

Absolventská práce

2004

Jiří VAŠEK

Prohlašuji že jsem absolventskou práci na téma 3D grafika a počítačové hry vypracoval samostatně a použil jen pramenů, uváděných v seznamu .

V Českých Budějovicích dne 19. května 2004

Studijní obor : Výpočetní technika a programování

3D grafika a počítačové hry

Absolventská práce

Autor : Jiří VAŠEK

Vedoucí absolventské práce : PaedDr. Petr PEXA

**Děkuji panu PaedDr. Pexovi za odborné vedení při zpracování absolventské práce.
Dále firmám L-Tronic, Třeboň a PC-Guardian, České Budějovice za umožnění přístupu na Internet, odkud jsem získával nejvíce informací.
A nakonec komunitě z diskusního fóra pctuning.cz.**

Obsah

| | |
|---|-----------|
| OBSAH | 5 |
| ÚVOD | 8 |
| 1 CO JE GRAFICKÁ PIPELINE A JAK FUNGUJE | 10 |
| 1.1 APLIKACE | 10 |
| 1.2 GEOMETRIE..... | 11 |
| 1.3 TRIANGLE SETUP | 12 |
| 1.4 TEXTUROVÉ OPERACE..... | 12 |
| 1.5 RASTER OUTPUTS (ROP) | 13 |
| 1.6 VĽATNÍ VYKRESLENÍ OBRAZU..... | 13 |
| 2 ZÁKLADNÍ POJMY 3D GRAFIKY A FUNKCE 3D AKCELERÁTORŮ ... | 15 |
| 3 ZÁKLADNÍ ČÁSTI 3D AKCELERÁTORU A ZÁKLADNÍ POJMY | 22 |
| 4 SBĚRNICE POUŽÍVANÉ GRAFICKÝMI KARTAMI | 26 |
| 5 TYPY PAMĚTÍ POUŽÍVANÝCH V AKCELERÁTORECH | 29 |
| 6 3D PROGRAMOVACÍ ROZHRANÍ (API) A JAZYKY | 31 |
| 6.1 3D PROGRAMOVACÍ API | 31 |
| <i>Microsoft Direct3D</i> | 31 |
| <i>S3i OpenGL</i> | 31 |
| <i>3Dfx Glide</i> | 32 |
| <i>S3 MeTaL</i> | 32 |
| 6.2 PROGRAMOVACÍ JAZYKY HLSL (HIGH LEVEL SHADING LANGUAGES) | 32 |
| 6.3 3D WRAPPERY | 33 |
| 6.4 3D BENCHMARKY | 33 |
| 7 TECHNICKÉ SPECIFIKACE GRAFICKÝCH KARET | 35 |
| 7.1 NULTÁ GENERACE (PRVNÍ AKCELERÁTORY) (1995-1996)..... | 35 |
| <i>Rendition Verité 1000</i> | 35 |
| <i>nVidia NV1</i> | 35 |
| <i>Matrox MGA1064SG</i> | 35 |
| <i>Matrox MGA1164SG</i> | 35 |
| <i>Videologic PowerVR PCX1</i> | 36 |
| <i>S3 Virge</i> | 36 |
| <i>ATi 3D Rage</i> | 36 |
| <i>ATi 3D Rage II</i> | 36 |
| <i>3DLabs Gigi</i> | 36 |
| <i>3DLabs Permedia</i> | 37 |
| 7.2 PRVNÍ GENERACE (1996-1997) | 38 |
| <i>3Dfx Voodoo</i> | 38 |
| <i>3Dfx Voodoo Rush</i> | 38 |
| <i>nVidia Riva 128</i> | 39 |
| <i>nVidia Riva 128 ZX</i> | 39 |
| <i>ATi Rage Pro</i> | 39 |

| | |
|---|----|
| <i>Number Nine Ticket To Ride</i> | 39 |
| <i>3DLabs Permedia 2</i> | 40 |
| <i>Rendition Verité 2100</i> | 40 |
| <i>Rendition Verité 2200</i> | 40 |
| <i>Videologic/NEC PowerVR PCX2</i> | 40 |
| <i>Matrox Productiva G100</i> | 41 |
| <i>Matrox MGA-2164W</i> | 41 |
| 7.3 DRUHÁ GENERACE (1998)..... | 42 |
| <i>3Dfx Voodoo2</i> | 42 |
| <i>Intel i740</i> | 42 |
| <i>nVidia Riva TNT</i> | 43 |
| <i>Matrox G200</i> | 43 |
| <i>Videologic PowerVR Second Generation</i> | 43 |
| <i>S3 Savage 3D</i> | 44 |
| <i>3Dfx Voodoo Banshee</i> | 44 |
| 7.4 TŘETÍ GENERACE (1999) | 45 |
| <i>nVidia Riva TNT2</i> | 45 |
| <i>ATi Rage 128</i> | 45 |
| <i>3Dfx Voodoo3</i> | 46 |
| <i>S3 Savage 4</i> | 46 |
| <i>3DLabs Permedia 3</i> | 47 |
| <i>Videologic/NEC PowerVR Neon 250</i> | 47 |
| <i>Number Nine Ticket To Ride IV</i> | 47 |
| <i>Matrox G400</i> | 47 |
| <i>BitBoys Oy Glaze 3D</i> | 48 |
| 7.5 ČTVRTÁ GENERACE (PRVNÍ DIRECTX 7 / T&L GENERACE) (2000) | 49 |
| <i>nVidia GeForce 256 (NV10)</i> | 49 |
| <i>S3 Savage 2000</i> | 49 |
| <i>ATi Rage 128 Pro</i> | 50 |
| <i>Matrox G450</i> | 50 |
| 7.6 PÁTÁ GENERACE (DRUHÁ DIRECTX 7 / T&L GENERACE) (2000) | 51 |
| <i>nVidia GeForce 2 (NV15)</i> | 51 |
| <i>nVidia GeForce 2 MX (NV11)</i> | 51 |
| <i>ATi Radeon - Radeon 7200 (R100)</i> | 52 |
| <i>ATi Radeon VE - Radeon 7000 (RV100)</i> | 52 |
| <i>3Dfx VSA-100</i> | 53 |
| <i>ST Micro / Imagination PowerVR3 - Kyro</i> | 54 |
| <i>S3 Savage XP</i> | 54 |
| 7.7 ŠESTÁ GENERACE (PRVNÍ DIRECTX 8 GENERACE) (2001) | 55 |
| <i>nVidia GeForce 3 (NV20)</i> | 55 |
| <i>ATi Radeon 8500, Radeon 9100 (R200)</i> | 56 |
| <i>ATi Radeon 7500 (RV200)</i> | 56 |
| <i>ST Microelectronics Kyro II</i> | 56 |
| <i>ST Microelectronics Kyro II SE</i> | 57 |
| <i>ST Microelectronics / PowerVR Series 4</i> | 57 |
| <i>PowerVR 5 (Kyro III)</i> | 57 |
| <i>Matrox G550</i> | 57 |
| 7.8 SEDMÁ GENERACE (DRUHÁ DIRECTX 8 GENERACE) (2001-2002)..... | 58 |
| <i>nVidia GeForce 4 Titanium (NV25 a NV28)</i> | 58 |
| <i>nVidia GeForce 4 MX, MX 4000 (NV17 a NV18)</i> | 58 |

| | |
|--|----|
| <i>Matrox Parhelia 512, Parhelia-LX</i> | 59 |
| <i>ATi Radeon 9000 (Pro) (RV250)</i> | 60 |
| <i>SiS 315</i> | 60 |
| <i>SiS Xabre</i> | 60 |
| 7.9 DNEŠNÍ AKCELERÁTORY (PRVNÍ DIRECTX 9 GENERACE) (2002-2003) | 62 |
| <i>ATi Radeon 9700 (Pro) (R300)</i> | 62 |
| <i>ATi Radeon 9500 (Pro) (RV300)</i> | 62 |
| <i>ATi Radeon 9200 (Pro) (RV280)</i> | 63 |
| <i>nVidia GeForce FX 5800 (Ultra) (NV30)</i> | 64 |
| <i>nVidia GeForce FX 5600 (Ultra, XT) (NV31)</i> | 65 |
| <i>nVidia GeForce FX 5200 (Ultra), 5500 (NV34)</i> | 65 |
| <i>SiS Xabre II (projekt)</i> | 66 |
| <i>Trident XP4 (projekt)</i> | 66 |
| <i>Trident XP 8 (projekt)</i> | 66 |
| 7.10 DNEŠNÍ AKCELERÁTORY (DRUHÁ DIRECTX 9 GENERACE) (2003-2004) | 67 |
| <i>ATi Radeon 9800 (Pro) (R350)</i> | 67 |
| <i>ATi Radeon 9800 XT (R360)</i> | 67 |
| <i>ATi Radeon 9550, 9600 (Pro) (RV350)</i> | 68 |
| <i>ATi Radeon 9600 XT (RV360)</i> | 68 |
| <i>nVidia GeForce FX 5900 (Ultra, XT) (NV35)</i> | 68 |
| <i>nVidia GeForce FX 5950 Ultra (NV38)</i> | 69 |
| <i>nVidia GeForce FX 5700 (Ultra) (NV36)</i> | 70 |
| <i>XGI Volari</i> | 70 |
| <i>S3 Deltachrome</i> | 71 |
| 7.11 NEJMODERNĚJŠÍ A BUDOUCÍ GRAFICKÉ ČIPY (2004) | 73 |
| <i>nVidia GeForce PCX</i> | 73 |
| <i>nVidia GeForce 6800 (Ultra) (NV40)</i> | 73 |
| <i>nVidia NV45</i> | 74 |
| <i>ATi Radeon X800 (R423,R420)</i> | 74 |
| <i>ATi RV380</i> | 74 |
| <i>ATi RV370</i> | 75 |
| <i>S3 Gammachrome</i> | 75 |
| ZÁVĚR | 76 |
| SEZNAM POUŽITÉ LITERATURY | 77 |
| <i>Časopisy</i> | 77 |
| <i>Zdroje všeobecných informací na internetu</i> | 77 |
| <i>Zdroje technických informací na internetu</i> | 77 |
| <i>Zdroje obrázků</i> | 77 |
| PŘÍLOHY | 78 |

Úvod

Zřejmě žádné technologické odvětví v posledních letech nedoznalo tak překotného vývoje jako počítače. Ještě před několika desetiletími byly počítače používány výhradně k armádním účelům, avšak s postupujícím technickým pokrokem se dostaly do kanceláří a nakonec i do domácností. Výroba počítačů drasticky zlevnila a dnes si může počítač dovolit prakticky každý. Počítače v různých podobách se staly součástí našeho života. A společně s rychlým vývojem samotné techniky se vyvíjely programy. Dříve sloužily lidem při práci, pak se však objevily počítačové hry. Z pouhého rozmaru programátorů při krácení dlouhé chvíle se vyvinuly plnohodnotné programy, jejichž tvorbou se v dnešní době už programátoři bez problému užívají. Původně široká veřejnost počítačovými hrami spíše opovrhovala; herní principy byly teprve mladé a také grafická stránka nebyla zrovna lákavá. Ovšem během posledních několika let vznikl plnohodnotný herní průmysl, v němž jsou investovány nepředstavitelné peníze a podle odhadů ekonomů obrat herního průmyslu dokonce během řádově desítek měsíců překoná obrat dnes ještě bezkonkurenčního průmyslu filmového. Společně se zdokonalováním počítačového hardware a spolu s ním i her se herní průmysl postavil po bok kinematografie jakožto další druh masové zábavy. I když v našich končinách střední a východní Evropy ještě penetrace široké veřejnosti počítačovými hrami nedosáhla takového stupně jako v západní Evropě, Americe nebo Japonsku, završují hry své vítězné tažení už i zde.

Svět počítačových her je fascinující. Díky obrovskému technologickému pokroku můžeme na monitorech počítačů vidět dech beroucí grafická kouzla, o jakých se ještě před několika lety nikomu ani nezdálo. Pokrok došel až tak daleko, že filmový a herní průmysl se do jisté míry přiblížily a provázaly; díky tomu je dnes možné i něco tak úžasného, jako filmy kompletně tvořené na počítači, bez účasti jediného živého herce; jako jediný příklad za všechny uvedu film Final Fantasy, který je tvořen - jak jinak - na motivy stejnojmenné hry.

Společně se stále dokonalejšími hrami výrobci neustále chrlí nové a nové modely grafických karet, určené k náhradě předchozích. Ve světě počítačových her platí doslova, že grafická karta je zastaralá už v okamžiku, kdy se objeví na pultech obchodů. Jakkoliv velký pokrok zaznamenaly ostatní počítačové komponenty, jsou grafické karty bezesporu tou nejsložitější komponentou v celém počítači; to, že dnešní grafické čipy jsou dvakrát komplexnější než dnešní procesory, mluví samo za

sebe. Ruku v ruce s tímto nepředstavitelně rychlým vývojem jde i marketing, neboť k prosazení už dneska nestačí jen dobrý výrobek, ale spotřebitel musí být přesvědčen o tom, že kupuje lepší výrobek než nabízí konkurence.

Počítačové hry mě fascinovaly již od chvíle kdy jsem s počítači začínal, díky hrám jsem vlastně se o počítače začal zajímat. A své zkušenosti bych chtěl předat dál. Tato práce by měla sloužit jako příručka běžnému uživateli. Množství údajů a termínů z oboru počítačově-herní grafiky splývá v jeden neprůhledný spletenec, jehož rozpletení jsem si stanovil jako hlavní úkol. Cílem je stanovit komplexní přehled technologií, které tvoří počítačové hry. Čtenář zde najde veškeré technické specifikace grafických karet, společně s jejich srovnávacími testy. Osobně si myslím, že tento přehled je důležitý, neboť marketingové snahy dnešních výrobců grafických karet způsobily na trhu obtížně průhledný zmatek. Druhým bodem práce je vysvětlení zásadních pojmů z oboru grafické akcelerace. Tato část by měla čtenáři osvětlit, co vlastně znamenají termíny používané v oblasti herní grafiky. Je důležité vědět, jak vlastně funguje grafika v počítačových hrách, protože díky tomu je možno vyladit počítač na maximální výkon.

Není dost dobře možné shromáždit veškeré informace o herní grafice na následujících stránkách; novinky z této oblasti se objevují prakticky denně, a velká část následujícího textu bude již za několik týdnů zastaralá. Proto bych chtěl tuto práci pojmut jako základ, na kterém budu v budoucnu dále pracovat.

1 Co je grafická pipeline a jak funguje

1.1 Aplikace

Příklad konečné podoby počítačové hry, tak jak ji vidí běžný uživatel, se skrývá hluboko ve strojovém kódu hry. Toto je úplný počátek putování dat z počítače směrem k uživateli. V kódu je samozřejmě to samé co uživatel vidí na monitoru, ovšem symbolizované změti čísel a dalších údajů. Tyto údaje se nazývají matematický model. Většinou se jedná o geometrické souřadnice jednotlivých prvků scény (jako například souřadnice hran těles a podobně), a další doplňkové informace. Za zpracování matematického modelu je zodpovědný renderovací motor hry, většinou nazývaný engine hry. I když v dnešní době více než jindy dochází k licencování engineů a mnoho her je vlastně jen jiná grafická a funkční slupka na stejné šabloně, pořád je engine jakýmsi unikátním postupem, jak tvořit matematický model. Touto fází se nebudeme zabývat, neboť z technického hlediska nemá přílišný význam. Aplikace musí nejprve provést základní operace, jako je simulace fyziky, umělé inteligence, a vlastně tak určit jaká scéna bude vlastně zobrazena, které objekty na ní budou, jak budou natočené, nasvícené a podobně. Tyto funkce jsou realizovány softwarem - jsou příliš závislé na programu a urychlit je mohou hlavně speciální instrukce 3DNow! (AMD) a SSE (Intel). 3D grafika také hodně závisí na výkonu FPU (Floating point unit), což je část procesoru provádějící výpočty v plovoucí desetinné čárce. Výkon FPU se uvádí ve Flops (počet operací v plovoucí čárce za sekundu). V této fázi se provádí i výpočet Higher-Order-Surfaces (teselace).

Tvorba 3D modelů - Základem 3D modelů je drátěná kostra, která se skládá z bodů (vertexů). Tři vertexy jsou spojeny hranami (edges), které definují trojúhelník (triangle), který je elementární plochou pro vykreslení. Některé akcelerátory umí i složitější polygony než trojúhelník, ale všechny se dají na trojúhelník rozložit. Se zakřivenými plochami je to složitější - musí se rozdělit na řadu trojúhelníků, proto je jejich implementace složitější.

Soubor vertexů a trojúhelníků tvoří drátěný model - mesh.

Drátěný model je možné použít mnoha způsoby - v nejjednodušším případě se vytvoří řada drátěných modelů poloh těla postavy (postur), které se střídají jako animační okénka ve dvourozměrné grafice. Tento systém je výpočetně nenáročný,

ale nevypadá příliš realisticky. Složitější systémy jsou vybaveny kostrou a počítač podle fyzikálních zákonů přepočítává tvar celé drátěné kostry. Tak je možné dosáhnout desítek tisíc odlišných poloh těla. Tento systém je výpočetně náročný a používá se tam, kde záleží třeba na přesné fyzikální reakci (bojové hry...). Nejsložitějším způsobem práce s modelem je variable mesh - drátěná kostra, jejíž složitost se mění v závislosti na síle procesoru. Zde je možno dynamicky měnit počet trojúhelníků a dosáhnout tak plynulosti animace i na pomalejších počítačích.

Na připravenou drátěnou kostru je ještě nutné natáhnout povrch (skin). Většinou je třeba připravit více skinů a počítač vyměňuje jednotlivé skiny v závislosti na situaci ve hře.

Když je hra připravena, má grafika z matematického modelu ještě daleko k zobrazení. Její putování vede přes mnoho míst, které ji zpracují. "Roura", kterou data prochází, se nazývá grafická pipeline. V prvních generacích grafických karet i her byla realizována softwarově, dnes se už výlučně používá hardwarové řešení.

1.2 Geometrie

Ve fázi výpočtů geometrie se provádí clipping (určení které objekty budou zobrazeny) a culling (ořezání neviditelných objektů), transformace objektů (natočení, zmenšení, zvětšení) a osvětlení. V dnešní době jsou tyto funkce prováděny hardwarově, děje se tak od doby co GeForce zavedly T&L funkce. Od zavedení DirectX 8 byly tyto funkce přeneseny na Vertex shader. U starších akceleratorů byly tyto funkce prováděny softwarově. Dnešní GPU mají kvůli kompatibilitě jak T&L, tak vertex shader.

Vertex Shader 1.0 - jedná se vlastně o druhou generaci T&L, implementovanou v DirectX 8. Zatímco první generace má na starosti transformace a nasvícení 3D scény, vertex shader dokáže 3D modely na základě zadaných dat sám modifikovat. Hlavním principem funkce vertex shaderu je zpracování vrcholů polygonů (vertexů). Současně je dostupných až 32 transformačních matic současně (dříve 2 až 4 matice). Toto mimo jiné umožňuje mnohem lépe modelovat živé tvory. Akcelerátory s vertex shaderem obvykle obsahují i pevnou T&L jednotku pro zachování zpětné kompatibility. Programově je vertex shader nízkourovňový kód, podobný assembleru; jedná se o poměrně komplexní záležitost, programátoři mohou provádět podmíněné skoky, volat podprogramy apod., i když existují omezení na velikost dat

nebo počet instrukcí. Vertex shadery se typicky využívají při modelování kolizí objektů, výpočtech realistických stínů z mnoha světelných zdrojů, nebo deformacích scény (např. pohledem přes sklo nebo vodu)

Vertex Shader 2.0 - jde o třetí generaci T&L, implementovanou v DirectX 9.

