonen. Ein Vieleck wird vom Beschleunigerchip mit Pixelinformationen gefüllt.

>> Zoom

Kopieren eines Bildausschnitts, der um einen bestimmten Faktor vergrößert skaliert und dargestellt wird. Die Koordinaten und der Zoomfaktor werden von der CPU geliefert, die die Ausführung dann dem Beschleunigerchip überträgt.

Die 2D-Beschleunigerkarten sind VGA-kompatibel. Das bedeutet, dass sie sich nach dem Einschalten des PC erst einmal wie eine übliche VGA-Karte verhalten und erst mithilfe eines speziellen Softwaretreibers ihre beschleunigten Funktionen zur Verfügung stellen können. Die Leistungsfähigkeit einer Beschleunigerkarte zeigt sich also erst bei der Verwendung von grafikintensiven Systemen und Programmen. Jede heute erhältliche Grafikkarte verwendet das Prinzip der Grafikbeschleunigung auch für 3D-Funktionen.

3.1.2 Aufbau und Funktionsprinzip

Neben dem eigentlichen Grafikchip befindet sich auf einer Grafikkarte ein D/A-Wandler (RAMDAC), der die digitale Information in die RGB-Werte (rot, grün, blau) für den angeschlossenen Monitor umsetzt. Des Weiteren sind noch Bauelemente für die Adressendekodierung nötig und ein RAM (Bildspeicher) für die Grafikdaten. Für dieses RAM werden auf einer Standard-VGA-Karte »normale« dynamische RAMs (DRAMs) eingesetzt.

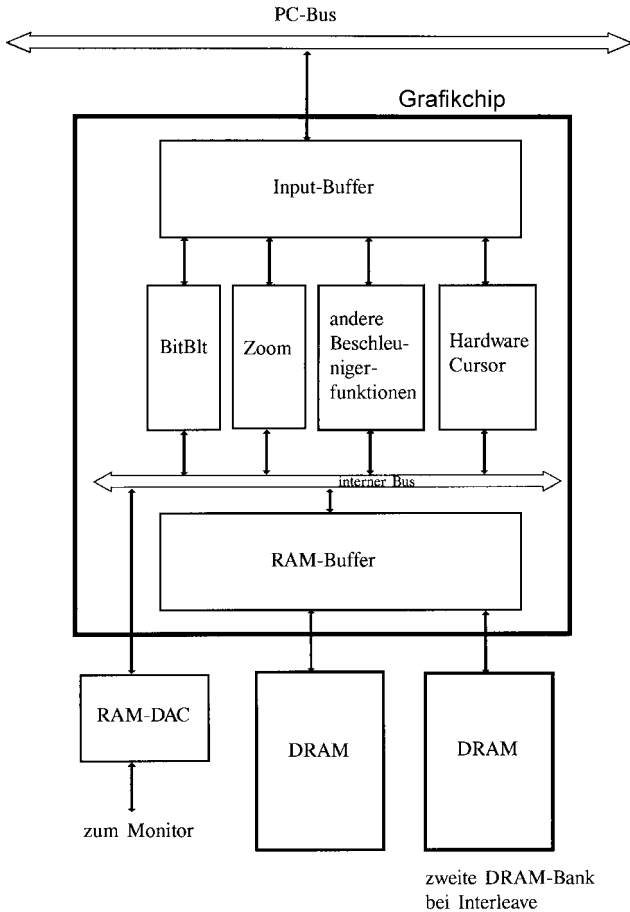


Abbildung 3.3: Der prinzipielle Aufbau einer Grafikkarte mit DRAMs.

Die zweite DRAM-Bank wird nur bei Grafikkarten verwendet, die mit Memory-Interleaving arbeiten.

