**NVIDIA GeForce**

　　Geforce（中文一般称为精视™）是一个英文产品的商标。 (Geometry-Force=Geforce)几何很强

　　GeForce是NVIDIA公司出品的显示芯片的一种系列

    1999年8月，NVIDIA发布了图形芯片领域开天辟地的产品GeForce 256，也由此提出了全新的GPU概念。此后经过将近十二年的发展，跨越五个DirectX时代(DX7-11)，GeForce终于迎来了这个里程碑的历史性时刻，**平均每天出货24多万颗，每小时就超过1万颗。**



　　目前已有GeForce256、GeForce2、GeForce3、GeForce4、GeForce FX、GeForce 6系列，GeForce 7系列，GeForce 8系列，GeForce 9系列和第十代产品“GeForce+定位+型号”系列。

目录

• [产品简介](http://www.techcn.com.cn/index.php?doc-view-153889.html#1)

• [累计出货10亿颗！NVIDIA GeForce创造历史](http://www.techcn.com.cn/index.php?doc-view-153889.html#3)

• [参考文献](http://www.techcn.com.cn/index.php?doc-view-153889.html#5)

产品简介[回目录](http://www.techcn.com.cn/index.php?doc-view-153889.html" \l "section)

　　本文所涉及到的规格与数据，均采自Nvidia公版显卡。

第一代产品：GeForce 256

　　DX7.0 OpenGL PCI/AGP4X（DDR版本只有AGP接口）

　　第一款256bit显示核心的可编程GPU,支持T&L立方环境映射、Dot3凹凸映射和HDTV动态补偿和硬件alpha混合。在显示接口方面，GeForce 256率先开始支持AGP4×快写模式。

**第二代产品：GeForce 2**

　　DX7.0 OpenGL AGP4X

　　GeFoece 2有明显高低端之分，MX系列为低端，其余系列为中高端。

　　加入数字振动控制(DVC)，TwinView(双头显示)功能。在当时，FSAA技术已经开始兴起，FSAA可以大量的改善画质，去除难看的锯齿，甚至是旧游戏也可以让它变得更光滑好看，此前只有3dfx的双VSA-100芯片Voodoo5支持。GeForce2 GTS也可以用硬件FSAA，它的方式是用SuperSampling，但是SuperSampling FSAA只能使用在低分辨率，很显然的，nVIDIA的SuperSampling方式比3dfx的MultiSampling的效率还要差，另外nVIDIA驱动程序没有DirectX的FSAA功能，这使得拥有FSAA技术的GeForce2 GTS并没有得到很好的改善画质的效果。

　　支持S3TC、FSAA、Pixels Shaders和硬件动态MPEG-2动态补偿功能。

**第三代产品：GeForce 3**

　　DX8.0 OpenGL AGP4X

　　光速存储结构（Lightspeed Memory Architecture），在高分辨率下，即使当时最快的DRAM模组也无法满足新一代的图形芯片这个事实已经不再是什么秘密。就因为这个原因ATI开发了HyperZ技术，它可以解决许多在存取数据和消除Z缓冲中浪费掉的带宽。它属于一种无损的压缩算法（即减小了数据量，但不降低画面精度）。nVIDIA的LMA结构的出发点和ATI十分类似，他们希望通过优化显存带宽的方法来达到最大限度利用板载230MHz DDR显存的目的。其中首要的创新是多路交叉显存控制器（Crossbar Memory Controller）。我们知道128位的DDR显存在一个时钟周期内可以同时传递2次数据，换句话来讲就是一次可以传递256位的数据包。但问题就出在这里：如果一个数据包的容量只有64位，也会占用一个数据传送周期，也就是在硬件端发送数据时仍然被当做256位来看待的，所以这里带宽的利用率实际上就只有25%非常低下。现在GeForce3改变了这个情况。它提供了并行工作的4个显存控制器，这些装置能够在与GPU通讯的同时互相联系，具体一点就是4个控制器分别掌管64位数据带宽，在遇到大数据包时（＞64位）可以整合在一起工作，遇到小的数据包时各自为战。这样的处理方式极大地提高了显存带宽的利用率。

　　在LMA架构中的第二项技术是无损的Z轴数据压缩算法，这个改进思想来源于ATI RADEON。由于芯片处理每个像素的时候要考虑到它们在三维场景中的深度坐标，所以Z缓冲是显存和芯片之间数据传递的关键部分，带宽占用的分量最重。nVIDIA已经开发了一个可以将Z轴数据压缩四倍的算法，在不会带来任何精度丧失的同时，也节省了许多不必要的带宽浪费。 nVIDIA称之为光速显存架构亦可以称为VS（Visibility Subsystem，可见子系统）。它和隐藏表面去除（Hidden Surface Removal）有着密切的关系，这是一种在PowerVR和Mosaic芯片设计中得到大量论述的关键技术。如果没有这种技术，一块图形芯片必须处理CPU传送来的每个3D对象，甚至那些在最终图像中被临时遮蔽的对象。采用了绘制隐藏对象或表面的过程称之为超量绘制（OverDraw），它意味着图形芯片在一个典型3D游戏中要渲染2到4倍于所需的像素。这次nVIDIA已经加入了一个被称之为Z轴吸收选择（Z-Occlusion Culling）的技术来达到和隐藏面去除技术类似的效果。GeForce3通过在绘制一个帧的早期检查深度值来取消隐藏的像素，也就是在应用转换和光源效果之后。nVIDIA声称它预防超量绘制的方法在实际处理过程中可以获得50%-100%的性能提升。