1.3 Triangle setup

V této fázi převezme procesor vypočtené údaje z geometrie a spočítá doplňkové informace, jako je míra barevné změny mezi jednotlivými hranami trojúhelníků. Podporuje hardwarově většina akceleratorů - jen ty úplně nejstarší ne.

1.4 Texturové operace

V této fázi dochází k vlastnímu renderingu. Nejdůležitější způsoby stínování jsou:

a) flat render - vytvoří jednoduché zobrazení modelů, každý polygon je zobrazen jednou barvou. Velmi rychlé, nenáročné, ale výsledek vypadá dost hnusně.

b) Gouraud shading - plynulé vystínování polygonů, což dává mnohem realističtější výsledek. Principem je dopočítání barevného odstínu mezi všemi třemi vertexy. Toto stínování je poměrně nenáročné, a tak jej podporuje většina akceleratorů již od těch nejstarších.

c) Phong shading - náročnější a složitější metoda stínování, která dává ještě realističtější výsledek. Je vlastně jen jednodušší formou ray-tracingu, protože nebere v úvahu lom a odrazy světla. Toto stínování je výpočetně velice náročné, proto jej podporují jen moderní GPU.

d) ray tracing - nejsložitější forma renderování, při kterém počítač musí pro každý zobrazený pixel vyslat paprsek, u kterého pak simuluje lom a odrazy od povrchů, podobně jako se chová skutečné světlo. Výsledek vypadá velice realisticky, ale výpočetně je tak náročný že se dnes prakticky nepoužívá.

V této fázi jsou aplikovány pixelové efekty jako například mlha (fog), je proveden výpočet alpha kanálů, kalkulace z-bufferu a stencil bufferu, dithering.

Pixel Shader (Fragment Shader) 1.x - jedná se o programovatelnou jednotku s omezenou velikostí, implementovanou v DirectX 8 a vyšších, která umožňuje tvorbu vlastních filtrů a efektů společně s těmi které už jsou implementovány napevno (bilineární filtrování, per-pixel efekty), a dále efektně pracovat se stíny. Principem funkce pixel shaderu je zpracování viditelných částí polygonů. V pipeline se pixel shadery nacházejí až za vertex shadery, z toho plyne že ke své funkci většinou

potřebují data zpracovaná vertex shadery. Pixel shadery obsahují dvě větve - jednu vektorovou pro zpracování barev, druhou skalární (alpha) pro zpracování průhlednosti. V porovnání s vertex shadery jde o poměrně omezenější jednotku, vzhledem k nutnosti co nejefektivnějších výpočtů je počet kroků dosti omezen (většinou na několik desítek až stovek).

Pixel shadery verze 1.1 (8 maximálně použitelných instrukcí) respektive 1.2 (12 instrukcí) byly použity na kartách nVidia GeForce 3 respektive 3 Ti. Verze 1.3 (12 instrukcí) byla implementována v GeForce 4 Ti. Verze 1.4 (14 instrukcí) v rodině ATi Radeon 8500 / 8500 LE / 9100.

Pixel Shader 2.0 - implementován v DirectX 9. Tato verze umí pracovat i s desetinnými čísly, což umožňuje vypočítat realističtější barvy. Dalším rozdílem je zvýšení počtu maximálně použitelných instrukcí na 96, z toho 64 aritmetických a 32 texturových.

Tato verze shaderů je využita v kartách ATi Radeon 9500,9600,9700,9800,X800

Pixel Shader 2.0 Extended - jde o rozšířenou verzi 2.0, využívanou v kartách nVidia GeForce FX. Minimální počet využitelných instrukcí je 96, maximální 512.

Pixel Shader 3.0 - jedná se o nejnovější specifikaci, standartizovanou v DirectX 9.0c a podporovanou zatím jediným GPU nVidia GeForce 6800. Minimální počet instrukcí je 512, maximální 32768. DX9.0c zatím existuje pouze ve vývojové verzi, tudíž oficiální specifikace nejsou úplně jasné.

1.5 Raster Outputs (ROP)

Zpracování informací z pixel shaderů, které končí obrazem scény ze zvoleného úhlu a zápisem tohoto obrazu do frame bufferu. V této fázi se mimo jiné provádí antialiasing nebo komprese barev.

1.6 Vlatní vykreslení obrazu

Vykreslení obrazu z frame bufferu na monitor.

Je nutné vysvětlit rozdíl mezi klasickým (traditional based) a dlaždicovým (tile based) renderováním. Tile based renderování vykresluje pouze viditelné pixely a jen si pro ně "sahá" do paměti. Eliminuje se tak z-buffer a sníží se nároky na propustnost sběrnice. Tento způsob je poměrně neobvyklý a byla na něm založena technologie PowerVR (platforma Sega Dreamcast, PC akcelerátory ST Micro Kryo).

Pixelová pipeline může být jednoduchá, ale také zdvojená, moderní akcelerátory mají čtyři nebo osm pipelines, a poslední generace akceleratorů počínaje nVidia GeForce 6800 zavádí až 16 pipelines. Také vertexové pipelines jsou vícenásobné, v poslední generaci karet se objevuje šest vertexových pipelines pracujících nezávisle na sobě. Více pipelines umožňuje zpracovávat paralelně více dat najednou - toto se projevuje zvláště při velmi složitých scénách. Dalším přínosem více pipelines je možná aplikace více textur během jednoho průchodu - zde se využívají buď pixel shadery nebo starší multitexturing.

Výroba dnešních čipů je složitá, a velká část čipů jsou zmetky – toto výrobci využívají a takto zmetkové čipy prodávají v pomalejších sériích karet – není problém například vypnout polovinu pipelines. Toto na druhou stranu otevírá prostor pro experimenty, některé čipy se proslavily vysokou šancí na znovuaktivaci vadných pipelines (vada nemusí být příliš patrná).

U karet postavených na architektuře CineFX 2.0 (řada GeForce FX 5900/5950 a 5700) se mezi API a renderováním ještě nachází kompilátor (nVidia Unified Compiler), který upraví instrukce tak, že se místo několika instrukcí použije jediná o velké délce slova (VLIW - Very Long Instruction Word). Tento kompilátor byl dodatečně přidělán k již hotové kartě a důsledkem byla nepříliš efektivní práce kompilátoru. U nové řady GF 6800 je VLIW nahrazeno novou architekturou SIMD/MIMD (Single / Multiple Instruction Multiple Data). Pipelines posledních grafických karet se vyznačují masivním paralelismem - dokáží zpracovávat velké objemy dat a několik instrukcí současně.

2 Základní pojmy 3D grafiky a funkce 3D akceleratorů

Pixel (Picture Element) - nejmenší vykreslovatelná jednotka na obrazovce - bod.

Texel (Texture Element) - Pixel z textury, který byl vykreslen na polygon. Jediný pixel na obrazovce může být vytvořen několika texely - například "kámen v dýmu" obsahuje pixel textury kamene, ale i textury dýmu. Tento pojem se používá při měření rychlosti multitexturingu.

Voxel (Volume Pixel Element) - Trojrozměrný pixel, poměrně málo využívaný způsob vyjádření trojrozměrné grafiky. Je výpočetně poměrně náročný, ale nevyžaduje (respektive neumožňuje) akceleraci klasickými akcelerátory. Na voxelové grafice byly založeny některé starší hry od firmy Novalogic.

Vertex - Bod, který definuje kus kostry trojrozměrného objektu

Edge - Hrana spojující dva vertexy, hrany tvoří wireframe

Mesh / Wireframe - definuje tvar 3D objektu

Polygon - Uzavřený plošný útvar, který je definován minimálně třemi body, což je vlastně trojúhelník. Trojúhelníky (triangles) tvoří základ pro tvorbu 3D prostředí. Vykreslená scéna obsahuje značné množství polygonů - na jediný objekt je zapotřebí několik tisíc i několik set tisíc polygonů.

Rasterization (rasterizace) - Rozklad trojúhelníků na pixely

Tessellation (teselace) - Rozklad trojúhelníků na menší trojúhelníky

Transforms - Transformace - operace, které mění polohu, velikost nebo natočení objektů. Objekt samotný se přitom nemění, pouze se zvětšuje, zmenšuje nebo natáčí.

Deforms - Deformace - podobají se transformacím, ale mění se podoba objektu, který se kroutí, naklání nebo křiví

Clipping - proces, který slouží k odstranění polygonů, které nejsou v obraze pozorovaném hráčem - zjednodušuje se tak výpočet.

Depth complexity - Míra složitosti - udává počet objektů, které se v zorném poli vzájemně překrývají. Například scéna se zdí má míru složitosti 1, pokud před ní stojí auto, má míru složitosti 2, pokud je před ním ještě člověk, pak je 3... ovšem pokud člověk stojí vedle auta a nepřekrývá se s ním, pak je míra složitosti opět jen 2. Průměrná míra složitosti v počítačových hrách je 3.

Flat shading - každý polygon má přiřazenu barvu stínu po celé ploše; mezi jednotlivými polygony jsou vidět ostré přechody v barvě stínů. Takto

stínovaný model vypadá velmi hranatě a uměle, algoritmus je ale výpočetně nenáročný.

Smooth shading - stín na jednotlivých polygonech plynule přechází plynule ve stín sousedního polygonu. Dnes jedno z nejpoužívanějších stínování, protože je podporováno hardwarem. Vzhled objektu je mnohem přirozenější a hrany mezi polygony se ztrácejí. Je náročnější než flat shading.

Smooth texturing - stín na polygonech se kombinuje s texturou a vytváří tak realisticky vypadající model.

Texture - Textura je dvourozměrný obrázek, který se mapuje na objekt složený z polygonů a dává mu realistický vzhled. Je to vlastně bitmapa, která se vytváří kreslením, focením či scanováním a následnými úpravami. Obrázky se tvoří hlavně v následujících formátech:

a) BMP - bitmapa je velmi rozšířený formát. Je nekomprimovaný a proto velmi náročný na paměť

b) JPG - rozšířený zvláště na webu. Využívá kompresi a je určený hlavně pro fotorealistické obrázky

c) GIF - formát využívající pouze 8 bitů a 256 barev. Je komprimovaný a zahrnuje i animovaný formát. Podporuje průhlednost.

Volumetric texture - volumetrická textura = trojrozměrná textura; příkladem trojrozměrné textury je například vnitřní struktura materiálu apod.

Texture compression - Texturová komprese je postup, který umožňuje snížit paměťové nároky textur za cenu "přijatelné ztráty kvality". Obvykle nezpomaluje rychlost renderování (dekódování se děje hardwarově) a ulehčuje požadavkům na paměť a snižuje zatížení sběrnice. To je možné využít například pro vytvoření detailnějších nebo větších textur. Existuje mnoho principů komprese, mezi nejznámější patří S3TC firmy S3, FXT1 firmy 3Dfx, vektorová kvantizace (vector quantization), 4CC či CLUT (Color Look-up Table), DXT1 a DXT3 užívané v Microsoft DirectX.

Mapping (Texture mapping) - mapování je proces, při kterém se jednotlivým polygonům přiřazují textury, které dávají objektu realistický vzhled. Tento proces se dá přirovnat k zabalení trojrozměrného předmětu do kusu balicího papíru; před mapováním vypadá objekt jako by byl vyroben "z plastu". Teprve textury mu dodají vzhled dřeva, kamene nebo kovu. Klíč k dokonalému vzhledu je přitom v nedokonalosti textur. Textura musí obsahovat nepravidelnosti, "škrábance" a

podobně, potom vypadá velice dobře. Mapování vyžaduje korekci perspektivy, mipmapping a bilineární nebo trilineární filtrování

Perspective correction - Korekce perspektivy zajišťuje, aby textury byly zobrazeny ve správné perspektivě vůči pozorovateli. Textury, které nejsou zobrazeny ve správné perspektivě, jsou zkroucené, "poskakují" a "plavou", což vypadá velice podivně. Korekce perspektivy je u dnešních her důležitá, protože dosud používají dost velké trojúhelníky, na které mapují textury - jinak totiž nemohou dosáhnout většího detailu.

Mip-mapping - zlepšuje kvalitu vzdálených objektů. Mipmapy představují několik textur v různých rozlišeních, které reprezentují stejný objekt v různých vzdálenostech. Jakmile se pozorovatel přibližuje ke vzdálenému objektu, akcelerátor využije textury s vyšším rozlišením. Tato technika zajišťuje, aby vzdálené objekty nevypadaly zmateně - příliš ostře apod.

Filtering - slouží k tomu, aby textury vypadaly vyhlazeně a nebyly na nich divoce vypadající zuby. Filtrování je základní technika, která zásadně vizuálně odlišuje staré hry od moderních akcelerovalých her. Filtrování jednak odstraňuje ostré přechody mezi "kostkami" na texturách, ale také odstraňuje efekt aliasingu na vzdálených objektech.

a) point sampling - vezme namapovanou texturu a jako výslednou barvu zobrazovaného pixelu vezme barvu texelu textury, který leží přímo pod zobrazovaným pixelem. Je to jednoduché, rychlé, nenáročné na paměť, ale nevypadá to moc hezky. Hlavní smysl má pro odstraňování efektu aliasingu.

b) bilinear filtering - "průměruje" čtyři sousední texely a vyhlazuje tak hrany - pro každý vykreslovaný pixel vezme jeho čtyři okolní sousedy a vypočte průměr, který pak zobrazí. Problém je v tom, že u ploch, které nejsou kolmé na obrazovku (a těch je většina), poněkud zkresluje barvy. Díky tomu pak získávají blízko ležící objekty "rozmlžený" vzhled, podle kterého je ihned poznat akcelerovalou hru. Bilineární filtrování vyžaduje 4 texely, je tedy čtyřikrát náročnější než point sampling.

c) trilinear filtering - v podstatě se jedná o vylepšenou techniku bilineárního filtrování. Vezme dvě po sobě jdoucí úrovně mipmap, pak vypočítá bilineární filtrování na každé z nich a nakonec tyto výsledky zprůměruje. Nevýhodou je zdvojnásobení paměťové náročnosti, a přitom se výsledek od bilineárního až zas tak moc neliší.

d) anisotropic filtering - je nejpřesnější, bere v úvahu fakt, že většina ploch neleží rovnoběžně s obrazovkou a s pomocí 16 až 32 texelů dává nejpřesnější výsledky. Toto filtrování nevyužívá výhradně pravidelné plochy, ale i atypické tvary (elipsy). Náročnost je ovšem taková, že u některých starších akceleratorů jej téměř nebylo možno implementovat. Využíváno je až s příchodem posledních generací akceleratorů s podporou DX9, které disponují dostatečným výkonem.

Lighting - osvětlení. Hlavní typy osvětlení jsou:

a) light map - nejedná se o skutečné osvětlení, na objekty se pouze nanese další textura která vytvoří tmavší a světlejší odstíny (light mapa)

b) omni lights - funguje jako bodový zdroj světla - ozařuje předměty ve všech směrech

c) spot lights - bodové osvětlení osvětluje objekty podobně jako reflektor nebo hledáček - ozařuje pouze část předmětu

d) ambient light - světlo je přítomné všude a přichází ze všech směrů

e) specular highlights - speciální typ odrazů od lesklých předmětů

Level of detail (LOD) - pokud se využije fakt, že vzdálené objekty není potřeba vykreslovat tak detailně jako blízké, je možno na celou scénu aplikovat LOD. Pokud bude akcelerator vykreslovat vzdálené objekty v nižším detailu nebo je dokonce nebude vykreslovat vůbec, výrazně se tím urychlí celý proces renderingu.

Dynamic level of detail - umožňuje dynamicky měnit LOD v závislosti na vytížení akceleratoru.

Depth cueing - tato funkce umožňuje měnit odstín barvy objektu v závislosti na vzdálenosti od pozorovatele

Embossing - technika, která simuluje drsný povrch na textuře trikem. Je to myšlenkově jednoduchý postup, který lze implementovat na každém akceleratoru. Výpočetně je ale poměrně náročný a složitý na programování, protože každý akcelerator ho podporuje trochu jinak

Bump mapping - technika umožňující simulovat drsný povrch textur pomocí speciální bump mapy, která vyjadřuje boule a prohlubně povrchu. Akcelerator podporující bump mapping dokáže s pomocí údaje o poloze pozorovatele a vytvořit tak plasticky vypadající povrch. Síla bump mappingu je u dynamického osvětlení, jako je třeba pochodeň - stíny se hýbou a vytváří iluzi drsného kamene či dřeva.

Environmental bump mapping - technika umožňující na drsný povrch promítat odrazy okolí - simuluje se jím například vodní hladina, lesklý povrch kovů nebo

tekoucí rtuť. Tuto techniku podporují akcelerátory Matrox G400 a počínaje GeForce3 všechny DirectX 8 kompatibilní.

Enviromental cube mapping - technika promítání obrazu okolí na povrch tělesa, u kterého je okolí definováno pomocí kostky, která těleso obklopuje. Definováno v DirectX 7.

Fogging - mlha je nejen efektní, ale umožňuje vývojářům také omezovat horizont viditelnosti, což snižuje zátěž procesoru a akcelérátoru. Princip mlhy je jednoduchý - objekty, které jsou blízko, jsou viditelné jasně, objekty vzdálenější jsou vybledlé. Nejčastěji se používá mlha lineární (zřetelnost objektů klesá rovnoměrně se vzdáleností) a exponenciální (zřetelnost klesá mnohem rychleji a viditelný prostor je menší).

Alpha blending - metoda míchání textur, která slouží pro tvorbu mnoha vizuálních efektů, z nichž nejdůležitější je průhlednost objektů.

Alpha-channel - číslo s hodnotou od 0 do 1, které určuje průhlednost textury.

Aliasing (jaggies, stairstepping) - zubový efekt, který je patrný na rozhraní tmavých a světlých linií, zvláště pokud jsou téměř vodorovné nebo téměř svislé. Aliasing je patrný zejména v nižších rozlišeních.

Anti-aliasing - technika odstraňování aliasingu. Principem je dopočítání odstínu mezi objektem a pozadím - zuby tak opticky zmizí. ATi tuto technologii nazývá Smoothvision. Používají se dvě metody počítání:

a) supersampling - každá scéna se renderuje několikrát, výsledný obraz je tvořen prolnutím těchto dílčích obrazů.

b) multisampling - využívá z-bufferu, ze kterého využije pouze postižené hrany, čímž dochází k úspoře výkonu.

Soft shadows - technika sloužící k rozpijení ostrých hran vržených stínů - nemá nic společného se stínováním polygonů

Motion blur - efekt rozmazání obrazu, který vzniká při rychlém pohybu a zvyšuje dojem filmovosti

Stencil buffer - je to další buffer velikosti obrazovky, není předem dáno pro jaké účely se má použít. Jeho hlavní smysl je v kombinaci s programovatelnými GPU a víceprůchodovým vykreslováním, využívá se například při počítání vržených stínů apod. Využití stencil bufferu se v budoucnu bude zvyšovat.

Depth of Field - každé oko či kamera ve scéně zaostřuje na určité místo a popředí či pozadí je rozmazané. Tato funkce vytváří podobné efekty ve 3D scéně a zvyšuje tak dojem filmovosti.

Z-occlusion culling - jedná se o předvídání z-informace. K vyhodnocení, zda pixel bude vidět nebo ne, dochází před samotným výpočtem pixelu, tímto se dá výrazně zrychlit rendering. Implementováno v DirectX 8.

Lossless Z-compression - přenos z-informace zabere velkou část šířky paměťové sběrnice. DirectX 8 akcelerátory dokáží tuto informaci bezztrátově komprimovat, obvykle v poměru 4:1 a lepším. Tato funkce není tak účinná jako tile-based rendering, ale i tak je nárůst výkonu značný. U ATi se tato technika nazývá HyperZ.