Ein DRAM verfügt über einen einzigen Datenport zur Kommunikation mit der CPU und dem RAMDAC. Das heißt, dass entweder Daten vom Bus empfangen oder Daten zum RAMDAC ausgegeben werden. Der Grafikchip sorgt dabei für die Umschaltung der »Datenwege«, wodurch die Geschwindigkeit der Grafikausgabe beeinflusst wird. Auf Grafikkarten sind auch VRAMs (Video RAMs) gebräuchlich, die prinzipiell einen schnelleren Datentransfer als DRAMs ermöglichen. VRAMs enthalten einen DRAM-Kern, der genauso wie ein herkömmliches DRAM aufgebaut ist. Daneben enthält ein VRAM ein Serial Access Memory – SAM –. Das SAM ist ein 8 Bit breiter serieller Port, der mit einem eigenen Takt arbeitet, welcher praktisch dem Pixeltakt des RAMDACs entspricht und demnach nach außen hin unabhängig vom integrierten DRAM arbeiten kann.

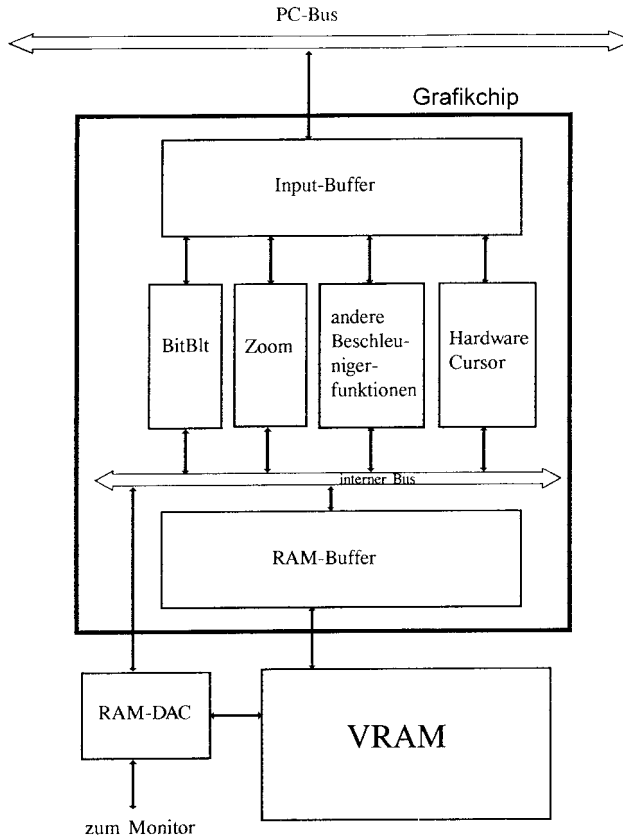


Abbildung 3.4: Eine Grafikkarte mit VRAMs kann gleichzeitig neue Daten vom Bus empfangen und Daten zum RAMDAC ausgeben.

Über den PC-Bus (ISA, PCI, AGP, PCIe) wird das DRAM mit Daten gefüllt, und gleichzeitig kann das SAM die vorigen Daten zum RAMDAC ausgeben. Über eine serielle Datenleitung des SAMs wird üblicherweise 1 Byte (nacheinander) zum RAMDAC gesendet. Ein RAMDAC mit einer Bitbreite von 64 Bit verfügt also über acht serielle Eingänge.

Weil VRAMs gegenüber den Standard-DRAMs teurer sind, haben einige Hersteller einen anderen Weg zur schnelleren Grafikausgabe mit DRAMs in Form des Speicher-Interleaving realisiert. Die dynamischen RAMs werden hierfür in zwei Bänken organisiert, wobei in der einen der Inhalt der geraden und in der anderen der Inhalt der folgenden (ungeraden) Adresse gespeichert wird. Somit kann ein Byte zum RAMDAC ausgegeben werden, während das folgende bereits von der CPU adressiert wird. Die Arbeitsweise mit DRAM-Interleaving bietet ferner den Vorteil, dass die DRAM-Konfiguration entsprechend der gewählten Auflösung per Software (Treiber) optimiert werden kann. Eine noch höhere Datenbandbreite wird realisiert, wenn VRAMs im Interleaving

betrieben werden. Das Prinzip des Speicher-Interleaving wird nicht nur im Zusammenhang mit Grafikkadaptern angewendet, sondern generell für einen versetzten Speicherzugriff, wo während der Datenverarbeitung bereits neue Daten gelesen oder auch geschrieben werden können, wie es etwa bei den aktuellen Mainboards mit Dual-Channel-Speicher-Architektur der Fall ist.

Neben DRAM- und VRAM-Bausteinen sind auf Grafikkarten auch Speicherchips zu finden, die als EDO-DRAM, SDRAM oder SGRAM bezeichnet werden. Bei allen aktuellen Grafikkarten werden für den Grafikspeicher verschiedene DDR-Chips (Double Data Rate) eingesetzt, die prinzipiell genauso arbeiten wie die DDR-Chips auf den Modulen für den PC-Arbeitspeicher. Näheres zu diesen Bausteinen, wie auch allen anderen Speicherchips, findet sich im Kapitel *Speicherbausteine und Module*.

Der Speicher auf einer Grafikkarte kann grundsätzlich in unterschiedlicher Bitbreite organisiert sein. Zum jeweiligen Bus des PC hin sind 16 Bit bei ISA oder 32 Bit bei PCI und AGP üblich, eben die maximal mögliche Adressierungsbreite, die das jeweilige Bussystem hergibt. Zur anderen Seite hin – zum RAMDAC – sind hier mindestens 64 Bit üblich und aktuell bis zu 512 Bit möglich.

Je breiter der Bus zwischen dem Grafikchip und dem Bildspeicher ausgeführt ist, desto schneller lassen sich die Bilder aufbauen, wobei der Grafikchip explizit für die jeweilige Bitbreite entwickelt sein muss. Insbesondere 3D-Anwendungen profitieren von den höheren Bitbreiten, wo dann mindestens 128 Bit üblich sind, wie es etwa mit dem älteren Chip RIVA 128 der Firma Nvidia der Fall ist, der mit SGRAM arbeitet.

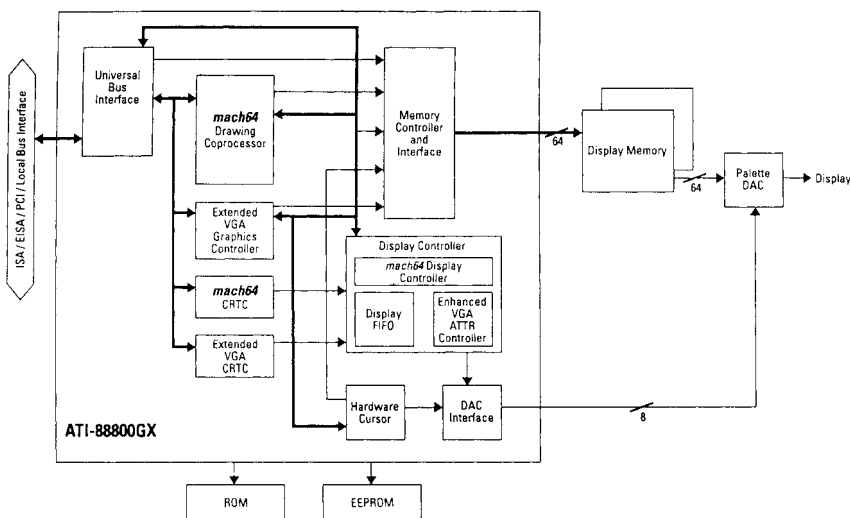


Abbildung 3.5: Die bekannte Grafikkarte MACH 64 der Firma ATI arbeitet mit einem eigenen ATI-Grafikchip und einem 64 Bit RAMDAC (Palette DAC).