Higher order surfaces / primitives - technika modelování 3D objektů implementovaná v DirectX 8. 3D modely nemusí být definovány pouze jako soustava polygonů, ale jeho jednotlivé části mohou být popsány Bézierovými křivkami a B-spline křivkami procházejícími několika kontrolními body. Takto je možné vytvořit i velmi složité tvary bez užití stovek polygonů a vertexů. Toto se děje hardwarově pomocí vertex shaderu, tudíž nedochází k dalšímu zatížení akcelerátoru či sběrnice.

Existují dva druhy HOS:

a) RT-patches - starší, podporovány v nVidia GeForce3 a Matrox Parhelia. Výrazně se neuchytily, protože když je akcelerátor nepodporuje, nevykreslí nic.

b) N-patches - novější, poprvé se objevily s příchodem ATi Radeon 8500, firma ATi je prezentovala pod názvem TruForm. Funkčně jsou podobné RT-patches, až na to že pokud je akcelerátor nepodporuje, vykreslí se prostě hranatý povrch. Principiálně jde o rozdělení trojúhelníků na menší (teselace), a dopočítání jejich normálových vektorů, čímž se docílí zaoblenějšího povrchu.

Keyframe animation, keyframe interpolation - automatické dopočítávání mezisnímků (jedna z funkcí vertex shaderu)

Výškově omezená mlha - jedná se o reálnější zpracování mlhy než původní mlha definovaná vzdáleností od pozorovatele (jedna z funkcí vertex shaderu)

3D Morphing - jedná se o počítání deformací objektů v reálném čase (jedna z funkcí vertex shaderu)

Hardware displacement mapping - jde o vylepšenou techniku stínování podobnou bump-mappingu, implementovanou v DirectX 9. Principem je nanesení

další textury v odstínech šedi, které určují "hloubku" textury - touto technikou lze efektivně modelovat prostorový dojem.

Ghost Shader - technika mapující na objekty barevné světlo místo textur.

Ultra Shadow - technologie umožňující pomocí z-indexů ohraničit pouze část scény která bude stínována, zbytek stínován nebude, což může výrazně zrychlit vykreslování a ušetřit výpočetní výkon. Je součástí architektury nVidia CineFX 2.0 a výš.

3 Základní části 3D akcelérátoru a základní pojmy

Videochipset / GPU - procesor, který realizuje grafické a akcelerační funkce. GPU je speciální jednotka, sloužící k urychlení výpočtů transformací a osvětlení. Dnes se používá termín GPU jako označení pro celý videočip.

Video BIOS - poskytuje základní funkce pro ovládání videokarty

Paměť - je v ní uložen framebuffer

Interface - zpočátku PCI, dnes už prakticky výlučně AGP, v nejbližší budoucnosti přechod na PCI-Express

16ti, 24ti a 32ti bitové barvy - barva každého pixelu může být uložena s různou přesností, neboli barevnou hloubkou. Nejstarší karty podporovaly barevné hloubky 1 bit (černá a bílá), 4 bity (16 barev), 8 bitů (256 barev). Přechody mezi odstíny jsou ale velmi ostré, a proto akcelérátory používají mnohem větší barevnou hloubku. Základní je 16 bitů (65536 barev), kterou podporují všechny akcelérátory již od těch nejstarších. Větší barevnou hloubku nabízejí všechny moderní akcelérátory (poslední pouze 16 bit bylo 3Dfx Voodoo3). 24 bitová i 32 bitová barevná hloubka nabízí 16,7 milionu barevných odstínů, což už je za hranicí barevné rozlišovací schopnosti oka. Tato kvalita zobrazení se označuje jako fotorealistická. Rozdíl mezi 24 bitovou a 32 bitovou grafikou je v rychlosti zpracování a paměťové náročnosti. 32 bitová grafika zabírá o čtvrtinu více paměti, ale rychleji se s ní pracuje.

Dithering - Technika prokládání obrazu tečkováním tak, aby opticky vznikla iluze většího počtu barev. Tato technika dává vizuálně horší výsledky než kdyby byl obraz skutečně tvořen vyšším počtem barev. V posledních letech se již příliš nepoužívá, protože ve hrách se využívá 32 bitové renderování.

Frame buffer - paměť, která obsahuje několik obrazů - front buffer, back buffer, z-buffer. Velikostí frame bufferu je omezeno rozlišení, kterého je akcelérátor schopen dosáhnout. Například akcelérátor se 4MB paměti dovede v 16 bitech zobrazit obraz 1600x1200, který vyžaduje 3,75MB. Pokud ale program používá z-buffer, bude velikost zobrazení omezena na 800x600x6 (2 byte pro front buffer, 2 byte pro back buffer, 2 byte pro z-buffer), dává 2,75 MB. Akcelerace v 1024x768 by vyžadovala 4,5 MB paměti, což na akcelérátoru s 4MB paměti nemůže fungovat. Na dnešních kartách je osazeno až 256 MB paměti, což stačí na akceleraci i ve vysokých rozlišeních. Tato velká paměť se využívá při dalším zpracování obrazu (antialiasing apod.).

Front buffer - paměť v níž je uložen právě viditelný obraz

Back buffer - paměť v níž je obraz, na němž akcelerátor právě pracuje. Po dokončení obrazu se front buffer a back buffer přehodí. Pokud se používají pouze dva, nazývá se tato operace double buffering.

Z-buffer - paměť, která obsahuje údaj o "hloubce" zobrazeného pixelu. Normálně jsou souřadnice zobrazených pixelů udávány pouze v matici X,Y - třetí souřadnice může být uložena v z-bufferu, ten umožňuje snadno zjistit který pixel je "blíže" a který "dále", což se efektivně využívá například při modelování mlhy. Moderní akcelerátory počínaje GeForce 3 využívají efektivně komprese z-bufferu.

W-buffer - jde o lineární alternativu k z-bufferu. Samotná bitová reprezentace z-informace v z-bufferu není lineární, ale objekty které jsou ve scéně blíže využívají více bitů a objekty které jsou dále využívají méně bitů. V některých aplikacích toto může působit problémy, proto se využije W-buffer, který ukládá z-informaci vždy ve stejném počtu bitů.

Blit rate - blitting je proces, při kterém se přesunuje velký blok dat z hlavní paměti počítače do paměti videokarty. Rychlost tohoto přenosu se nazývá blit rate.

Fill rate - rychlost, udávající počet vykreslených pixelů za sekundu. Dříve jeden z nejdůležitějších parametrů akcelerátoru, který silně souvisí se schopností dosahovat velkého množství snímků za sekundu - karty s nízkým fill rate je prostě nedokážou vykreslit. Dnes již tento parametr ustoupil do pozadí, dnešní karty jsou dostatečně výkonné aby dokázaly vykreslit složité scény, a výrobci se soustředují spíše na optimalizace výpočtů. Fill rate může být udávána jako dvě hodnoty - jednak pixel fill rate a jednak texel fill rate - texel fill rate je násobkem počtu texturovacích jednotek a pixel fill rate. Tento parametr je důležitý při multitexturingu (jeden pixel může být vytvořen z více texelů).

Frame rate - rychlost, s jakou je akcelerátor schopen vykreslovat snímky za sekundu.

Refresh rate - rychlost, s jakou videokarta obnovuje obraz na obrazovce monitoru. Jde o zcela jinou věc než je frame rate.

Dlouhou dobu byla vedena debata o tom, zda je šedesát snímků za sekundu opravdu lepší než třicet. Zdá se, že doopravdy ano, ale vysvětlení je mnohem složitější než tom, co je uvedeno v základních knížkách o biologii a psychologii vnímání.

Za absolutně nejnižší obnovovací frekvenci je považováno 24 snímků za sekundu, což je obnovovací frekvence filmu v kině. Televize má o něco vyšší obnovovací frekvenci - 30 snímků za sekundu. Teoreticky oko vyšší počet snímků nerozezná, je tu ovšem jedno "ale".

Nízký obnovovací kmitočet v kině stačí hlavně proto, že film je promítán ve tmě a snímky se mění jeden po druhém. Obraz hodně září a vytváří na sítnici efekt "afterimage" (zbytkový obraz). Protože se filmové obrazy mění najednou, projeví se efekt zbytkového obrazu v plné síle a vjem je mnohem plynulejší. Dalším faktorem je motion blur - na jednotlivých snímcích filmového pásku jsou pohybující se předměty rozmazané tak, že jsou skoro k nepoznání. Ale když je film puštěn plnou rychlostí, oko vnímá obraz jako ostrý a plynulý.

Televize, video i DVD používají kmitočet 60 Hz, ale obnovovací kmitočet obrazu je pouze 30 snímků za sekundu, takže každý obrázek je obnoven dvakrát. Dvojnásobek je v tomto okamžiku důležitý - při obnovovacím kmitočtu 72 a neceločíselném násobku by byl dojem z přehrávání méně plynulý. Vyšší kmitočet je nezbytný, protože televize zobrazuje v prostředí, kde je více světla. Jednotlivé snímky nejsou vykreslovány najednou, ale po řádcích elektronovým dělem. Díky přítomnosti motion bluru ovšem oko vnímá dvě obnovení téhož obrázku jako dva různé obrázky.

Počítačové zobrazení je naproti tomu zcela jiné - jednotlivé snímky jsou dokonale ostré, bez motion bluru. Chybějící motion blur způsobuje, že obraz vypadá méně spojitě - a to zvláště ve chvílích kdy se framerate mění. Lidské oko se lépe adaptuje na nižší obnovovací kmitočet, který je ale stejnoměrný, než na kmitočet který se mění. Za maximální možnou hranici rozlišení "pomalejšího" a "rychlejšího" framerate bez motion bluru se považuje zhruba 75 snímků za sekundu.

Single buffering - pokud je nastaven single buffering, akcelerátor vykresluje obraz přímo do front bufferu - paměti ze které načítá RAMDAC. Nevýhodou je, že zobrazovaný obraz "teče" (tearing), protože akcelerátor mění obsah bufferu RAMDACu. Efekt tearingu je tím horší, čím víc se odlišuje refresh rate od frame rate - ať už je frame rate rychlejší nebo pomalejší než refresh rate.

Double buffering - tento způsob bufferingu eliminuje efekt "tečení". Zatímco RAMDAC vykresluje obsah front bufferu, akcelerátor vykresluje obraz do back bufferu. Jakmile je obraz v back bufferu dokončen, čeká videokarta na VSYNC a v tomto okamžiku se front buffer a back buffer přehodí. Nevýhodou tohoto přístupu je

právě čekání na VSYNC, kdy akcelerátor stojí a snižuje tak frame rate. Vlivem "aktivního čekání" je totiž většinou nižší frame rate než refresh rate - v nejhorším případě akcelerátor dokončí frame těsně po zahájení vykreslování obrazu RAMDACem a musí s výměnou bufferů čekat na další VSYNC, což může průměrnou rychlost frame rate stáhnout dost hluboko pod refresh rate.

Triple buffering - nejpokročilejší technika využívá tři bufferů - jeden je front buffer, ze kterého RAMDAC kreslí obraz na monitor. Rozdíl oproti double buffering nastane v okamžiku, kdy akcelerátor domaluje celý obrázek, ale ještě není možné vyměnit front buffer a back buffer, protože RAMDAC je při vykreslování kdesi "uprostřed" obrazovky a VSYNC nastane až poměrně za dlouho. Zde ale problém nenastane - akcelerátor začne ihned vykreslovat do druhého back bufferu, takže výrazný pokles frame rate nenastává. I tak je ale frame rate limitována refresh rate, videokarta nemůže vyměňovat buffery častěji než je refresh rate. Paměťově jde o nejnáročnější způsob renderování.

VSYNC On / Off - při zapnuté vertikální synchronizaci vyměňuje akcelerátor front buffer a back buffer jen v okamžiku, kdy se elektronový paprsek vrací z pravého dolního do levého horního rohu obrazovky - tím se zabraňuje, aby obraz "plaval". Zároveň se ale snižuje frame rate maximálně na refresh rate, ale často i o dost níže.

4 Sběrnice používané grafickými kartami

ISA (Industry Standard Architecture) - stará ISA sběrnice neumožňovala použít výkonné grafické karty, proto byla časem rozšířena o VLB, vylepšena na MCA nebo EISA, a nakonec na plné čáře překonána PCI.

VESA Local Bus (VLB) - jednalo se o nadstavbu klasické ISA sběrnice, kterou využily úplně nejstarší 3D akcelerátory (3D Blaster VLB). Principiálně byl VLB přímo rozšířením paměťové sběrnice CPU. Neprosadila se jednak kvůli těsnému svázání frekvence s frekvencí CPU - zde byl problém s prudce narůstajícím rušením na vyšších frekvencích (málokteré zařízení dokázalo fungovat nad 40 MHz), jednak kvůli rychlému nástupu PCI.

PCI - dnes již zastaralá sběrnice, která se prosadila na trhu přibližně před deseti lety, když nahradila v té době nejrozšířenější ISA, EISA a VLB sloty. Původní frekvence PCI byla 33 MHz, v revizi 2 byla zdvojnásobena na 66 MHz. Její maximální přenosová rychlost je 133 MB/s (266 MB/s ve druhé revizi), kterou navíc sdílejí všechna zařízení včetně řadiče IDE, zvukové karty atd. PCI dobře posloužilo pouze prvním generacím grafických karet, než bylo zavedeno AGP. V současnosti začíná nástup nové sběrnice PCI-Express, která na základních deskách nahradí jak PCI, tak AGP.

AGP (Accelerated Graphics Port) - Speciální vysokorychlostní propojení mezi systémovým chipsetem a grafickou kartou. AGP prošlo vývojem od verze AGP 1x (266 MB/s) přes verze 2x (528 MB/s) a 4x (1 GB/s), až po v současné době standardní 8x (ve skutečnosti však není patrný rozdíl ve výkonu AGP 4x a 8x z důvodů omezení architekturou samotných karet, rozdíl je spíše teoretický a výhodný hlavně pro marketingová oddělení výrobců karet). Toto označení ve skutečnosti znamená, kolik operací za jeden pracovní cyklus je možno provést. Základní frekvence AGP je 66 MHz, ve verzi 8x se frekvence efektivně zosminásobuje (provádí se 8 operací za jeden pracovní cyklus). AGP umožňuje používat velké textury a je schopné v jisté míře kompenzovat nedostatek lokální paměti akcelerátoru. Nabízí také další funkce jako zařazování instrukcí do fronty nebo texturování z hlavní paměti počítače. Některé staré karty ovšem tento AGP texturing nepodporovaly (3Dfx Voodoo3). Původně mělo AGP ušetřit drahou paměť na kartě, dnes je jeho smysl spatřován jednak ve větší špičkové přenosové kapacitě než má PCI (hodí se v okamžicích kdy se náhle vymění všechny textury ve scéně), a jako

"záchranný pás", díky němuž nemusí uživatel často kupovat akcelerátor s větší pamětí - AGP dokáže s jistou výkonovou penalizací simulovat velkou paměť. V současné době začíná nástup nové sběrnice PCI-Express, která na základních deskách nahradí jak PCI, tak AGP.

PCI-X - jde o projekt IBM, HP a Compaq, za účelem nahradit starou PCI sběrnici. Druhým cílem bylo přesvědčit Intel, aby nový standard přijal (ten však nesouhlasil a zahájil vývoj 3GIO, které nakonec souboj marketingově vyhrálo). Sběrnice podporuje ECC, a je zaměřena na zařízení s extrémními datovými přenosy. Původní verze PCI-X (1.0) měla frekvenci 66 (PCI-X 66MHz) a 133 MHz (PCI-X 133MHz). Pro zařízení pracující na 64-bit technologii to znamená přenosovou kapacitu 1 GB/s. Dnešní revize 2.0 nabízí rychlejší PCI-X 266 MHz a PCI-X 533 MHz, přičemž vývoj i přes tvrdou konkurenci ze strany PCI-Express pokračuje.

PCI-Express - jedná se o přejmenovanou sběrnici nové generace, vyvinutou firmou Intel pod krycím jménem 3GIO. Narozdíl od PCI a AGP sběrnice je PCI-Express sériová sběrnice, proto je jednodušší, levnější a snadno upgradovatelná; díky sériovému rozvržení obsahuje daleko méně "vodičů". Minimálně softwarový model by měl být zpětně kompatibilní se sběrnici PCI. Narozdíl od PCI tato sběrnice nepodporuje bus-mastering, ale zajišťuje rovnou přímé připojení díky point-to-point architektuře. PCI-Express je navržena s ohledem na snadnou upgradovatelnost - je možno zvýšit frekvenci i snadno díky sériovému designu přidat "vodiče". Další výhody této sběrnice jsou nižší spotřeba energie, možnost hot-swap i hot-plug, kontrola chyb a integrity dat. V současné době již jsou uváděny na trh první základní desky s touto sběrnici (s chipsety Intel Grantsdale a Alderwood), zajímavost je že již na této první generaci desek nové rozhraní úplně nahradilo AGP. Ve verzi PCI-Express X16 má datovou propustnost 4 GB/s, což je dvojnásobek AGP 8x. V budoucích verzích se počítá se zvýšením rychlosti na 16 GB/s a výš.

RAMDAC (Random Access Memory Digital-Analog Converter) - zařízení, které převádí obsah front framebufferu na analogový výstupní signál. V paměti akcelerátoru je obraz uložený v digitální podobě, ale většina monitorů je analogových - proto musí RAMDAC mnohokrát za sekundu načíst obsah paměti a překódovat ho na videosignál. Kvalita obrazu závisí na tom, zda je RAMDAC zabudovaný přímo na čipu nebo externí (ty bývaly lepší, ovšem pak se rozdíly smazaly a dnes se používají pouze integrované RAMDAC), a také na tom, na jaké frekvenci pracuje - 400 MHz RAMDAC je ve vyšších rozlišeních lepší než 250 MHz

RAMDAC. V dnešní době už mají některé karty pouze digitální DVI výstup; zde RAMDAC není potřeba, protože digitální data z videopaměti jsou posílány přes digitální výstup přímo do monitoru (většinou digitální panel TFT)

AGP Graphics Aperture Remapping Table (GART) - zařízení, umožňující efektivnější přenos dat přes AGP. Princip je rozdělení velkých bloků na menší, se kterými se manipuluje rychleji - zde je souvislost s DIME. GART pak předkládá tato data grafické kartě, jako by byla vzata z lokální paměti na kartě.

AGP Direct Memory Execute (DIME) - termín označující využití systémové paměti rozhraním AGP - systémová paměť RAM se tak chová jako by byla na kartě.

AGP Aperture Size - souvisí se schopností AGP simulovat grafickou paměť pomocí lokální paměti RAM (DIME). Jedná se o velikost paměti, která se dodatečně vyhradí pro grafické operace, když paměť na grafické kartě nestačí. Nevýhodou tohoto řešení je nižší rychlost (operační paměť je pomalejší než paměť na kartě), výhodou naopak to, že se tato paměť použije pouze když je třeba.

AGP Fast Writes (FW) - tato technologie umožňuje přenášet data z procesoru přímo do grafické karty, kdežto bez využití Fastwrites jdou data přes operační paměť. Tato funkce rozumně funguje pouze na grafických kartách nVidia, na konkurenční ATi jsou hlášeny spíše problémy. Jako taková je tato funkce zajímavá, nicméně v praxi se její přínos příliš neprojevuje.

AGP Sideband Addressing (SBA) - jedná se o zařazení až 32 instrukcí do fronty AGP, čímž se minimalizují prodlevy při přenosu. Opět jako u Fast Writes spíše teoretická záležitost, která v praxi způsobuje poměrně často problémy.

Integrované grafické čipy - dnešní integrované čipy jsou vnitřně napojeny na AGP, tudíž disponují teoreticky stejným výkonem jako jejich protějšky na kartách. Jejich nevýhodou je sdílení systémové paměti, která je několikanásobně pomalejší než rychlé paměti používané na kartách, a má také užší sběrnici (většinou 64-bit, v duálním režimu 128-bit. Dnešní nejvýkonnější karty mají 256-bit paměťovou sběrnici).

UMA - Unified Memory Architecture - jedná se o sdílení systémové paměti integrovanými zařízeními (grafická karta, zvuková karta apod.)

5 Typy pamětí používaných v akcelerátorech

FPM DRAM (Fast Page Mode) - Nejjednodušší, nejpomalejší a nejzákladnější paměť, využívaná v nejstarších akcelerátorech. Umožňuje pouze jedno čtení nebo zápis najednou a má malou rychlost a šířku přenosu.

EDO DRAM (Extended Data Out) - Dovoluje zahájit čtecí cyklus před tím, než skončil předchozí. Tím se získává malé vylepšení výkonu - mezi 5 a 20%. Tyto paměti využívaly například karty 3Dfx Voodoo2.

VRAM (Video RAM) - původní jednoportové paměti umožňovaly pouze jedno čtení nebo zápis, což způsobilo kolize mezi čipsetem, který do front bufferu zapisuje, a RAMDACem, který z něj čte data. Řešení se objevilo v podobě paměti VRAM, která má porty dva a dovoluje tak současný zápis i čtení. Toto řešení je výrazně rychlejší, ale také dražší než klasické paměti DRAM.

WRAM (Window RAM) - vylepšená paměť VRAM, má o 25% vyšší propustnost než VRAM a je levnější. Byla používána v akcelerátorech Matrox.

SDRAM (Synchronous DRAM) - pracuje na mnohem vyšších kmitočtech než DRAM. Je pouze jednoportová, ale vysoká rychlost pamětí tento nedostatek kompenzuje. Oproti EDO DRAM získává asi 20% rychlosti navíc díky vylepšené vnitřní architektuře. Dříve byla používána na většině akceleratorů.

SGRAM (Synchronous Graphics DRAM) - jde o SDRAM vylepšenou o některé speciální funkce, například Block Write a Masked Write, dovolující rychleji mazat framebuffer a rychleji modifikovat požadované pixely. Obvykle se hůře přetaktovává než SDRAM, je to dáno složitější vnitřní strukturou. Byly používány společně s SDRAM, i když kvůli vyšší ceně nakonec nepřevážily.

DDR RAM (Dual Data Rate RAM) - jde o vysokorychlostní paměť SDRAM, která je schopná pracovat s daty dvojnásobnou rychlostí než SDRAM - využívá při přenosu vzestupnou i sestupnou hranu hodinového signálu. Dnes nejpoužívanější paměti na akcelerátorech, u posledních generací však už i ony přestávají stačit a v budoucnu budou nejspíše úplně vytlačeny paměťmi GDDR-III.

GDDR-II - nová generace pamětí DDR, kvůli vysoké ceně využívána zatím pouze nejvýkonnějšími kartami dneška. Jedná se vlastně o paměti DDR-II uzpůsobené pro použití v grafických kartách. Její výhody jsou nižší pracovní napětí a schopnost dosáhnout vysokých frekvencí (frekvence prvních generací se pohybují kolem 500 MHz DDR - efektivně 1 GHz). Jsou zatím použity na kartách nVidia

GeForce FX 5800 a 5700 Ultra a některých ATi Radeon 9800 XT. Zřejmě bude využita jako jakýsi mezistupeň mezi DDR SDRAM a GDDR-III.

GDDR-III - novější konkurence GDDR-II, je pinově kompatibilní s GDDR-II. Začíná být osazována na nejnovější generaci grafických karet počínaje GeForce 6800, a zpětně i na mainstreamových GeForce FX 5700 Ultra, a s jejich podporou počítá také ATi, zřejmě jde o grafickou paměť blízké budoucnosti.

GDDR-IV - bude se jednat o další vylepšení GDDR-III, ovšem využita bude nejdříve u přespříští generace grafických akceleratorů.

U paměti je nejdůležitější údaj o šířce paměťové sběrnice - dnešní karty mají standardně 64, 128 nebo 256ti bitovou paměťovou sběrnici. Tato hodnota je důležitá, protože se od ní odvozuje celková paměťová propustnost karty. Dalším důležitým parametrem je časování, hodnota udávaná v nanosekundách (čím nižší hodnota, tím vyšší je pracovní frekvence paměti).

6 3D programovací rozhraní (API) a jazyky

6.1 3D programovací API

Microsoft Direct3D

- vyvinula jej společnost Microsoft jako součást balíku DirectX v roce 1995 v souvislosti s nástupem Windows 95. Direct3D bylo od počátku navrženo pro využití v herní oblasti, a jeho největší výhodou oproti OpenGL je nižší náročnost na programátory. Po pádu 3Dfx Glide se Direct3D prosadilo jako prakticky monopolní 3D API, drtivá většina her dneška je postavena na Direct3D. Nevýhodou oproti OpenGL byla naopak nižší flexibilita při programování, i když tento rozdíl se během vývoje postupně smazával; dnes jsou tato rozhraní víceméně rovnocenná. Vývoj Direct3D v souvislosti s DirectX:

DirectX5 - obsahovalo z herního hlediska nejstarší použitelnou verzi D3D. Obsahovalo rutiny pro softwarové transformace a osvětlení. Brzy po uvedení přestalo kvůli ostrému konkurenčnímu boji mezi výrobci 3D karet stačit a Microsoft začal pracovat na nové verzi.

DirectX6 - bylo uvedeno v létě 1998. Podporovalo multitexturing, antialiasing a bump-mapping.

DirectX7 - uvolněno v roce 1999. Nejdůležitější inovací byla podpora hardwarového T&L.

DirectX8 - vyšlo v roce 2001. Hlavní novinkou byly vertex a pixel shadery, zároveň došlo k velkému zjednodušení a zrychlení celého rozhraní, ovšem na úkor zpětné kompatibility.

DirectX9 - zatím poslední verze vyšla v roce 2002. Z programátorského hlediska je nejdůležitější fakt, že konečně umožňuje výpočty v plovoucí desetinné čárce, což znamená teoreticky fotorealistickou grafiku.

SGI OpenGL

- bylo vyvinuto společností Silicon Graphics v roce 1992 za účelem renderingu CAD. Jeho předchůdcem bylo rozhraní SGI IrisGL. Původně bylo OpenGL využíváno v průmyslové a obchodní sféře, ale s nástupem Windows 95 a her se stalo též herním API. Největší zásluhu na prosazení OpenGL má firma iD Software a její

hlavní programátor John Carmack, který vydal první akcelerovanou hru Quake právě na OpenGL, a dodnes jsou všechny hry iD Software vyvíjeny primárně pro OpenGL. Hlavní výhodou OpenGL oproti konkurenčnímu Direct3D je v platformové nezávislosti - OpenGL je prakticky jedinou možností jak renderovat v newindowsových systémech (podpora OpenGL obsahuje např. OS Linux), ovšem z důvodů větší složitosti jej používá jen malá skupinka programátorů. V současnosti je poslední verze 2.0, která se funkčně vyrovná posledním DirectX.

3Dfx Glide

- v počátcích 3D akcelerace se jednalo o nejrozšířenější 3D API, které se rozšířilo díky kartám Voodoo, vydaným v roce 1996. Podstatou bylo silně zjednodušené a upravené rozhraní OpenGL, toto ovšem umožňovalo velmi jednoduché programování. Nevýhodou bylo omezení na karty 3Dfx, a proto s úpadkem a nakonec krachem 3Dfx skončila i vláda Glide.

S3 MeTaL

- nepřiliš rozšířené rozhraní, podporované kartami S3 Savage 3D a lepšími. Hlavního rozšíření doznalo ve hrách založených na enginu Unreal. S odstoupením S3 od 3D grafiky a odkoupením firmou VIA skončily i pokusy zařadit MeTaL mezi standardní API.

6.2 Programovací jazyky HLSL (High Level Shading Languages)

Protože zpočátku bylo nutné GPU programovat pomocí nízkoúrovňového assembleru, byly vytvořeny programovací jazyky vyšší úrovně. Prvním z nich bylo Cg od nVidie, po ní přišla s vlastním řešením firma ATi. Nakonec byly tyto jazyky standardizovány Microsoftem pod názvem HLSL a zahrnuty do DirectX 9. Tyto jazyky jsou jakýmsi prostředníkem mezi GPU a API, přenáší funkce GPU do programovacího jazyka.

Cg je odvozeno z jazyka C, má podobnou syntaxi i deklarace. Výhody oproti assembleru jsou zřejmé - rychlejší a efektivnější vývoj a ladění, přenositelnost, a optimalizace proveditelná kompilátorem. Program napsaný jazykem Cg je schopný detekovat možnosti hardwaru a když zjistí, že grafický čip je plně kompatibilní se shaderem, může nadefinovat DirectX tak, aby dokázala použít více funkcí. V praxi to

znamená, že pokud program dokáže pozměnit DirectX podle schopností grafické karty. Velká výhoda tohoto systému je v tom, že pokud je hra naprogramována tímto způsobem, přizpůsobuje se sama novějšímu hardwaru bez nutnosti upgradu či aktualizace. Cg je podporováno více platformami (Windows, Linux, Mac, Xbox).

Cg se skládá z několika částí: Je to programovací jazyk samotný, dále Cg Compiler, Cg Standard Library, Cg Runtime Libraries, Cg Browser 4.0 a Cg Manual. Celý balík je volně ke stažení na stránkách nVidie.

Program, který musel být dříve zapsán dvaceti řádky instrukcí, je nyní možné přepsat do dvou krátkých řádků v Cg. Je tedy velice podobný klasickému C/C++ a mnoho od něj přebral. Velice dobrá věc je také možnost použití více shader programů na jeden objekt

6.3 3D wrappery

Jedná se o programy, umožňující převádět volání jednoho 3D API do druhého. Největší rozmach wrapperů nastal v době úpadku 3Dfx, kdy umožňovaly spustit Glide-only hry na kartách podporující pouze Direct3D nebo OpenGL. Dnes je jejich význam minimální, protože všechny dnešní akcelerátory podporují Direct3D a OpenGL, na jejichž základě jsou programované všechny aplikace.

6.4 3D benchmarky

Benchmark je program, jehož účelem je otestovat rychlost - v tomto případě grafické karty - nebo některé další vlastnosti jako například kvalitu obrazu. Existují dva druhy benchmarků - syntetické a reálné.

a - Syntetické - jde o programy speciálně konstruované k testování rychlostí grafických karet. Jejich tvůrci se většinou snaží využít všech možných funkcí, které nabízí nejmodernější generace grafických karet, a neřídka se stává, že nové benchmarky se dají s rozumnou rychlostí spustit až na kartách které jsou o generaci novější než samotný benchmark. Výhodou syntetických benchmarků je, že se dají použít jako univerzální testovací nástroj pro celý sortiment grafických karet na trhu. Nevýhodou je, že benchmarky přece jenom představují jakýsi "ideální stav", a výsledky v konkrétních reálných aplikacích se mohou někdy i výrazně lišit.

Nejpoužívanějším benchmarkem současnosti je 3DMark 2003, vyvinutý společností Futuremark.

b - Reálné (herní) - v tomto případě se jako test rychlosti využije konkrétní hra - většina špičkových her má v sobě obsažen test rychlosti, čítač fps nebo něco podobného. Hlavní výhodou proti syntetickým benchmarkům je fakt, že test v konkrétní hře přesně ukáže, jak se bude určitá grafická karta v této hře chovat. Problémem naopak je, že tyto testy jsou závislé na způsobu naprogramování hry, a test v jiné hře může vykazovat jiné výsledky.

Víceméně vždycky se jako herní benchmark používají hry, které využívají nejmodernější technologie - dnes jsou těmito hrami Unreal Tournament 2004, Doom 3 a Far Cry.

Na CD přiloženém k této práci jsem umístil instalační balíky benchmarků rozšířených dnes i v minulosti. Přidal jsem též ukázky z výše zmíněných her – na screenshotsch je poměrně dobře vidět, jak technologie pokročila od nejstarších her.

7 Technické specifikace grafických karet

7.1 Nultá generace (první akcelerátory) (1995-1996)

Rendition Verité 1000

Jeden z prvních použitelných 3D čipů. Jako první podporoval triangle-setup, ploché a Gouraudovo stínování, bilineární i trilineární filtrování, alpha-blending, mip-mapping, z-buffering a dokonce i anti-aliasing. Nejznámější kartou byl Creative Labs 3D Blaster.

Nevýhodou byl velice pomalý VGA mód, neakcelerované hry byly fakticky nehratelné.

Technické parametry: VLB nebo PCI sběrnice, 4 MB EDO-RAM, 135 MHz RAMDAC

nVidia NV1

Karty obsahovaly i integrované 32-kanálové, 16-bitové prostorové audio, což v té době bylo něco nevídaného. Podporován byl gamepad z konzole Sega Saturn. MIDI dokázalo využít až 6 MB systémové paměti.

Karta však pracovala s kvadratickými texturami; zanedlouho po jejím vydání Microsoft uvolnil DirectX 1.0, který používal výhradně polygony, a tím se karty s čipem NV1 staly neprodejnými.

Matrox MGA1064SG

Čip byl určen pro low-end trh, nepodporoval bilineární filtrování, funkce pro práci s texturami byly omezené. Karty Matrox Mystique s tímto čipem se však prodávaly v nádherné krabici, a to nalákalo mnoho zákazníků.

Technické parametry: 170 MHz RAMDAC, 50 MHz jádro, 100 MHz SGRAM

Matrox MGA1164SG

Vylepšená verze MGA1064SG.

Technické parametry: 220 MHz RAMDAC, 60 MHz jádro, 100 MHz SGRAM

Videologic PowerVR PCX1

Filozofie PowerVR spočívala v tile-based renderingu, což zrychlilo práci s pamětí. Nepodporoval však bilineární filtrování, zatěžoval CPU, vzhledem ke koncepci renderování chyběla podpora z-bufferu a nebyl stabilní v Direct3D. Ve svém nativním API SGL byl naopak velice rychlý a kompatibilní, navíc s postupem času na rychlejších procesorech došlo k výraznému zvýšení výkonu a snížení ceny.

Technické parametry: 66 MHz jádro, 66 MHz 32-bit SDRAM, 0.5 mikronový výrobní proces, 264 MB/s paměťová propustnost

S3 Virge

Masově rozšířený čip, přestože jako novinka při svém uvedení stál neuvěřitelných 25000 Kč. Virge je vlastně S3 Trio 64V+ doplněné o 3D. Osazované byly dvě verze - Virge (2 nebo 4 MB RAM, 135 MHz RAMDAC) a Virge VX (až 8 MB RAM, 220 MHz RAMDAC). Čip uměl spoustu funkcí, včetně bilineárního filtrování, celkově vzato však byl neuvěřitelně pomalý, zvláště ve verzi Virge. Zajímavostí je bezproblémový chod pod Windows XP, první podporované hry byly Tomb Raider, Descent 2, Terminal Velocity, MechWarrior 2.

ATi 3D Rage

První 3D akcelerátor firmy ATi, jádro bylo založené na slavném ATi Mach64, ale malá paměť limitovala použití na 1024x768-16bit při 75 Hz. Podporován byl mip-mapping, gourard shading, alpha blending, perspective correct texture mapping.

Technické parametry: 1 nebo 2 MB EDO DRAM

ATi 3D Rage II

Vylepšená verze 3D Rage, podporoval z-buffering, bilineární a trilineární filtraci, paměťová propustnost se zdvojnásobila, výkon zvýšil až o 50%. Ve verzi Rage IIc podporoval AGP, měly až 8 MB SDRAM, SGRAM nebo EDO DRAM.

3DLabs Gigi

První čistě 3D čip od 3DLabs, neuměl bilineární filtrování, a celkově chudá výbava a sběrnice VLB ho odsoudila k zániku.

3DLabs Permedia

Čip už podporoval bilineární filtrování, z-buffering, anti-aliasing, mip-mapping a více podporovaných API. Sběrnice byla PCI a zajímavostí je použitý GLINT Delta procesor.

7.2 První generace (1996-1997)

3Dfx Voodoo

Karty s čipy 3Dfx způsobily revoluci v 3D grafice; využily pádu v ceně paměti a díky poměrně nízké ceně si je mohl dovolit každý. Díky geniální marketingové strategii firmy 3Dfx se 90% her dělalo pro Voodoo - zástupci 3Dfx osobně seznámili vývojáře her s jejich čipem i rozhraním Glide. Karty byly čistě 3D, a zapojovaly se mezi 2D kartu a monitor. 3Dfx zavedlo vlastní API Glide, a díky jeho nižší složitosti se většina nových her dělala pro 3Dfx. Problémem bylo, že renderování fungovalo jen ve fullscreen režimu, v okně ne. Karty byly osazovány 4 MB RAM. Nejznámější kartou byl Diamond Monster 3D, nejstarší byl Orchid Righteous 3D. 3Dfx Voodoo se stalo standartem mezi 3D kartami a odstartovalo hon na "3Dfx killera", každý chtěl Voodoo překonat. Ve své době bylo Voodoo úžasné pokrokové a radikální řešení.

Nevýhodou byla poměrně výrazná degradace kvality obrazu (zvláště ve 2D), a kvůli konceptu 3D přídavné karty bylo nemožné tento nedostatek eliminovat. Nezáleželo, jaká 2D karta byla použita, protože Voodoo si obraz přebralo podle svého, a bylo proto profesionály zavržováno. Rozlišení renderovaného obrazu bylo omezeno na 800x600 bez z-bufferu a 640x480 s z-bufferem, navíc podporovalo pouze 16-bitové renderování. Kromě toho Voodoo nadlouho zabrzdilo vývoj v oblasti 3D akcelerace, jednak proto, že dlouho nebylo potřeba koncept radikálně inovovat, jednak proto, že v něm bylo několik omezení, které prozatím nevadily, ale z hlediska budoucnosti byly slepou uličkou.

Technické parametry: PCI, až 16 MB EDO DRAM (v praxi se však používalo pouze 4 MB), 0.35 mikronová výrobní technologie

Podpora API: Glide, Direct3D, OpenGL (miniport), RenderWare, SurRender, Brender

3Dfx Voodoo Rush

Rush bylo vlastně Voodoo spojené s 2D kartou. Bylo však drahé, nekompatibilní a problémové, proto se neprosadilo. Výkon ve 3D byl slabší než u Voodoo, a 2D bylo spíše průměrné.

Technické specifikace: PCI sběrnice, 6 MB EDO DRAM (4 MB texturové a 2 MB obrazové), 0.35 mikronová výrobní technologie

Podpora API: Glide, Direct3D, OpenGL (miniport), RenderWare, SurRender, Brender

nVidia Riva 128

Plnohodnotné 2D/3D řešení, měl vlastní FPU, podporoval TV-out, byl výkonnější než Voodoo. Kvalita obrazu byla o něco lepší než u 3Dfx Voodoo, i když na profesionální řešení (Matrox) neměla. Nejznámější kartou bylo STB Velocity 128.

Technické specifikace: 128-bit sběrnice, PCI nebo AGP 1X, 4 MB 100 MHz SDRAM nebo SGRAM, 230 MHz RAMDAC, 0.35 mikronová technologie

Podpora API: Direct3D, OpenGL

nVidia Riva 128 ZX

Omezením Rivy 128 bylo pouze 4 MB podporované paměti, ZX verze podporovala 16 MB, a díky tomu se prosadila u těch, kteří Voodoo nebo Rivu 128 odmítali kvůli nízkému použitelnému rozlišení.

Technické informace: 128-bit sběrnice, PCI nebo AGP 2X, 8 nebo 16 MB SDRAM nebo SGRAM, 0.35 mikronová technologie, 250 MHz RAMDAC

Podpora API: Direct3D, OpenGL

ATi Rage Pro

Ve své době velmi rychlé řešení, výkonem někdy výrazně překonávající 3Dfx Voodoo, v jiných hrách naopak. 2D obraz byl na velmi vysoké úrovni, kvalita 3D obrazu však byla horší než na 3Dfx Voodoo. Podporován byl TV-out, hardwarové dekódování MPEG a hardwarová podpora přehrávání MPEG2, ve spojení s TV tunerem od ATi se stalo základem řady ATi All-In-Wonder.