Die Abbildung 3.6 zeigt, dass der Speicher beispielsweise bei einer Grafikkarte mit dem Chip Vision964 der Firma S3 auf unterschiedliche Art und Weise aufgebaut sein kann. Die hier maximal mögliche 64-Bit-Datenbreite wird nur dann erreicht, wenn beide VRAM-Bänke bestückt sind, andernfalls erhält man lediglich einen 32-Bit-Datenzugriff. Auch wenn auf dieser Seite – zwischen Grafikchip und VRAM – 64 Bit gegeben sind, besagt dies noch nicht automatisch, dass die andere Seite – vom VRAM zum RAMDAC – ebenfalls in dieser Breite ausgeführt ist, sondern sie kann auch größer (128 Bit) oder kleiner (32 Bit) sein.

Aus diesen Erläuterungen mag ersichtlich werden, dass Grafikkarten verschiedener Hersteller, die alle auf dem gleichen Grafikchip basieren, nicht zwangsläufig die gleiche Performance erreichen können, wenn sich bereits die Speicherarchitektur voneinander unterscheidet.

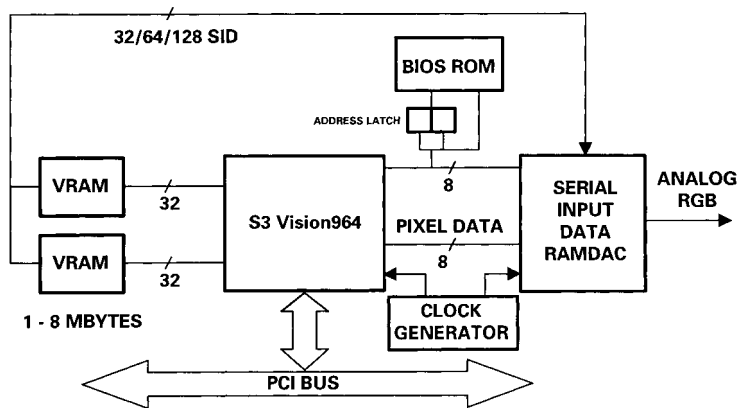


Abbildung 3.6: Die Applikation einer Grafikkarte mit »S3 Vision964« lässt einige Optionen für die Realisierung des Grafikspeichers offen.

Der zweite Grund für die unterschiedliche Qualität scheinbar identischer Grafikkarten liegt in der Verwendung verschiedener RAMDACs. Je höher die mögliche maximale Pixelrate, desto teurer ist auch der RAMDAC.

Die Abbildung 3.7 zeigt einen direkt zum S3 Vision964 passenden RAMDAC, den Typ 86708 der Firma S3. Das Innenleben dieses RAMDACs lässt erkennen, dass für jede Farbe (Analog Red, Analog Green, Analog Blue) ein eigener 8-Bit A/D-Umsetzer vorhanden ist, wobei die digitalen Farbsignale über einen Multiplexer (Mux) geschaltet werden.

Für die Steuerung jeder Farbe ist vor dem Multiplexer ein separater Speicher (Palette RAM) vorhanden, damit die Daten vom CPU-Bus-Interface zeit-synchron mit den vom Grafikspeicher gesendeten Pixeladressen (PA 15:0) zu verarbeiten sind. Dieser Schaltungsteil wird als *Color Look-up Table* (CLUT) bezeichnet und bei *True Color* (24 Bit) schaltungstechnisch umgangen (bypass).