Technické parametry: 4 nebo 8 MB RAM, podpora AGP 2X.

Number Nine Ticket To Ride

Karta byla určena pro profesionální aplikace, ne jako konkurence herních chipsetů, ovšem pro hraní byl použitelný. Výkon v DOSu byl slabý, ve Windows95 a NT výborný.

Technické parametry: 4 nebo 8 MB WRAM, SGRAM nebo EDO-DRAM, 0.35 mikronová technologie, AGP 1X nebo PCI, 128-bit sběrnice, 220 MHz RAMDAC

Podpora API: OpenGL, Direct3D, Heidi

3DLabs Permedia 2

Karta byla navržena hlavně pro profesionální aplikace, cílem bylo produkovat levnou kartu s podporou OpenGL a jiných profesionálních API v programech jako 3DStudio nebo CAD. Díky podpoře všech důležitých rozhraní bylo možno kartu úspěšně proužít i ve hrách. Předností karty byl výborný 2D obraz a vysoké obnovovací frekvence při vysokých rozlišeních

Technické informace: AGP 1X nebo 2X nebo PCI, 64-bit sběrnice, 4 nebo 8 MB SGRAM nebo SDRAM, 230 MHz RAMDAC, 0.35 mikronová technologie

Podpora API: OpenGL, Direct3D, Heidi

Rendition Verité 2100

Levnější verze 2200.

Technické parametry: 4 MB SDRAM nebo SGRAM, PCI, 170 MHz RAMDAC, 50 MHz jádro, 0.35 mikronová technologie

Rendition Verité 2200

R2200 se vyznačovala výbornou kvalitou obrazu a příznivou cenou. Výhodou byly také nízké nároky na hardware, pracovaly dobře i s procesory Cyrix. Nevýhodou však byl slabý 3D výkon, zvláště v DOSu.

Technické parametry: 16 MB SDRAM nebo SGRAM, AGP 1X, TV-out, 230 MHz RAMDAC, 60 MHz jádro, 0.35 mikronová technologie

Podpora API: Direct3D, OpenGL, RRedLine (nativní pro Win95), Speedy3D (nativní pro DOS)

Videologic/NEC PowerVR PCX2

Stejně jako 3Dfx Voodoo šlo o přídavné 3D řešení; narozdíl od Voodoo však spolupráce s 2D kartou neprobíhala přes externí kablík, ale interně přes PCI sběrnici. Čip byl velmi náročný na výkon počítače.

Technické informace: PCI sběrnice, 0.35 mikronová technologie, 4 MB SDRAM

Podpora API: Direct3D, OpenGL, nativní PowerSGL

Matrox Productiva G100

G100 byly určeny k náhradě předchozích karet Mystique II. Zachována byla výborná kvalita 2D obrazu, kvalita 3D obrazu však úplně špičky nedosahovala.

Technické informace: AGP 2X, 64-bit sběrnice, 8 MB SDRAM nebo SGRAM, 0.25 mikronová technologie

Matrox MGA-2164W

Chipset určený pro profesionální aplikace (CAD, DTP). Specialitou byla možnost vložit do systému 4 karty s chipsetem MGA-2164W, pak došlo k rozdělení výstupu na čtyři obrazy a každý obstarávala jedna karta. Tím bylo možno dosáhnout extrémně vysokých rozlišení (3800x2400 pixelů) i obnovovacích frekvencí. Pro hráče tato karta příliš zajímavá nebyla - byla poměrně drahá a výkon ve hrách nijak excelentní nebyl. Karty byly prodávány pod názvem Millenium II. Bylo možno je doplnit přídatným 3D modulem m3D od Matroxu, který byl osazen čipem PowerVR PCX2 a podstatně tak urychlit 3D rendering a přidat nové funkce.

Technické informace: AGP 1X, 64-bit sběrnice, 0.35 mikronový výrobní proces, 16 MB WRAM

Podpora API: Direct3D, OpenGL

7.3 Druhá generace (1998)

3Dfx Voodoo2

Druhá generace Voodoo, stejně jako první Voodoo byly pouze přídavné 3D karty. Karty měly 2 procesory TEXELfx pro nanášení textur a jeden procesor řídicí. Osazovány byly 8 MB nebo 12 MB paměti, z toho 4 MB pro framebuffer a z-buffer a 2 nebo 4 MB pro každý texturový procesor. Použitelné rozlišení se oproti Voodoo zvedlo na 800x600 se z-bufferem. Výkon stoupl na dvoj- až trojnásobek Voodoo.

Legendární a dodnes unikátní byla možnost propojit dvě karty Voodoo2, které pak spolupracovaly při vykreslování obrazu a efektivně se tím zdvojnásobil výkon. Toto se nazývalo SLI (Scan Line Interleave) - jedna karta kreslila sudé řádky a druhá liché. Na dlouhou dobu to bylo dostatečně rychlé řešení pro všechny hry, a v některých počítačích bylo toto občas problémové a přehřívající se monstrum ještě docela nedávno. Spíše mezi technické kuriozity by se dal počítat Obsidian X24, který integroval dvě Voodoo2 a 2D čip na jedné kartě, ovšem vzhledem k tomu, že byl ještě problémovější a dražší než SLI, byl odsouzen k zapomenutí.

Výkon karet v této době závisel čistě na kmitočtu procesoru a kmitočtu jádra karty - víceméně neexistovaly pokročilé optimalizační technologie, které by ulehčily čipu práci při výpočtech, a tak jedinou možností zvýšení výkonu bylo zvýšení frekvence jádra. V době Voodoo2 došlo k prvnímu masovějšímu využití sběrnice AGP, ta byla ovšem využívána pouze jako jakási "záchranná brzda", pokud karta potřebovala rychle načíst jiné textury, které neměla v lokální paměti; samotné Voodoo2 ovšem AGP přímo nepodporovalo.

Karty oproti Voodoo podporovaly trilineární filtrování, sub-texel opravu, bump-mapping, plně hardwarový triangle-setup.

Technické parametry: 8 nebo 12 MB EDO DRAM, 0.35 mikronová výrobní technologie, PCI sběrnice, pouze 16-bitové renderování

Podpora API: Direct3D, OpenGL (miniport), Glide, RenderWare, SubRender, Brender

Intel i740

První karta která bezzbytku využila předností AGP, navíc při velmi výhodné ceně. Výkonově překonávala Voodoo, ovšem na Voodoo2 nestačila. i740 byl

navržen na intelovské chipsety a procesory Pentium II, na ostatních platformách (zvláště VIA chipsety) byly hlášeny různé problémy.

Technické parametry: AGP 2X, 8 MB SGRAM nebo SDRAM, 0.25 mikronová technologie

Podpora API: Direct3D, OpenGL

nVidia Riva TNT

Čip známý jako "Voodoo 2 killer" - výkonem dokonce překonávající Voodoo2 SLI při nesrovnatelně lepší kvalitě obrazu, ceně a menší problémovosti. TNT byla zkratka z "TwiN Texel" (měla 2 texturovací jednotky). Podporováno bylo 4 až 16 MB paměti, a plné využití možností AGP. Byly také použity nové efekty, jako bump-mapping nebo reflection-mapping. TNT bylo původně plánováno s ohledem na 0.25 mikronový výrobní proces, ale do zahájení výroby se nepodařilo jej plně odladit a tak byly karty TNT vyráběny 0.35 mikronovým výrobním procesem a se sníženou frekvencí jádra i paměti. Původně plánovaná podoba TNT byla dotažena do konce až z nástupkyní TNT2. 2D část čipu byla převzata z Rivy 128, 3D část byla nová.

Technické parametry: 16 MB 125 MHz SGRAM nebo SDRAM, AGP 2X nebo PCI sběrnice, 0.35 mikronový výrobní proces

Podpora API: Direct3D, OpenGL

Matrox G200

Vynikající řešení pro náročné uživatele. G200 se vyznačovalo vynikajícím 2D obrazem a slušným 3D výkonem na úrovni Voodoo2, navíc už podporovalo z-buffer. Zvláštností bylo, že rendering ve 32 bitech byl rychlejší než v 16 bitech, což bylo dovedeno do konce u ATi Rage 128. Osvědčily se zvláště v kombinaci s Voodoo2 nebo SLI, navíc uměly hardwarově komprimovat MPEG.

Technické parametry: 16 MB SGRAM nebo SDRAM, AGP 2X

Podpora API: Direct3D, OpenGL

Videologic PowerVR Second Generation

Výkonově se pohyboval na úrovni Voodoo2, uměl hardwarový bump-mapping, plnou podporu DirectX 6.0, bilineární, trilineární i anizotropní filtrování,

implementována byla vlastní komprese textur (8:1). Výborná byla i práce s videem (hardwarová dekomprese MPEG2).

Technické informace: AGP 2X, 0.25 mikronová výrobní technologie, 32 MB SGRAM nebo SDRAM

S3 Savage 3D

Zvláštností karty byla komprese textur S3TC (6:1), která se v mírně pozměněné podobě používá dodnes v DirectX a podporují ji všechny moderní karty. Navíc byl integrován zvláštní ditherovací algoritmus, který umožňoval zobrazit ve 24 bitech i hry projektované pro 16 bitů. Ve své době to byl ojedinělý pokus zvýšit kvalitu obrazu při srovnatelném výkonu, ovšem mizerné ovladače kartu odsoudily k neúspěchu; když se konečně objevily kvalitní vyladěné ovladače, byla karta již zastaralá.

Technické parametry: 0.25 mikronová výrobní technologie, pouze AGP 2x nebo 4X, 8 MB SGRAM nebo SDRAM

Podpora API: Direct3D, OpenGL, MeTal

3Dfx Voodoo Banshee

Voodoo Banshee obsahovalo nejrychlejší 2D jádro na světě plus 3D renderovací základ z Voodoo2 a bylo přímou reakcí firmy 3Dfx na Rivu TNT. Výhodou byla nižší cena, nevýhodou o něco nižší výkon než u Voodoo2 SLI nebo Rivy TNT. Karty měly navíc problémy s kompatibilitou. Podporována byla jak AGP tak PCI sběrnice, ovšem karta byla šita na míru PCI a AGP verze nejenže nebyla rychlejší, navíc byla problémovější.

Technické parametry: 16 MB SGRAM nebo SDRAM, AGP 1X nebo PCI, 0.35 mikronová výrobní technologie, pouze 16-bitové renderování, 250 MHz RAMDAC

Podpora API: Direct3D, Glide, OpenGL

7.4 Třetí generace (1999)

nVidia Riva TNT2

Velice úspěšná série karet, TNT2 bylo vlastně to, co původně nVidie plánovala s TNT. Postupem času nVidie diferencovala čipy TNT2 na několik řad, odstupňovaných podle ceny a výkonu:

TNT2 - původní verze TNT2, 0.25 výrobní proces, 16 nebo 32 MB RAM, 125 MHz jádro, 150 MHz paměť, AGP 4X

TNT2 Ultra - nejrychlejší a také nejdražší karta s TNT2, 175 MHz jádro, 200 MHz paměť

TNT2 Pro - inovovaná TNT2, vyráběná modernějším výrobním procesem 0.22 mikronů, výkonově mezi TNT2 a TNT2 Ultra

TNT2 M64 - ochuzená verze (64-bitová sběrnice místo 128-bitové u plnohodnotné TNT2)

TNT2 Vanta - nejlevnější a nejpomalejší karty s čipy TNT2, výkonově pomalejší než původní TNT, jde o pomaleji taktovanou TNT2 M64 s 16MB 64-bit paměti.

TNT2 Vanta LT – Vanta osazená 8 MB paměti

Díky skvělému výkonu, kompatibilitě a bezproblémovosti se karty TNT2 masově rozšířily a v mnohých počítačích běží dodneška. nVidie touto řadou karet započala svůj specifický výrobní cyklus - vždy dobu přišla s novou technologií (TNT, GeForce, GeForce 3), a po určité době ji zdokonalila aby se na trhu masově prosadila (TNT2, GeForce2, GeForce4).

ATi Rage 128

Od svých začátků propagovali inženýři z 3Dfx 16-bitové renderování - podle jejich slov je 32-bitů "frame killer", i když všechny ostatní karty od konkurence podporovaly i 32 bitů, ovšem při o něco nižší rychlosti. Toto tvrzení definitivně vyvrátila firma ATi, protože její čip Rage 128 byl ve 32 bitech stejně rychlý jako v 16 bitech. Výkon byl o něco vyšší než u TNT. ATi také poprvé představilo 0.25 mikronovou výrobní technologii (Rage 128 přišel na trh o něco dříve než TNT2), a jako první byla využita DDR RAM. ATi dále představila hardwarové dekódování DVD. Vyskytly se však problémy se špatnou implementací AGP (v některých

testech bylo rychlejší i Voodoo3 PCI), karty vyžadovaly procesor Pentium II, a také kvůli o něco vyšší ceně se karty Rage 128 vůči konkurenční TNT2 neprosadily.

Karty se prodávaly ve dvou verzích - **Rage 128 GL** pro OpenGL pracovní stanice, a **Rage 128 VR** pro mainstream.

Technické parametry: 128-bit sběrnice, 32 MB DDR SGRAM nebo SDRAM, AGP 2X nebo 4X, 250 MHz RAMDAC

Podpora API: Direct3D, OpenGL ICD

3Dfx Voodoo3

Voodoo3 bylo realizací plánu s kódovým jménem Rampage. V3 představovalo lehce vylepšené a zrychlené Voodoo Banshee, v podstatě tedy pořád staré Voodoo, výkonově odpovídalo Voodoo2 SLI. Prodávaly se ve verzi PCI i AGP, a opět plně nevyužívaly AGP ani nepodporovaly 32-bitové renderování. Výhodou byla menší závislost na výkonu CPU, ale kvalita obrazu, i když od dob Voodoo o něco vylepšená, nesnesla srovnání s konkurenční TNT2 nebo ATi Rage 128. Další omezení bylo v maximální použitelné velikosti textur (256x256 pixelů, konkurence podporovala 2048x2048 i více). V3 bylo prodáváno ve čtyřech verzích, přičemž frekvence jádra a paměti byla stejná:

STB Velocity 100 - jde vlastně o Voodoo3 2000, ovšem pouze s 8 MB RAM a s jednou texturovací jednotkou.

Voodoo3 2000 - nejpomalejší verze (143 MHz), podle zkušeností i nejméně problémové a cenově nejvýhodnější, lehce přetaktovatelné na úroveň V3 3000

Voodoo3 3000 - rychleji taktované V3 2000 (166 MHz), často se však přehřívalo a bylo poměrně drahé, navíc mělo S-VHS výstup

Voodoo3 3500 - specialitou nejrychlejší verze (183 MHz) byl TV tuner, Digital VCR a FM tuner přímo od 3Dfx, doplněné MPEG2 kódérem/dekódérem v reálném čase, bylo však drahé, problémové, a nakonec překonané ATi All-in-Wonder.

Technické parametry: 16 MB SDRAM, AGP 2X

S3 Savage 4

První akcelerátor s podporou AGP 4X, S3 vylepšila kompresní algoritmus S3TC, bylo podporováno hardwarové přehrávání DVD. Výkon karty odpovídal přibližně Voodoo3 2000 nebo pomalejším TNT2, navíc poskytovala vysoký výkon v nízkých

rozlišeních a mizerný ve vysokých rozlišeních. Neschopnost udělat bezproblémové ovladače nakonec jinak nadějnou kartu pohřbila.

Technické parametry: 32 MB RAM, AGP 4X

Podpora API: Direct3D, OpenGL, MeTal

3DLabs Permedia 3

3DLabs vylepšili herní stránku čipu Permedia 2. Inovací byla taktéž podpora obrovských textur (8192x8192 pixelů) pro profesionální aplikace. Kvůli nekompatibilitě se však neprosadila.

Technické parametry: AGP 4X, 32 MB RAM

Videologic/NEC PowerVR Neon 250

Karta nabízela nižší výkon než TNT2 nebo Voodoo3, hlavně kvůli 64-bitové paměťové sběrnici. Implementován byl opět Tile-based rendering. Kvůli vyšší ceně a nepříliš přesvědčivému výkonu se na trhu neprosadily.

Number Nine Ticket To Ride IV

Vylepšená verze Ticket To Ride, opět určená pro profesionální práci.

Technické parametry: AGP 2X nebo PCI, 128-bit sběrnice, 0.25 mikronová technologie, 32 MB WRAM nebo EDO DRAM nebo SGRAM, 250 MHz RAMDAC

Matrox G400

Karta byla vylepšením předchozí G200. Na vynikající úrovni byla kvalita obrazu. Matrox přišel s technologií DualHead - na jedné kartě byly dva výstupy na monitor a tím bylo možné bezproblémově provozovat dva monitory (proto Matrox diferencoval řadu G400 na Single Head - pouze s jedním výstupem a Dual Head se dvěma). Další unikátní vlastností byl Enviromental Bump-mapping - na dlouhou dobu byly G400 jedinými kartami s podporou této technologie (s další implementací přišly až čipy Matrox G450 a v mainstreamu nVidia GeForce3). Zpočátku byly malé problémy s ovladači, které Matrox časem vyladil a výkonově se dostal na úroveň TNT2. Matrox prodával dvě série karet s čipy G400 (G400 MAX - 166 MHz jádro, 200 MHz paměti, a G400 - 125 MHz jádro, 166 MHz paměti).

Technické parametry: 32 MB RAM, AGP 4X

Podpora API: Direct3D, OpenGL ICD

BitBoys Oy Glaze 3D

Firma BitBoys Oy je známa svými návrhy na čipy, které nikdy nespátřily světlo světa. V době svého představení byl čip Glaze 3D něčím technologicky převratným, díky unikátní zapouzdřené (embedded) paměti přímo v jádře, což mělo být počátkem nového směru ve vývoji 3D čipů. Mediální bublina splaskla, když se ukázalo, že BitBoys nejsou schopni své návrhy realizovat. Přibližně za rok po blamáži Glaze3D s ním BitBoys přišli znovu, opět technologicky na špičce, a opět vše vyznělo do ztracena.

7.5 Čtvrtá generace (první DirectX 7 / T&L generace) (2000)

nVidia GeForce 256 (NV10)

GeForce se stala první kartou postavené na koncepci GPU (Graphics Processing Unit), postavené za účelem ulehčit CPU a přenést více práce s renderingem na čip karty. Ve své době byla GeForce 256 dosti nepochopena - proto, že klasické TNT2 karty podávaly stejný výkon v tehdejších hrách které ještě nebyly stavěny s ohledem na GPU. Situace se zlepšila až postupem času, a nástupem DDR verze která byla podstatně výkonnější.

Další novinkou byly funkce Hardware Transform & Lighting, které souvisely s koncepcí GPU. V principu šlo o přenesení geometrických a nasvětlovacích postupů z CPU na GPU. Jako první čip také podporoval AGP FastWrites a měl čtyři texturovací jednotky (staré čipy měly dvě). Počínaje GeForce se zastavil růst frekvence jádra - namísto získávání výkonu hrubou silou začaly být prováděny optimalizace. GPU mělo původně úplně odbourat závislost výkonu grafické karty na CPU - praxe je ovšem složitější, rendering neobsahuje jen grafické operace, ale i fyziku apod., kterou musí počítat CPU.

Vizionářská koncepce nVidie se prokázala časem - všechny firmy začaly nejpozději s odstupem jedné generace od GeForce 256 chrlit čipy s podporou T&L. Celá technologie GeForce 256 byla Microsoftem převzata a standartizována v podobě DirectX 7.