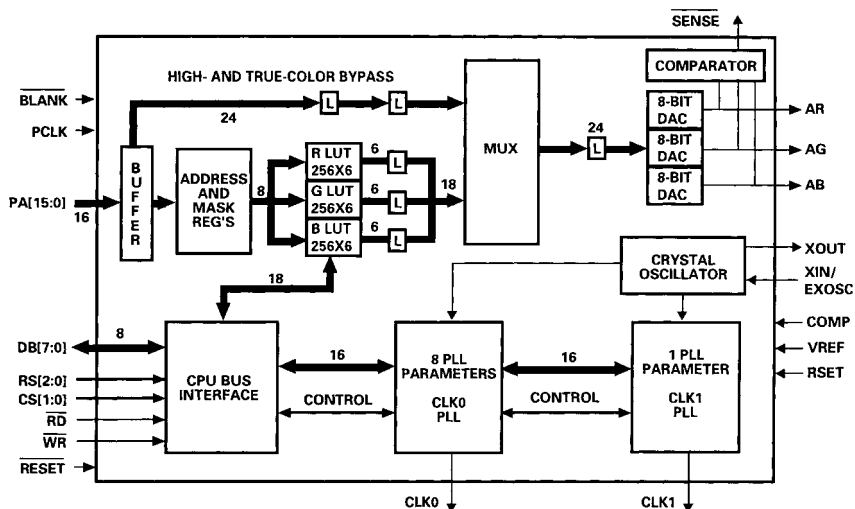


Abbildung 3.7: Der interne Aufbau eines RAMDACs zeigt die Verarbeitung der Grafikinformati

In Abhängigkeit von der jeweiligen Auflösung und der Geschwindigkeit der eingesetzten Speicherbausteine können Pixel- (PCLK) und Speichertakt (CLK) unabhängig voneinander festgelegt werden, was eine möglichst optimale Pixelrate ermöglicht, die bei diesem Typ bei 135 MHz liegt und eine Auflösung von 1280 x 1024 Bildpunkten mit einer Bildwiederholrate von 75 Hz erlaubt.

Grafikauflösungen und Speicherbedarf

Aktuelle Grafikkarten werden mit so viel Speicher ausgeliefert, wie vor noch gar nicht langer Zeit als Hauptspeicher in PCs eingebaut wurde. Das Aufrüsten des Speichers einer Grafikkarte ist daher nur noch bei älteren Modellen sinnvoll, wobei die Anschaffung einer neuen Grafikkarte – mit mehr Speicher als die alte – demgegenüber preiswerter kommt. Allerdings ist es erstaunlich, wie wenig Speicher prinzipiell für bestimmte Auflösungen benötigt wird, wie es die folgende Tabelle zeigt.

Auflösung (Bildpunkte)	Farbanzahl	Grafikspeicher (Minimum)
640 x 480	16	256 kByte
640 x 480	256	512 kByte
640 x 480	65536	1 MByte
640 x 480	16,7 Mio.	1 MByte
800 x 600	16	256 kByte
800 x 600	256	512 kByte
800 x 600	65536	1 MByte
800 x 600	16,7 Mio.	1,5 MByte

Tabelle 3.4: Der Grafikspeicherbedarf in den üblichen Abstufungen für verschiedene Auflösungen

Auflösung (Bildpunkte)	Farbanzahl	Grafikspeicher (Minimum)
1024 x 768	16	512 kByte
1024 x 768	256	1 MByte
1024 x 768	65536	2 MByte
1024 x 768	16,7 Mio.	2,5 MByte
1280 x 1024	16	1 MByte
1280 x 1024	256	1,5 MByte
1280 x 1024	65536	3 MByte
1280 x 1024	16,7 Mio.	4 MByte

Tabelle 3.4: Der Grafikspeicherbedarf in den üblichen Abstufungen für verschiedene Auflösungen

Der Speicherbedarf einer Grafikkarte kann allgemein nach der folgenden Formel berechnet werden:

Speicherbedarf (in Byte) =
$$\frac{\text{Auflösung} \times \text{Bitbreite für die Farbanzahl}}{8}$$

Die »Bitbreite für die Farbanzahl« – auch als Farbtiefe bezeichnet – gibt dabei an, wie viele Bits nötig sind, um jeweils die gewünschte Anzahl an Farben zu erhalten (siehe auch folgende Tabelle). Wird eine Farbtiefe von mindestens 24 Bit verwendet, spricht man vom True-Color-Modus

Farbtiefe	Anzahl der Bits
2	1
4	2
16	4
256	8
65536	16
16,7 Mio.	24

Tabelle 3.5: Der Zusammenhang zwischen der Farbtiefe und der hierfür nötigen Anzahl der Bits. Zur Darstellung von 65536 Farben werden beispielsweise 16 Bit benötigt (2 hoch 16).