Technické parametry: AGP 4X nebo PCI, 256-bit sběrnice, 128 MB SDRAM nebo DDR SDRAM, 350 MHz RAMDAC, 0.25 mikronový výrobní proces

Podpora API: Direct3D, OpenGL

S3 Savage 2000

Další pokus S3 o konkurenceschopnou kartu, který opět skončil fiaskem. Papírově moderní karta s GPU a podporou T&L, ale první série karet a ovladačů byly tak problémové, že už při prodeji měly deaktivovanou T&L jednotku, která byla zprovozněna až po několikanásobné revizi ovladačů. Čipy samotné nebyly nijak zvlášť výkonné, a ani cena nebyla příliš lákavá. Šlo o zatím poslední produkt samostatné firmy S3, po neúspěchu Savage 2000 ji koupil čipový moloch VIA za účelem výroby levných integrovaných grafických čipů do základních desek.

Savage 2000 se prodávalo ve dvou verzích - **Savage 2000** a **Savage 2000+**, které se lišily taktovací frekvencí čipu a paměť.

Technické parametry: AGP 4X, 350 MHz RAMDAC, 64 SDRAM RAM, 128-bit sběrnice

Podpora API: Direct3D, OpenGL, MeTal

ATi Rage 128 Pro

Vylepšený čip Rage 128. Samotný čip je přibližně na úrovni TNT2, ale ATi představila projekt Aurora - dva čipy na jedné kartě spolupracovaly a každý střídavě vykresloval jeden obraz (sudý-lichý) (AFR - alternate frame rendering). Silnou stránkou čipu byla tradičně úplná akcelerace DVD, nevýhodou naopak náročnost na hardware a velká problémovost a chybovost ovladačů.

Technické informace: 2 x 32 MB SDRAM (pro každý čip), 0.25 mikronová technologie, AGP 2X nebo 4X

Podpora API: Direct3D, OpenGL ICD

Matrox G450

Karta je vylepšením G400. Zachována a vylepšena byla kvalita obrazu i technologie DualHead, ovšem 3D výkon byl na stejné úrovni nebo spíše horší než u G400 – tento nedostatek připravil čip o přízeň hráčů a dále byly využity v profesionálním sektoru.

7.6 Pátá generace (druhá DirectX 7 / T&L generace) (2000)

nVidia GeForce 2 (NV15)

Největší inovací oproti GeForce 256 byla 0.18 mikronová technologie, a použití výhradně DDR pamětí. Vylepšené jsou také renderovací pipeline, GF2 dokáže zpracovat oproti GF 256 dvě textury najednou, což znamená teoreticky až dvojnásobný výkon, praxe je pochopitelně jiná. nVidie už tradičně odstupňovala GF2 podle výkonu a ceny na několik řad:

GeForce 2 GTS - původní GeForce2, frekvence jádra 200 MHz

GeForce 2 Pro - rozdíl oproti GTS je v rychlejších pamětech (5ns), výkonově mezi GTS a Ultra

GeForce 2 Ultra - rychleji taktovaná GTS, frekvence jádra 250 MHz, frekvence pamětí 460 MHz

GeForce 2 Titanium - inovovaná verze v důsledku přechodu nVidie na řadu Titanium a marketingové číselné značení

Technické parametry: AGP 4X, 32 nebo 64 MB DDR SDRAM

Podpora API: Direct3D, OpenGL

nVidia GeForce 2 MX (NV11)

Podle původních plánů se mělo jednat o čip pro notebooky, ovšem kvůli vysoké ceně GF2 GTS byla koncepce změněna na levnější GTS, za účelem náhrady přežívajících TNT2. MX verze byla v podstatě GeForce2 GTS, ovšem jen se dvěma renderovacími pipeline (GTS má čtyři), a s nižší frekvencí jádra i pamětí. Naopak oproti GF2 měly výhodu v technologii TwinView, která umožňovala pohodlně připojit dva monitory (ovšem oproti DualHead od Matroxu bylo toto dosti amatérské řešení, které navíc mnoho výrobců ignorovalo). Postupem času se opět objevilo několik řad GF2 MX:

GeForce 2 MX - původní verze GF2 MX, frekvence jádra byla 170 MHz, použitá byla SDR paměť a 128-bit sběrnice. Někteří výrobci využili marketingového triku a začali prodávat DDR verzi karet, ovšem s 64-bit sběrnici, což bylo ve výsledku ještě pomalejší než 128-bit SDR. Aby nVidie jasně odlišila pomalejší a rychlejší verze MX, zavedla číselné značení - 200 a 400

GeForce 2 MX 400 - 128-bit sběrnice, frekvence jádra 200 MHz, paměti 64 MB 166 MHz SDR, prakticky přetaktovaná GeForce 2 MX.

GeForce 2 MX 200 - 64-bit sběrnice, 175 MHz jádro, 32 MB 166 MHz SDR paměti.

GeForce 2 MX 100 - plánovaná nejpomalejší verze, nikdy nezačala být sériově vyráběna

Technické parametry: AGP 4X nebo PCI sběrnice

Podpora API: Direct3D, OpenGL

ATi Radeon - Radeon 7200 (R100)

Radeony byly první skutečně herní karty, které obstály v přímém měření s konkurencí (nVidia GeForce 2 GTS). S Radeonem slavily premiéru některé nové funkce, jako např. Keyframe interpolation nebo TruForm. Radeony podporovaly jako první karty od ATi T&L sadu, pod marketingovým názvem Charisma Engine. Radeon výkonově odpovídal GeForce 2 GTS, ale ATi ještě neměla dostatečně bezchybné ovladače a také nebyla zpočátku schopna uspokojit poptávku po nových kartách, a tak na Radeony bylo pohlíženo s nedůvěrou a byly neprávem opomíjeny. Na vynikající úrovni byla tradičně multimediální stránka čipu - přehrávání DVD, video vstupy a výstupy a speciální All-In-Wonder verze. ATi představila také HydraVision, což je obdoba konkurenčního TwinView nebo DualHead, čili připojení dvou monitorů k jedné kartě. Ovšem narozdíl od konkurence je HydraVision přítomen na všech kartách, nejen na těch nejvýkonnějších.

Radeon byl prodáván ve dvou sériích s SDR nebo DDR pamětí. SDR verze výkonově přibližně odpovídaly GeForce2 MX, DDR verze GeForce 2 GTS. První generace Radeonu se stala základem pro další řady a byla prapočátkem dnešní převahy ATi na trhu grafických akceleračních karet. Po revizi marketingových označení dostal Radeon číslo 7200, a po inovaci výrobní technologie 7500.

Technické parametry: 0.25 mikronová technologie, AGP 4X, 32 až 64 MB SDRAM nebo DDR SDRAM

Podpora API: Direct3D, OpenGL

ATi Radeon VE - Radeon 7000 (RV100)

VE edice Radeonu byla mířena proti GeForce 2 MX. Má pouze jednu renderovací pipeline se třemi texturovacími jednotkami a nepodporuje hardwarové T&L, je tedy

čipem hardwarově kompatibilním pouze s DirectX 7. Jako všechny karty však obsahuje HydraVision. Po revizi marketingového značení dostal jméno Radeon 7000.

Technické parametry: AGP 4X, 32 nebo 64 MB DDR SDRAM, 0.18 mikronový výrobní proces, 166 až 183 MHz jádro i paměť, 300 MHz RAMDAC

3Dfx VSA-100

VSA-100 byly konečným důsledkem radikálních vizí inženýrů z 3Dfx. Čipy byly vyvíjeny pod kódovým názvem Napalm a 3Dfx poprvé upřednostnili optimalizace a efekty před hrubým výkonem, ovšem hardwarové T&L nadále ignorovali. Srdcem čipů byl tzv. T-buffer (marketingová náhrada framebufferu), který byl zodpovědný za speciální efekty jako motion-blur nebo full-scene anti-aliasing. Principiálně šlo o vyrenderování více obrazů a následné spojení do jednoho. Karty s čipy VSA-100 byly posledními produkty firmy 3Dfx, po jejich neúspěchu byla firma 3Dfx nucena vyhlásit bankrot a její technologie byly odkoupeny konkurenční nVidií.

VSA byly montovány na několik sérií karet, odstupňovaných podle počtu čipů, potažmo výkonu a ceny. Oficiálně byly prodávány dvě série V4-4500 a V5-5500, ovšem výrobci užívali své značení a své frekvence karet (příkladem budiž Voodoo 4 4200 nebo 4800)

Voodoo 4 4500 - nejlevnější série, osazené jedním čipem VSA-100, 32 MB SDR RAM, 350 MHz RAMDAC, PCI nebo AGP sběrnice, nepodporuje T-buffer (ten potřebuje minimálně dva VSA-100). Výkonově odpovídaly GeForce 256 SDR, ovšem při vyšší ceně a výrazně horších ovladačích. Tyto karty byly naprostým propadákem.

Voodoo 5 5500 - hlavní tahoun nových karet od 3Dfx. Byly osazeny dvěma čipy VSA-100, oba měly aktivní chlazení, podporovaly T-buffer. Osazeno bylo 64 MB paměti SDRAM (32 MB pro každý VSA-100), podporovala AGP nebo PCI sběrnici. Ve své době se jednalo o jedinou kartu plné délky, do mnoha skříní se prostě nevešla. Karty měly vlastní přídavné napájení (obdobně jako nejvýkonnější karty dneška, ovšem tehdy byla celá technologie ještě nedokonalá a často docházelo k tomu, že karta vinou špatného napájení doslova vybourala celý počítač). Výkonnostně a cenově ovšem nestačily na GeForce2. Několik měsíců po vydání karty byly vypuštěny nové ovladače, které využívaly tzv. occlusion-culling (marketingový

název HSR - Hidden Surface Removal), a tím došlo k výraznému nárůstu výkonu, ovšem Voodoo 5 již tou dobou byly odepsané.

Voodoo 5 6000 - karta měla obsahovat 4 čipy VSA-100 a 128 MB SDRAM, ovšem do výroby už se vzhledem ke krachu 3Dfx nedostala. Koncepte byla prodána firmě Quantum3D, která plánovala monstrem s 8 čipy VSA-100, ovšem do sériové výroby se nikdy nedostalo.

Technické parametry: 0.25 mikronová technologie

Podpora API: Glide, Direct3D, OpenGL

ST Micro / Imagination PowerVR3 - Kyro

ST Microelectronics Kyro se objevilo jako blesk z čistého nebe, když 3Dfx zbankrotovalo, S3 bylo odkoupeno VIA, Matrox odstoupil a dále vyráběl karty jen pro profesionální aplikace, ATi Radeon nebyl k dostání, a zdálo se, že monopolu nVidie nic nestojí v cestě. Kyro je založeno na principu čipu PowerVR, a základem je Tile-based renderování. Čip nemusí mít vysoké frekvence ani převratný výkon, TBR nepotřebuje ani T&L ani DDR paměti. První generace Kyra se příliš neprosadila z důvodů nedostatečně doladěných ovladačů a hlavně nedůvěry výrobců i spotřebitelů.

Technické parametry: AGP 4X, 64 MB 115 MHz 128-bit SDRAM nebo SGRAM, 0.25 mikronová technologie, frekvence jádra 115 MHz

Podpora API: Direct3D, OpenGL

S3 Savage XP

Karty měly být zlepšením Savage 2000 s funkčním T&L, ovšem do sériové výroby se nikdy nedostaly.

Technické parametry: 32 MB RAM

7.7 Šestá generace (první DirectX 8 generace) (2001)

nVidia GeForce 3 (NV20)

GeForce 3 se stala završením vývoje započatého GeForce 256. Čip obsahoval 57 milionů tranzistorů, což byl ještě více než má Pentium 4. Někteří vývojáři si stěžovali, že hardwarové T&L je pevně dané, a proto se dá efektivně využít pouze někdy. GeForce 3 je kartou masivních optimalizací. T&L jednotka je flexibilnější, vývojáři můžou naprogramovat vlastní renderovací skripty pro Vertex Shader až o 128 instrukcích. Čip umí nanést 4 textury na pixel v jednom průchodu. GeForce 3 podporuje Pixel Shadery verze 1.1. Implementována je transparentní bezztrátová komprese z-bufferu. Narozdíl od předchozí architektury 256-bitové sběrnice (128-bit DDR), GeForce 3 přináší crossbar-architekturu, která pracuje se čtyřmi 64-bit bloky dat, a potenciálně tak dochází ke čtyřnásobení výkonu. Poprvé je také implementován rozumně využitelný Full-scene anti-aliasing a je to první karta která implementuje vynález Matroxu Enviromental-head bump-mapping. Vzhledem k paměťové náročnosti výpočtů je GeForce 3 první kartou, plně využívající rychlosti AGP 4X.

Zpočátku nebyly karty GeForce 3 příliš oblíbené, protože ve stávajících aplikacích byly někdy i pomalejší než GeForce 2 (v důsledku pomalejšího taktu jádra), ovšem časem s podporou vývojářů se její využitelnost rapidně zvýšila a dnes je architektura GeForce 3 prakticky minimum, pod které nemá vzhledem k budoucnosti smysl jít. Koncepce byla Microsoftem standartizována pro DirectX 8.0.

Půl roku po zahájení výroby GeForce 3 nVidia inovovala výrobní řadu GeForce 3:

GeForce 3 Ti 200 - podtaktovaná původní GeForce 3 (175 MHz jádro, 400 MHz DDR SDRAM), díky tomu šla lehce přetaktovat, nVidia navíc nabídla výbornou cenu.

GeForce 3 - původní verze GeForce 3 (200 MHz jádro, 64 MB 460 MHz DDR SDRAM)

GeForce 3 Ti 500 - prakticky přetaktovaná GF3 (240 MHz jádro, 500 MHz DDR SDRAM)

Technické parametry: AGP 4X, 256-bit sběrnice, 0.15 mikronová technologie

Podpora API: Direct3D (DX 8.0), OpenGL

ATi Radeon 8500, Radeon 9100 (R200)

R200 je druhou generací Radeonu. Konstrukčně i výkonově je prakticky shodné s GeForce 3. Zpočátku kvůli špatným ovladačům trochu zapadl, ale postupem času plně dotáhl GeForce 3. Poprvé se u ATi představil 0.15 mikronový výrobní proces, na kterém staví dodneška. ATi překonala nVidii i v počtu tranzistorů (60 milionů). R200 má čtyři vylepšené renderovací pipeline, které zvládnou aplikovat buď 4 textury při každém taktu, nebo šest na jeden průchod, technologicky je tedy R200 na vyšší úrovni než GeForce 3. Vzhledem k pozdějšímu datu vypuštění R200 podporuje Pixel Shadery verze 1.4, čili plně vyhovuje specifikacím DirectX 8.1.

Po vzoru nVidie ATi časem zavedla výkonově a cenově odstupňované řady:

Radeon 8500 LE - v podstatě pouze pomaleji taktovaný R8500, funkčně naprosto shodný. 250 MHz GPU, 64 nebo 128 MB DDR SDRAM o různé frekvenci

Radeon 8500 - 275 MHz GPU, 64 MB 550 MHz DDR SDRAM

Radeon 8500 XT - 300 MHz GPU, 64 nebo 128 MB 600 MHz DDR SDRAM

Radeon 9100 - jde pouze o přejmenovaný R8500 LE, z důvodů zdůraznění výkonu oproti pozdějšímu Radeonu 9000. Může být osazen 64 nebo 128 MB paměti.

Technické parametry: AGP 4X, 256-bit sběrnice, 0.15 mikronová technologie, 400 MHz RAMDAC, 2 vertex shadery, 4 pipeline po 2 TMU, 6 textur na TMU

Podpora API: Direct3D (DX 8.1), OpenGL

ATi Radeon 7500 (RV200)

RV200 byl určen pro levnější segment trhu, měl pouze dvě pipeline, výkonově lehce převyšoval původní Radeon (7200), technicky mu vlastně odpovídala až na pokročilejší 0.15 mikronovou výrobní technologii. Nepodporuje Pixel Shader, tudíž není hardwarově kompatibilní s DirectX 8.

Technické parametry: AGP 4X, 256-bit sběrnice, 290 MHz jádro, 64 MB 460 MHz DDR SDRAM, 0.15 mikronová technologie, 350 MHz RAMDAC

ST Microelectronics Kyro II

Čip zůstal nadále postaven na technologii PowerVR 3, zlepšena byla výrobní technologie. Hlavní výhodou zůstalo Tile-based renderování, čip tudíž mohl zůstat poměrně nevykonný. Díky obchodnímu partnerství ST Micro s Herculesem mohly být karty Kyro II rozumně distribuovány a také ovladače byly na solidní úrovni.

Výkonově Kyro II odpovídalo GeForce 2 Pro. Nepodporovalo ovšem nové funkce DirectX 8, jako hardwarové T&L, a proto ve hrách dělaných na DX8 ztrácí dech.

Technické parametry: AGP 4X, 0.18 mikronová technologie, jádro 175 MHz,

ST Microelectronics Kyro II SE

Kvůli chybnému hardwarovému T&L u původních čipů Kyro a Kyro II byl plánován Kyro II SE alespoň se softwarovým T&L (tzv. EnTnL), ovšem zřejmě kvůli problémům se ani nedostaly do prodeje.

ST Microelectronics / PowerVR Series 4

Čip měl být nástupcem PowerVR 3, ale kvůli oprodeji ST Microelectronics byl zrušen.

PowerVR 5 (Kyro III)

Čipy měly podporovat DirectX 9 (možná jen částečně), a skloubením Tile-based renderingu a pixel shaderu docílit zajímavého výkonu, navíc měly podporovat pokročilou techniku Hidden Surface Removal, což mělo snížit nároky na hrubý výkon. Celý koncept se však nedočkal realizace.

Předpokládané technické parametry: 0.13 mikronová technologie, 250-350 MHz GPU, 128 nebo 256-bit DDR RAM, 8 pipeline

Matrox G550

Další vylepšení zastaralého G400/450. Čip je vyráběn 0.18 mikronovou technologií, je pinově kompatibilní s G450. Oproti G450 má dvě pipeline a 64-bit DDR paměťové rozhraní. Přítomna je jakási neúplná verze T&L, chybí Pixel Shader. Výkonově odpovídá přibližně GeForce 2 MX 200. Karta byla určena pro profesionální aplikace, podporovala DualHead

7.8 Sedmá generace (druhá DirectX 8 generace) (2001-2002)

nVidia GeForce 4 Titanium (NV25 a NV28)

Sudé číslo série znamená u nVidie tradičně vylepšení předchozí převratné koncepce. Architektura byla podobná GeForce 3, čtyři renderovací pipeline umí až dvě textury najednou. GeForce 4 obsahuje i několik nových funkcí, jako nView (náhrada nepovedeného TwinView). Geometrický výkon je oproti GF3 teoreticky zhruba trojnásobný. Karty s GeForce 4 byly určeny k masovému prosazení na trhu, což se nVidii víceméně podařilo, protože konkurence od ATi nebyla schopna dostatečně rychle nabídnout nástupce Radeonu 8500. Po nějaké době od zahájení prodeje nVidie inovovala řadu GF4 přidáním podpory AGP 8X, ovšem tento tak byl spíše efektní než efektivní, současné karty zatím nejsou schopny potenciál AGP 8X využít a tak je výhoda spíše papírová. Inovované čipy dostaly jméno NV28

Karty byly prodávány v několika verzích:

GeForce 4 Ti 4200 - 250 MHz GPU, 444 MHz paměť (NV25)

GeForce 4 Ti 4200 8X - stejný model s podporou AGP 8X (NV28), 500 MHz paměť

GeForce 4 Ti 4400 - 275 MHz GPU, 550 MHz paměť (NV25)

GeForce 4 Ti 4600 - 300 MHz GPU, 650 MHz paměť (NV25)

GeForce 4 Ti 4800 SE - stejný model jako 4400 Ti s podporou AGP 8X (NV28)

GeForce 4 Ti 4800 8X - stejný model jako 4600 Ti s podporou AGP 8X (NV28)

Technické parametry: 128-bit DDR RAM

Podpora API: Direct3D (DX 8.1), OpenGL

nVidia GeForce 4 MX, MX 4000 (NV17 a NV18)

GF4 MX je hlavně marketingový tah nVidie spíše než odlehčená verze GF4 Ti. Má výrazně zjednodušenou architekturu, která je na úrovni GeForce 2 obohacené o některé prvky architektury GF3 a GF4. Nepodporuje pixel shader a funkce vertex shaderu je omezená, tudíž hardwarově nevyhovuje specifikacím DirectX 8. Výkonově odpovídá přibližně GeForce 2 až GeForce 3, v DX8 hrách ovšem ztrácí. Architektura je postavena na dvou pipeline, každá zpracovává až dvě textury. Z modernějších prvků nechybí nView.