Für die True-Color-Darstellung (24 Bit) bei einer Auflösung von 1280 x 1024 Bildpunkten (Pixeln) errechnet sich der Speicherbedarf (1 kByte = 1024 Byte) mit der obigen Formel wie folgt:

$$\frac{1280 \times 1024 \times 24}{8} = 3.932.160 \text{ Byte} = 3.840 \text{ kByte} = 3,75 \text{ MByte}$$

Der Speicher müsste demnach über eine Kapazität von 4 MByte verfügen, da dies einer standardisierten Größe des Grafikspeichers entspricht. Ein größerer Grafikspeicher bedeutet auch einen Geschwindigkeitszuwachs bei der Bilddarstellung, da sich dadurch die Datenübertragungsrate quasi erhöht, weil größere Teile der Grafikinformation im Speicher der Grafikkarte lokalisiert sind und nicht erst über den (PCI-)Bus geladen werden müssen. Ein möglichst großer Grafikspeicher hat demnach nicht nur höhere Auflösungen zur Folge, sondern kann auch für eine schnellere Grafikverarbeitung sorgen.

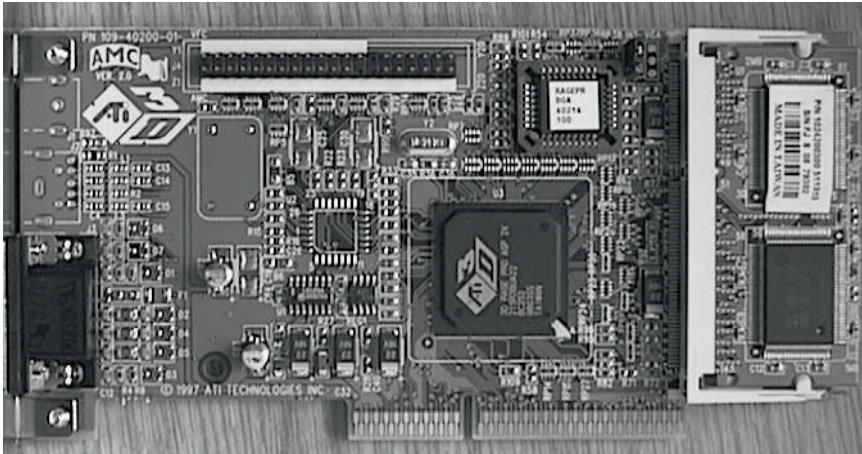


Abbildung 3.8: Die Grafikkarte ATI 3D Rage verwendet ein SO-DIMM (rechts) als Grafikspeicher.

Der Grafikspeicher ist bei den verschiedenen Karten unterschiedlich aufgebaut und kann oftmals auch nur mit den speziell hierfür ausgelegten Modulen oder Chips erweitert werden. Einen gewissen Standard stellten ein Zeit lang zumindest die SO-DIMMs dar, wie sie auch in Notebooks als PC-Arbeitspeicher zum Einsatz kommen.

Die erläuterten Zusammenhänge für die jeweilige Größe des Grafikspeichers betreffen nur die 2D-Funktion und gelten nicht für die dreidimensionale Darstellung. Diese benötigt bereits für relativ einfache Effekte die drei- bis vierfache Speichergröße (z.B. Z-Buffer) bei gleicher Auflösung im Vergleich zur einer traditionellen 2D-Grafikkarte.

Der Z-Buffer ist ein für 3D-Karten zusätzlicher Speicherbereich, der für die Tiefeninformation verwendet wird. Zu jedem Bildpunkt gehört mindestens eine mit 16 Bit aufgelöste Tiefeninformation. Heutzutage beherrschen selbst die einfachen, im Chipset integrierten Grafikadapter dreidimensionale Funktionen in unterschiedlicher Qualität, was den »ambitionierten Gamer« jedoch selten zufrieden stellt und daher – in den folgenden Kapiteln – noch einer genaueren Betrachtung bedarf.