Karty GeForce 4 MX byly již při svém uvedení zastaralé, ovšem nVidie při své tradiční inovaci přidala podporu AGP 8X, která nemá u takto slabě výkonných čipů žádnou váhu; série dostala název NV18. Po uvedení GeForce FX nVidie ještě jednou inovovala už tak zastaralou GF4MX pod názvem GF MX 4000.

Karty byly prodávány v několika verzích:

GeForce 4 MX 420 - 250 MHz GPU, 64 MB 166 MHz SDRAM (NV17), GPU na stejné úrovni je integrován v čipsetu nForce 2.

GeForce 4 MX 440 - 270 MHz GPU, 64 MB 400 MHz 128-bit DDR SDRAM (NV17)

GeForce 4 MX 460 - 300 MHz GPU, 64 MB 550 MHz DDR SDRAM (NV17)

GeForce 4 MX 440 SE - 250 MHz GPU, 64 MB 166 MHz SDRAM (NV18)

GeForce 4 MX 440 8X - 250-275 MHz GPU, 64 MB 333 MHz DDR SDRAM (NV18)

GeForce 4 MX 4000 - 270 MHz GPU, 64 MB 400 MHz 64- nebo 128-bit DDR SDRAM (NV18)

Technické parametry: 0.15 mikronový výrobní proces

Podpora API: Direct3D (DX 7.1a), OpenGL

Matrox Parhelia 512, Parhelia-LX

Matrox vydal nový čip Parhelia, když už byl na herní scéně prakticky odepsaný. Nový čip se výkonově vyrovnal pomalejším modelům GeForce 4 Ti, navíc přišel s novou technologií TripleHead, která umožňovala připojit tři monitory zároveň (Parhelia je název pro vzácný jev, kdy lom světla způsobuje zdání tří sluncí vedle sebe). Novou je také funkce hardware displacement mapping pro modelování plastického povrchu. Architektura je postavená na 256-bitovém paměťovém řadiči a dokáže aplikovat čtyři textury v jednom průchodu. Tyto karty byly od počátku navrhovány pro multi-monitorové systémy, proto Matrox souběžně vyrábí i PCI verze karet za účelem spolupráce dvou karet v jednom počítači. Parhelie byly po dlouhé době nové karty od Matroxu, ovšem kvůli vyšší ceně se prosadily spíše do profesionálního sektoru. Kvůli tomuto faktu Matrox produkuje i levnější série Parhelie-LX. V současné době přichází na trh inovovaná série Parhelia 2.

Matrox P650 (Parhelia-LX) - 64 MB 128-bit DDR SDRAM, DualHead, 2 pipelines

Matrox P750 (Parhelia-LX) - 64 MB 128-bit DDR SDRAM, TripleHead, 2 pipelines

Parhelia 512 - původní plná verze Parhelie, 128 nebo 256 MB 256-bitové paměti, 4 pipelines

Technické parametry: 256-bit sběrnice, AGP 4X nebo PCI, 0.15 mikronový výrobní proces, 220 MHz jádro, 550 MHz DDR RAM

Podpora API: Direct3D (DX8.1), OpenGL 1.3

ATi Radeon 9000 (Pro) (RV250)

Odpověď ATi na GeForce 4 MX. RV250 je zjednodušené jádro R200, Radeony 9000 jsou tedy odvozeninami Radeonů 8500. Rozdíl je v jedné chybějící texturové jednotce v každé pipeline. V praxi to znamená, že výkon při single-texturingu je podobný Radeonu 8500, při multitexturingu horší. Když ATi vydala Radeony 9000, suverénně předčila konkurenční GeForce 4 MX, a to ve všech směrech. Potvrdil se tím smutný fakt, že GeForce 4 MX byly zastaralé již při svém uvedení na trh a byly pouze marketingovým trikem na zákazníky. Série 9000 byla velice úspěšná a dočkala se inovace v podobě série 9200, a s menšími změnami se s nimi počítá i do budoucnosti jako s low-endovými kartami.

V prodeji byly dvě verze, lišící se pouze taktovací frekvencí:

Radeon 9000 - 250 MHz GPU, 64 MB 400 MHz DDR RAM

Radeon 9000 Pro - 270 MHz GPU, 64 nebo 128 MB 550 MHz DDR SDRAM

Technické parametry: 0.15 mikronový výrobní proces, AGP 4X, 2 vertex shadery, 4 pipeline po 1 TMU, 6 textur na TMU

Podpora API: Direct3D (DX 8.1), OpenGL

SiS 315

První pokus SiS soupeřit s MX řadou od nVidie, šlo dokonce o nejlevnější GPU vůbec; výkonově ani kvalitou ovladačů se však neprosadil.

Technické informace: AGP 4X, 143 MHz GPU, 286 MHz DDR SDRAM

Podpora API: Direct3D (DX 7), OpenGL

SiS Xabre

Vážná konkurence pro GeForce 4 MX a Radeon 9000, i přes nedoladěné ovladače jde o cenově zajímavé karty. Jako vůbec první SiS využila 0.13 mikronový výrobní

proces (u Xabre 600), jako první uvedla AGP 8X, a karty podporují Pixel Shadery verze 1.3 (DX 8.1 kompatibilní). Naopak vertex shader je pouze softwarový a ani T&L jednotka není plně hardwarová. Hrubým výkonem se karty dokonce přibližují GeForce 4 Ti 4200, ale shadery zaostávají a jsou spíše na úrovni GeForce 4 MX, kterou však na druhou stranu přebíjí nižší cenou. Celkově se Habre příliš neprosadily kvůli špatným ovladačům a tvrdé konkurenci ze strany ATi i nVidie.

Série Xabre:

Xabre 80 - čip SiS 328 - 200 MHz GPU, 64 MB 128-bit 166 MHz SDRAM nebo 64-bit 333 MHz DDR SDRAM, 0.15 mikronový proces, AGP 4X

Xabre 200 - čip SiS 332 - 200 MHz GPU, 64 MB 128-bit 400 MHz DDR SDRAM, 0.15 mikronový proces, AGP 8X

Xabre 400 - čip SiS 334 - 250 MHz GPU, 64 MB 128-bit 500 MHz DDR SDRAM, 0.15 mikronový proces, AGP 8X

Xabre 600 - čip SiS 338 - 300 MHz GPU, 64 MB 128-bit 600 MHz DDR SDRAM, 0.13 mikronový proces, AGP 8X

Technické informace: 4 pipeline po 1 TMU, 8 textur na TMU

Podpora API: Direct3D (DX 8.1), OpenGL

7.9 Dnešní akcelerátory (první DirectX 9 generace) (2002-2003)

ATi Radeon 9700 (Pro) (R300)

Jádro R300 se stalo základem první generace karet, které výkonnostně i cenově jasně převálcovaly konkurenční nVidii a hlavním dílem se zasadily na dnešní převaze ATi na trhu s grafickými kartami. ATi vydala tyto nové karty v době kdy nVidie ještě neměla protizbraň (NV30) a GeForce 4 Ti byly jasně překonány. Narozdíl od nVidie ATi zůstala u osvědčené 0.15 mikronové technologie, a to se ukázalo společně se zajímavou cenovou politikou hlavním důvodem úspěchu. Poté, co nVidie vydala NV30, jen ztěžji se jí podařilo dotáhnout na R300, a ATi lehkými optimalizacemi a zrychlením opět odskočila ve formě R350. R300 byly první čipy podporující DirectX 9.0, dokonce se objevily na trhu ještě před vydáním samotných DX9. ATi také poprvé v praxi předvedlo jistý vliv AGP 8X na výkon. Všechny high-end karty od ATi i nVidie mají přídatné napájení, jako kdysi Voodoo 5; doporučen je 300W zdroj. Už při testech v novém 3DMarku 2003 se ukázalo, že oba hlavní rivalové nVidia a ATi používají trochu jiné cesty jak dosáhnout výkonu. ATi využívá spíše tradiční 24-bit float interní přesnosti; nVidia vsadila na inovaci v podobě 32-bit nebo 16-bit float, a 12-bit integer; bude ještě zajímavé, jak se celá situace vyvine. Jasně je, že vývojáři si nemohou dovolit ignorovat žádnou koncepci.

Karty s GPU R300 jsou na trhu ve dvou provedeních, lišících se frekvencí GPU a pamětí:

Radeon 9700 - 275 MHz GPU, 540 MHz 256-bit DDR SDRAM, 8 pixelových pipeline

Radeon 9700 Pro - 325 MHz GPU, 620 MHz 256-bit DDR SDRAM, 8 pixelových pipeline

Technické specifikace: 0.15 mikronový výrobní proces, AGP 8X, 400 MHz RAMDAC, 4 vertexové pipeline

Podpora API: Direct3D (DX 9.0), OpenGL

ATi Radeon 9500 (Pro) (RV300)

RV300 se stalo základem prvních karet využívajících DirectX 9 za rozumnou cenu. Výkonově se vyrovnají GeForce 4 Ti. Navíc existuje možnost softwarově

odemknout neaktivní prvky čipu a fakticky z RV300 udělat R300, ovšem s postupem času existuje čím dál menší pravděpodobnost že budou odemčené prvky bezchybné. Když firmy začínaly dodávat Radeony 9500, zřejmě pouze deaktivovaly některé prvky, a prostou aktivací bylo možno rekonstruovat plnohodnotný Radeon 9700, ovšem postupem času logicky začínaly být využívány zmetkové součástky, které však pro výrobu takto odlehčených karet vyhovovaly. Pokud na těchto čipech provedeme softwarové odemknutí, většinou dochází k degradaci obrazu. Obzvláště populární byly Radeony 9500 od Sapphire, které měly nadstandartní 256-bitový paměťový řadič. Karty s 9500 nabízely výborný poměr cena-výkon, ale poměrně záhy po svém uvedení ATi vydala řadu 9600 a starší 9500 marketingově „zazdila“.

Radeon 9500 - 275 MHz GPU, 64 nebo 128 MB 275 MHz 128 nebo 256-bit DDR SDRAM, 4 pixelové pipeline

Radeon 9500 Pro - 275 MHz GPU, 128 MB 275 MHz 128-bit DDR SDRAM, 8 pixelových pipeline

Technické specifikace: 0.15 mikronový výrobní proces, AGP 8X, 400 MHz RAMDAC, 4 vertexové pipeline

Podpora API: Direct3D (DX 9.0), OpenGL

ATi Radeon 9200 (Pro) (RV280)

Jde o GPU RV250 (Radeon 9000), doplněný o podporu AGP 8X, stejně jako u GeForce 4 MX jde spíše o obchodní trik. GPU řady RV280 nepodporují DirectX 9 (narozdíl od hlavní konkurence GeForce FX 5200), ovšem u takto relativně málo výkonných čipů stejně podpora DX9 nedává přílišný smysl, protože aplikace stavěné na DX9 zřejmě budou mít vyšší hardwarové požadavky než by byly tyto čipy schopny nabídnout, a proto ignorace podpory DX9 je od ATi spíše rozumným řešením než nedostatkem. Výkon karet s R9200 odpovídá přibližně původnímu R9000, je tedy srovnatelný s konkurenční GeForce FX 5200. V současnosti je řada 9200 zřejmě nejoblíbenějším low-endovým čipem vůbec.

Nabízen je ve dvou verzích lišících se taktovací frekvencí:

Radeon 9200 SE - 200 MHz GPU, 333 MHz 64-bit DDR RAM

Radeon 9200 - 250 MHz GPU, 400 MHz 128-bit DDR RAM

Radeon 9200 Pro - 300 MHz GPU, 600 MHz 128-bit DDR RAM

Technické parametry: 0.15 mikronový výrobní proces, AGP 8X, 2 vertex shadery, 4 pipeline po 1 TMU, 6 textur na TMU

Podpora API: Direct3D (DX 8.1), OpenGL

nVidia GeForce FX 5800 (Ultra) (NV30)

První GPU GeForce FX se měl stát demonstrací technologické převahy nVidie a zadupat opovážlivou konkurenci ATi a její Radeon 9700 do země. Karta byla doslova prošpikována nejnovějšími technologiemi, v čele s nejmodernější 0.13 mikronovou výrobní technologií a GDDR-II paměťmi s vysokým kmitočtem a jako prvním u nVidie podporou DirectX 9. Pro efekty byl použit engine s marketingovým názvem CineFX, kompresní technologie a optimalizace obrazu zastupuje termín IntelliSample. Ale všechno bylo špatně. GPU výkonem sotva dotáhlo Radeon 9700, čip se přílišně hřál, a tak nVidie navrhla systém FXFlow, který spočíval v masivním chladiči přesahujícím až do další pozice pod AGP kartou. Šlo o velice monstrózní, těžké, a hlavně hlučné řešení (mimořádně víceméně okopírované od chladiče OTES, který používá na svých kartách Abit, ten je ovšem o dost tišší). Při práci ve 2D běžel větrák na standardní frekvenci, srovnatelné s ostatními chladiči. Ve 3D se však rozeběhl do vysokých otáček a ne náhodou se GeForce FX 5800 začalo říkat vysavač. Paměti GDDR-II jsou sice velice rychlé, ale také zatím poměrně drahé. nVidie využila pouze 128-bitové paměťové rozhraní, které ATi přebila levnějším a stejně efektivním 256-bitovým rozhraním a DDR SDRAM. FX 5800 byl naprostým debaklem, a mnozí prorokovali nVidii konec a přirovnávali FX 5800 k 3Dfx Voodoo 5. nVidie vyrobila jen několik set tisíc čipů, pak výrobu přeorientovala na modernější NV35, a dnes dokonce na FX 5800 nenajdeme ani odkaz na firemních stránkách nVidie. Dokonce i generální ředitel nVidie uznal chybu a na tiskové konferenci si činovníci nVidie tropili z FX 5800 žerty. FX 5800 přišel na trh zřejmě příliš brzy, neodladěná výrobní technologie a nezvládnuté chlazení společně s neefektivním využitím paměti odsoudily FX 5800 ke krachu.

Karty se dodávaly ve dvou provedeních, lišících se frekvencí GPU a paměť:

GeForce FX 5800 - 400 MHz GPU, 128 MB 800 MHz 128-bit GDDR-II

GeForce FX 5800 Ultra - 500 MHz GPU, 128 MB 1000 MHz 128-bit GDDR-II

Technické parametry: 0.13 mikronový výrobní proces, AGP 8X, 8 renderovacích pipeline po 1 TMU, 16 textur na TMU

Podpora API: Direct3D (DX 9.0, vertex a pixel shadery v2.0+), OpenGL 2.0

nVidia GeForce FX 5600 (Ultra, XT) (NV31)

Jedná se o jádro NV30, ochuzené o polovinu renderovacích pipeline. Při výrobě jsou použity standartní DDR SDRAM, a GPU pracuje na nižších frekvencích, takže není potřeba FXFlow, což zřejmě zachránilo FX 5600 od osudu FX 5800. Ovšem v porovnání s konkurenční ATi Radeon 9600 je nVidie opět o trochu pomalejší a dražší. FX 5600 však byly primárně určeny také k nahrazení přežívajících GeForce 4 Ti. Výkonem je překonávají při použití optimalizací a hlavně pod Direct3D, naopak GeForce 4 Ti jsou díky vyššímu hrubému výkonu rychlejší pod OpenGL a při vypnutých nebo snížených optimalizacích.

Řada GeForce FX 5600 obsahuje čtyři modely lišící se frekvencí (model XT byl reakcí na podobné značení ATi, ovšem zde znamená naopak nižší výkon).

GeForce FX 5600 XT - 235 MHz GPU, 128 MB 64-bit (výjimečně 128-bit) DDR SDRAM

GeForce FX 5600 - 325 MHz GPU, 550 MHz 128-bit DDR SDRAM

GeForce FX 5600 Ultra - 350 MHz GPU, 700 MHz 128-bit DDR SDRAM

GeForce FX 5600 Ultra rev.2 - 400 MHz GPU, 800 MHz 128-bit DDR SDRAM

Technické informace: 0.13 mikronový výrobní proces, AGP 8X, 4 renderovací pipeline po 1 TMU

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

nVidia GeForce FX 5200 (Ultra), 5500 (NV34)

Tyto čipy byly na architektuře NV30 založeny pouze částečně, liší se starším 0.15 mikronovým výrobním procesem. Jsou určeny k náhradě zastaralých GeForce 4 MX a pomalejších GeForce 3 a 4 Ti, ovšem podobně jako u FX 5600 na ně nestačí hrubým výkonem. Zajímavé je, že díky svým kmitočtům jsou v jistých podmínkách rychlejší karty GF FX 5200 Ultra než FX 5600. Diskutabilní výhodou oproti těmto starším kartám a konkurenčnímu Radeonu 9200 je plná podpora DirectX 9.0, otázkou totiž je, zda bude na DX9 aplikace stačit hrubý výkon těchto karet, teoreticky jsou však ideálním prostředkem pro prosazení DX9 do všech počítačů, a pro připravení půdy pro budoucí 3D akcelerované GUI Windows Longhorn, které poběží na DX9. Model 5500 je pouze inovovanou verzí původní FX 5200.

Nabízeny jsou v těchto verzích:

GeForce FX 5200 - často s pasivním chlazením, 250 MHz GPU, 400 MHz 128-bit DDR SDRAM

GeForce FX 5500 - 270 MHz GPU, 400 MHz 128-bit DDR SDRAM

GeForce FX 5200 Ultra - 325 MHz GPU, 650 MHz 128-bit DDR SDRAM

Technické informace: 0.15 mikronový výrobní proces, AGP 8X, 4 renderovací pipeline po 1 TMU

Podpora API: Direct3D (DX 9.0), OpenGL

SiS Xabre II (projekt)

Projekt SiS, který se měl vyrovnat konkurenčním high-end čipům s podporou DirectX 9. Vývoj byl však nakonec zrušen a veškeré kapacity obráceny na vývoj a výrobu XGI Volari. Xabre II měl být nabízen ve dvou variantách lišících se výkonem:

SiS 340 - 8 pipeline, 375 MHz GPU, 1000 MHz GDDR-II

SiS 341 - 4 pipeline, 375 MHz GPU, 1000 MHz GDDR-II

Technické informace: 0.13 mikronový výrobní proces, AGP 8X

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

Trident XP4 (projekt)

Trident měl tímto čipem slavit návrat mezi grafickou elitu. Výkonově měl soupeřit s levnějšími verzemi Radeonů (9500), nebo GeForce 4 Ti za velmi nízké ceny. Čipy měly svůj výkon stavět na pokročilých optimalizacích obrazu, utajovaném multiplex-renderingu a 0.13 mikronové technologii, podporován měl být DirectX 8.1, a série odstupňována podle výkonu na čtyři řady (T1, T2, T3, T4). Nakonec se vše vyvinulo jinak a Trident se stal součástí XGI a podílí se na vývoji čipů Volari.

Trident XP 8 (projekt)

Mělo se jednat o rychlejší alternativu XP4, vyráběnou možná už 0.10 mikronovou technologií. V souvislosti s přesunem kapacit k XGI byl nakonec nerealizován.

7.10 Dnešní akcelerátory (druhá DirectX 9 generace) (2003-2004)

ATi Radeon 9800 (Pro) (R350)

Jde o přetaktovaný a vylepšený R300, rychlostně vyrovnaný s GeForce FX 5900 (Ultra). Technologie výroby zůstala na 0.15 mikronech, podporována je buď DDR SDRAM, nebo i nový standard GDDR-II, ovšem na dnešních kartách jsou rozdíly od konkurenčních GeForce FX 5800 použity klasické DDR SDRAM, protože GDDR-II je zatím příliš drahá. Dá se však očekávat, že příští generace karet od ATi (R400) bude postavena na stejném jádru a GDDR-II. U karet R9800 SE lze opět s jistotou šancí na úspěch aplikovat softwarové odemknutí pipeline, a dostat tak plnohodnotný R9800, výhodný je zejména Hercules který vyrábí SE verzi s 256-bitovým řadičem paměti. Poslední modely Radonů 9800 se již vyrábí s novějším jádrem R360. V současné době je řada 9800 symbolem dominance firmy ATi na trhu s 3D akcelerátory. Tradiční rozdělení karet od ATi na modely o různých frekvencích vypadá následovně:

Radeon 9800 SE - 380 MHz GPU, 680 MHz 128 nebo 256-bit DDR SDRAM, 4 renderovací pipeline po 1 TMU

Radeon 9800 - 380 MHz GPU, 680 MHz 256-bit DDR SDRAM, 8 renderovacích pipeline po 1 TMU

Radeon 9800 Pro - rychlejší a mírně optimalizovaná verze R9800,

Technické specifikace: 0.15 mikronový výrobní proces, AGP 8X, 400 MHz RAMDAC, 4 vertexové pipeline

Podpora API: Direct3D (DX 9.0, vertexové a pixel shadery v2.0+), OpenGL 2.0

ATi Radeon 9800 XT (R360)

R360 je pouze lehce vylepšené a přetaktované R350. Další inovací je tepelná dioda, která umožňuje automatickou regulaci frekvence GPU. Jedná se o poslední high-end GPU od ATi vyrobené 0.15 mikronovou technologií.

Technické parametry: 0.15 mikronový výrobní proces, AGP 8X, 400 MHz RAMDAC, 4 vertexové pipeline, 8 renderovacích pipeline po 1 TMU, 412 MHz GPU, 256 MB 730 MHz 256-bit DDR SDRAM

Podpora API: Direct3D (DX 9.0, vertexové a pixel shadery v2.0+), OpenGL 2.0

ATi Radeon 9550, 9600 (Pro) (RV350)

GPU RV350 je prvním od ATi vyráběným 0.13 mikronovou technologií. Jde vlastně o karty konstrukčně i výkonově dost podobné Radeonu 9500, ovšem na rozdíl od R9500 se nejedná o zmetky které neprošly při výrobě výkonnější série, nýbrž o zcela novou řadu. ATi se podařilo opět o kousek předstihnout konkurenci, protože Radeony 9600 jsou výkonnější než GeForce FX 5600. Řada 9550 byla původně určena pouze pro východní trhy, ale zdá se že se bude prodávat po celém světě. Řada 9600/9550 se vyrábí v několika verzích lišících se pouze taktovacími frekvencemi a šířkou paměťové sběrnice:

Radeon 9550 SE - 250 MHz GPU, 400 MHz 64-bit DDR SDRAM

Radeon 9600 SE - 325 MHz GPU, 400 MHz 64-bit DDR SDRAM

Radeon 9550 - 250 MHz GPU, 400 MHz 128-bit DDR SDRAM

Radeon 9600 - 325 MHz GPU, 400 MHz 128-bit DDR SDRAM

Radeon 9600 Pro - 400 MHz GPU, 600 MHz 128-bit DDR SDRAM

Technické parametry: 0.13 mikronový výrobní proces, AGP 8X, 2 vertex shadery, 4 pipeline po 1 TMU, 400 MHz RAMDAC

Podpora API: Direct3D (DX 9.0), OpenGL

ATi Radeon 9600 XT (RV360)

Radeon 9600 XT je výsledkem optimalizací jádra RV350. Je vyráběn novější technologií low-dielectric, která umožnila vyšší taktovací frekvence a tím i vyšší výkon. V budoucnosti se počítá s tím, že díky optimalizacím shaderů tato architektura bude schopna rozbíhat hry dokonce rychleji než starší jádro R300 (Radeon 9700). Radeon 9600 XT byl od začátku mířen jako protiváha GeForce FX 5700/Ultra, a výkonově jsou obě architektury víceméně vyrovnané.

Technické parametry: 0.13 mikronový výrobní proces, AGP 8X, 500 MHz GPU, 600 MHz 128-bit DDR SDRAM, 4 renderovací pipeline po 1 TMU, 400 MHz RAMDAC

Podpora API: Direct3D (DX 9.0), OpenGL

nVidia GeForce FX 5900 (Ultra, XT) (NV35)

NV35 byl čipem, na kterém závisela záchrana pověsti nVidie po debaklu FX 5800, a záchrana byla úspěšná. NV35 byl vlastně optimalizovaný čip NV30, čili původní FX 5800, ovšem využit byl 256-bitový paměťový řadič s klasickými DDR

SDRAM. Kmitočet GPU i paměti je snížen, ovšem výkonem překonává FX 5800 i Radeon 9700 a vyrovná se Radeonu 9800. Další optimalizace se týkají kvality obrazu a paměťové propustnosti, CineFX je ve vylepšené verzi 2.0.

NV35 jsou k dostání ve čtyřech verzích (model XT byl reakcí na podobné značení ATi, ovšem zde znamená naopak nižší výkon. Verze XT se stala poměrně populární díky výbornému poměru cena/výkon, ovšem s uvedením FX5700 Ultra byly tyto marketingově "zazděny" a postupně začaly být stahovány z trhu. SP verzi vyráběla pouze firma MSI a okamžitě po jejím uvedení způsobila nákupní horečku, ovšem pak začala být také nahrazována verzí XT na levnějším PCB):

GeForce FX 5900 XT - 300 MHz GPU, 128 MB 700 MHz 256-bit DDR SDRAM, 4 nebo 8 pipelines

GeForce FX 5900 SE - 400 MHz GPU, 128 MB 700 MHz 256-bit DDR SDRAM, 8 pipelines

GeForce FX 5900 SP - 400 MHz GPU, 128 MB 850 MHz 256-bit DDR SDRAM, 8 pipelines

GeForce FX 5900 - 400 MHz GPU, 128 MB 850 MHz 256-bit DDR SDRAM, 8 pipelines

GeForce FX 5900 Ultra - 450 MHz GPU, 128 nebo 256 MB 850 MHz 256-bit DDR SDRAM, 8 pipelines

Technické parametry: 0.13 mikronový výrobní proces, AGP 8X, 16 textur na TMU

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

nVidia GeForce FX 5950 Ultra (NV38)

NV38 je v podstatě pouze rychleji taktovaný NV35, za účelem překonat konkurenční Radeon 9800 XT. Skutečnost je taková, že v aplikacích vyžadujících hrubý výkon mají mírně navrch čipy nVidie, a v aplikacích spíše stavěných na optimalizacích si o něco lépe vedou čipy ATi. Obsahuje 135 milionu tranzistorů (pro porovnání - procesor Athlon XP má 54 milionů tranzistorů).

Technické parametry: 0.13 mikronový výrobní proces, AGP 8X, 8 renderovacích pipeline po 1 TMU, 16 textur na TMU, 475 MHz GPU, 256 MB 950 MHz 256-bit DDR SDRAM

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

nVidia GeForce FX 5700 (Ultra) (NV36)

NV36 je určen jako rychlejší alternativa k GeForce FX 5600, s tím že časem ji zřejmě nahradí úplně. Díky tomu, že NV36 je nikoliv odvozenina NV30 jako FX 5600, ale úplně novým čipem odvozeným z NV35, obsahuje mnoho konstrukčních prvků z high-end GeForce FX 5900 a 5950. Mechanismus ochuzení je stejný jako u FX 5600, GPU má pouze polovinu renderovacích pipeline (i fyzicky - nejde o odpadové čipy ale plně 4-pipelines architekturu) a poloviční šířku paměťové sběrnice. Nejvýznamnějším rozdílem oproti FX 5600 je však engine CineFX 2.0, který umožňuje výrazně rychlejší vykonávání pixel shader kódu, a dále rychlejší antialiasing. Čip NV36 je prvním produktem spolupráce nVidie s IBM (CPU ostatních modelů jsou vyráběny u TSMC), a proto se GPU vzhledově dosti liší od všech ostatních GeForce FX. Výkon FX 5700 přibližně odpovídá Radeonu 9600, rychlejší model 5700 Ultra vybavený GDDR-II pamětmi se vyrovná Radeonu 9600 XT.

Po čase od uvedení NV36 nVidia opět vydala další dvě karty LE a VE, jedná se o odpadové čipy které nebyly schopny fungovat na standardních frekvencích. Zároveň byly uvedeny první karty FX 5700 Ultra vybavené pamětmi GDDR-III.

Specifikace řady NV36:

GeForce FX 5700 VE - 235 MHz GPU, 128MB 128-bit 400 MHz DDR SDRAM

GeForce FX 5700 LE - 250 MHz GPU, 128MB 128-bit 400 MHz DDR SDRAM

GeForce FX 5700 - 450 MHz GPU, 550 MHz 128-bit DDR SDRAM

GeForce FX 5700 Ultra - 475 MHz GPU, 900 MHz 128-bit DDR SDRAM nebo GDDR-II nebo GDDR-III

Technické informace: 0.13 mikronový výrobní proces, AGP 8X, 4 renderovací pipeline po 1 TMU, 16 textur na TMU

Podpora API: Direct3D (DX 9), OpenGL

XGI Volari

Čipy Volari jsou výsledkem spojení grafické divize SiS a Trident. XGI znamenalo Xabre Graphics Incorporated, později eXtreme Graphics Innovation. Čipy Volari se dají považovat za jakési skutečné Xabre II. Plánem je vyrovnat se současné špičce nVidia a ATi. Prozatím jde o novou a nezavedenou značku, výkon nejvýkonnějších Duo V8 Ultra se z důvodu nedoladěných ovladačů pohybuje maximálně na úrovni

Radeonu 9700, možná bude časem ještě lepší. Zajímavostí Duo verzí je nutnost napájet je hned dvěma napájecími konektory, což byla dosud nevídaná věc.

Karty s čipy Volari budou na trhu v několika provedeních:

Volari Duo V8 Ultra - 2 procesory V8, 2x8 pipeline, 350 MHz GPU, 750 MHz DDR RAM nebo 1000 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari Duo V5 Ultra - 2 procesory V5, 2x4 pipeline, 350 MHz GPU, 750 MHz DDR RAM nebo 1000 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari V8 Ultra - 1 procesor V8, 8 pipeline, 350 MHz GPU, 750 MHz DDR RAM nebo 1000 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari V8 - 1 procesor V5, 8 pipeline, 300 MHz GPU, 650 MHz DDR RAM nebo 900 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari V5 Ultra - 1 procesor V8, 4 pipeline, 350 MHz GPU, 750 MHz DDR RAM nebo 1000 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari V5 - 1 procesor V5, 4 pipeline, 300 MHz GPU, 650 MHz DDR RAM nebo 900 MHz GDDR-II, DirectX 9.0, OpenGL 2.0

Volari V3 - 1 procesor V3, 2 pipeline, 300 MHz GPU, 600 MHz DDR RAM, DirectX 8.1, OpenGL

Technické parametry: AGP 8X, 0.13 mikronová technologie

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

S3 Deltachrome

Deltachrome je po dlouhé době nový čip S3, mířící mezi mainstreamové karty (S3 je fakticky vlastněno firmou VIA a mezi Savage a Deltachrome byly kapacity S3 využity k výrobě integrovaných grafických čipů jako např. Unichrome). Výkon tradičně srážejí špatné ovladače, výkonnostně je nejvýkonnější Deltachrome někde výkonnější, někde méně výkonná než XGI Volari. Pro S3 ale hovoří výrazně lepší cena.

Karty Deltachrome jsou vyráběny v těchto provedeních:

Deltachrome S4 - 128 MB 64 nebo 128-bit DDR paměti, 4 pipelines

Deltachrome S4 Nitro - 128 MB 64 nebo 128-bit DDR paměti, 4 pipelines

Deltachrome S8 - 300 MHz jádro, 128 MB 300 MHz paměti, 8 pipelines

Deltachrome S8 Nitro - 315 MHz jádro, 128 MB 315 MHz paměti, 8 pipelines

Deltachrome F1 - 256 MB DDR RAM, 8 pipelines

Deltachrome F1 Pole - 256 MB DDR RAM, 8 pipelines

Technické parametry: AGP 8x, 0.13 mikronová technologie, 128-bit paměťové rozhraní, DDR

Podpora API: Direct3D (DX 9.0), OpenGL 2.0

7.11 Nejmodernější a budoucí grafické čipy (2004)

nVidia GeForce PCX

Jedná se vlastně o zavedené karty GeForce FX a GeForce 4 MX, které narozdíl od svých předchůdců využívají sběrnice nové generace - PCI-Express. Tato podpora narozdíl od konkurenční ATi není nativní, ale stará se o ni speciální bridge HSI (high-speed-interconnect). Vliv HSI na výkon karet není zatím úplně jasný, protože tyto karty se zatím neprodávají – důvodem je minimální rozšíření desek podporujících PCI-Express.

nVidia GeForce PCX 5950 - Jde o FX 5950, ovšem na sběrnici PCI-Express. Má stejné technické parametry.

nVidia GeForce PCX 5300 - Jde o FX 5200, ovšem na sběrnici PCI-Express. Má stejné technické parametry.

nVidia GeForce PCX 5750 - Jde o FX 5700, ovšem na sběrnici PCI-Express. Má stejné technické parametry.

nVidia GeForce PCX 4300 - Jedná se o starou GeForce4 MX440 / MX440-8XSE / MX4000, ovšem na sběrnici PCI-Express. Má stejné technické parametry.

nVidia GeForce 6800 (Ultra) (NV40)

Nová generace GPU od nVidie. GPU obsahuje rekordních 222 milionů tranzistorů a podporuje nejmodernější GDDR-III paměti. Nová revize CineFX architektury pracuje až s 16 virtuálními pipelines. nVidia z názvu nové karty dokonce vynechala „FX“, toto je očividně zapříčiněno nepříliš přesvědčivými prodejními výsledky řady FX. V současné době se karty s těmito čipy již začínají prodávat ve světě.

Verze 6800 Ultra Extreme je reakcí na nejrychlejší modely X800 XT od konkurenční Ati. Jedná se o přetaktovanou 6800 Ultra, která se zřejmě nebude vyrábět ve velkém počtu a její význam je spíše marketingový. Tyto čipy jsou dnes nejvýkonnějším řešením od nVidie, a v některých podmínkách (rendering OpenGL) nejrychlejším řešením vůbec.

6800 - 325 MHz GPU, 128 MB 256-bit 700 MHz GDDR-III RAM, 12 pipelines

6800 GT - 350 MHz GPU, 128 MB 256-bit 1000 MHz GDDR-III RAM, 16 pipelines

6800 Ultra - 400 MHz GPU, 256 MB 256-bit 1100 MHz GDDR-III RAM, 16 pipelines

6800 Ultra Extreme - 450 MHz GPU, 256 MB 256-bit 1100 MHz GDDR-III RAM, 16 pipelines

Technické parametry: 0.13 mikronový výrobní proces, vertex a pixel shader 3.0

nVidia NV45

Jedná se o PCI-Express verzi NV40, zřejmě dojde ke zvýšení frekvencí jádra i paměti díky zefektivnění výrobního procesu.

ATi Radeon X800 (R423,R420)

Nová generace GPU od ATi. Počet tranzistorů se vyšplhal na 160 milionů. Vzhledem k tomu, že předchozí generace Radeonů měla označení 9xxx, nová generace je označena Xxxx, velké "X" symbolizuje desítku. Na rozdíl od konkurenční GeForce 6800 se nejedná o žádnou převratnou technologii, ale o další vylepšení předchozích Radeonů. Karty s čipy řady X800 se v současné době začínají prodávat i u nás. Karty s čipem X800 XT jsou v současné době vůbec nejrychlejším řešením pro počítačové hry – zvláště při použití optimalizací obrazu. Karty s čipem X880 XT se prozatím nezačaly vyrábět z důvodů skoro nulového rozšíření desek podporujících PCI-Express.

X800 SE - jádro R420, AGP 8x, 450 MHz GPU, 128 MB 128-bit 800 MHz DDR SDRAM, 8 pipelines

X800 Pro - jádro R420, AGP 8x, 475 MHz GPU, 256 nebo 512 MB 256-bit 900 MHz GDDR-III RAM, 12 pipelines

X800 XT - jádro R420, AGP 8x, 520 MHz GPU, 256 nebo 512 MB 256-bit 1120 MHz GDDR-III RAM, 16 pipelines

X880 XT - jádro R423, PCI-Express x16, 600 MHz GPU, 256 nebo 512 MB 256-bit 1120 MHz GDDR-III RAM, 16 pipelines

Technické parametry: 0.13 mikronový low-k výrobní proces, vertex a pixel shader 2.0

ATi RV380

Následník čipu RV360, určený pro rozhraní PCI-Express. Karty budou zřejmě pojmenovány R X600 a výkonově budou následovat současnou sérii R9600.

ATi RV370

Low-end řada chipů, odvozená od RV280, určená pro rozhraní PCI-Express. Karty zřejmě ponesou název R X200 a výkonem se budou podobat současným R9200.

S3 Gammachrome

PCI-Express verze Deltachrome.

Závěr

Závěrem bych rád upozornil na druhou část absolventské práce, přiloženou na CD. Shromáždil jsem ilustrační obrázky grafických karet od historie až po současnost, dále ukázky z moderních akcelerovaných her jako demonstraci toho, o čem vlastně v této práci píšu, a nejrozšířenější benchmarkovací programy. Místo vymezené na CD je bohužel příliš malé na to, aby se na něj vešlo vše popsané v této práci – ideálním řešením by byly demoverze nejnovějších her současnosti. Druhým zdrojem informací je proto internet – doporučuji návštěvu serverů uvedených v seznamu zdrojů; zde se novinky z oblasti herní grafiky objevují každý den.

Cílem této práce nebyly složité matematické postupy a rozbor algoritmus používané při renderingu. Účelem bylo poskytnout čtenáři přehledný popis technologií a samotných grafických karet dneška, což se myslím podařilo. Plánuji tuto práci volně zveřejnit na internetu za účelem doplňování dalších informací.

Seznam použité literatury

Časopisy

GameStar (ISSN 1212-3331), ročník 1999 až 2004

LEVEL (ISSN 1211-6777), ročník 1997 až 2004

SCORE (ISSN 1210-7522), ročník 1994 až 2004

Zdroje všeobecných informací na internetu

www.pctuning.cz

www.pctuning.cz/forum2

www.pretaktovani.cz

www.svethardware.cz

www.technet.cz

www.tomshardware.com

www.anandtech.com

Zdroje technických informací na internetu

<http://msdn.microsoft.com/directx>

<http://opengl.org>

www.chipanalyst.com/3D

www.nvidia.com/Developer

www.ati.com/developer

internetové stránky jednotlivých výrobců hardware

Zdroje obrázků

www.pctuning.cz

www.svethardware.cz

www.bonusweb.cz

internetové stránky jednotlivých výrobců hardware

Přílohy

K této absolventské práci přikládám CD, obsahující:

- fotografie grafických karet
- diagramy znázorňující postupy popsané v této práci
- benchmarkovací programy
- demonstrační screenshoty z moderních her
- srovnávací testy výkonu grafických karet