　　其次是nVIDIA的nfinite FX引擎，nVIDIA在标准的T&L引擎旁边增加了一个高度可编程的动画和效果引擎，这就是“nfinite FX”引擎。它是GeForce3中最精密、复杂的新单元，与纯粹的T&L引擎相比它的应用更为广泛。虽然可编程性听起来更像是针对开发商而言，玩家或许不会对此关注，不过可以肯定一点的是没有人喜欢一再看到同样的3D特效！根据这样设计,游戏设计师们可以创造出独特和充满变化的效果。

　　nfiniteFX引擎分为2个部分：一个专用于处理像素的，另一个则专门用来处理顶点或几何效果。我们记得在GeForce2 GTS发布的时候，它的像素着色器（Pixel Shaders）的技术已经非常眩目。这次NVIDIA决定让它增值，新的技术叫做顶点着色器（Vertex Shaders）。新的顶点流水线可以优先传送实时可见的视觉效果。GeForce3中的可编程Vertex Shaders技术允许开发者有更多的选择。过去一个顶点数据包含了位置坐标（X，Y，Z）、颜色、光影和纹理结构数据，现在又多了一个Vertex Shaders数据。它可以在不需要CPU经常发送指令的前提下改变3D对象。顶点Shader效果尤其生动，因为它们可以控制一个物体的形状、动画或光线。很方便地操控一个顶点数据浮现在别的表面，表现不同的透明度以及不同的色彩等。当然了并不是每个顶点都需要进入到Vertex Shaders中进行处理，只有当有特殊要求的时候，这才是必要的。

　　就像你从附图中看到的那样，顶点投影和硬件T&L加速单元并行执行。如果顶点阴影器正在运行则T&L单元就空闲下来。这并不是一个问题，因为其实阴影器本身也包含有硬件加速顶点转换的功能，可以说Vertex Shader是性能更强大，功能更多的硬加速转换和光影引擎。之所以让两者同时存在则是为了兼容的原因。Vertex Shader是可编程的，这里有大量的效果可以产生，唯一的限制是开发人员的编程技能。NVIDIA已经提供了一个手册介绍了整个效果库，其中包括了近百种效果。下面我们简单地来看看Vertex Shader包含的效果。

　　还有就是顶点效果，角色动画最早被使用在Vertex Shaders中。如果你记得，ATI已经在Radeon引入了一些硬件编码的特点，例如游戏角色的关节皮肤或关键帧插补技术。GeForce3可以进行更加精密的皮肤处理，它允许利用内部骨架建造角色，而美工（或者游戏的物理引擎）来实现骨架的运动，同时仍然以真实方式保持关节附近的纹理。

**第四代产品：GeForce 4**

　　MX：DX7.0 OpenGL1.3 AGP4X/8X 桥接PCI-E（PCX4300）

　　TI：DX8.0 OpenGL AGP4X/8X

　　GeFoece 4有明显高低端之分，MX系列为低端，TI系列为中高端。

　　GeForce4 MX(研发代号为NV17)。GeForce4 MX并不是源自GeForce4 Ti的核心而源自于GeForce2 MX所采用的NV11图形核心。与GeForce2 MX相比，GeForce4 MX最主要的改良是拥有较高的频率与两段交错式内存控制器的LMA II(GeForce3与GeForce4 Ti有四段)，可以增加可利用的内存带宽。

　　GeForce4 Ti它仍采用4条渲染流水线，每条流水线包含2个TMU(材质贴图单元)，但GeForce4 Ti是依靠其他方面的改进和增强来实现提高性能的目的。在T&L单元方面拥有nfiniteFX II引擎，它是从GeForce3时代开创的nFiniteFX引擎改进而来的。在GeForce3中只有一个顶点着色引擎，而GeForce4 Ti配备了两个。两个平行的顶点着色引擎是可以在同一时间处理更多的顶点。两个单位都是多线程与多线程处理，并在芯片上执行的，因此可以通过应用程序或API来获得性能上的好处。指令的分配是由NV25所处理的，但是那必须去确认每个顶点着色引擎处理的是不同的顶点，这样才会使平行处理变得较实用。顶点着色引擎是从GeForce3内的原始版本所优化而来的，因此降低了指令的延迟。

　　此外，GeForce4 Ti也同时引入了LightSpeed Memory Architecture II(LMA II)光速显存构架II技术，通过优化渲染过程及数据压缩等技术，大大节省了显存带宽，提升了显卡的性能表现。在全屏反锯齿方面，GeForce4 Ti采用了新的Accuview AA技术。

　　nVIDIA使用「新的取样位置」来改善AA的品质，因为可以减少错误堆积，特别是使用Quincunx AA的情况下。新的过滤技术会在每次(多重)取样时被一起执行来产生反锯齿的帧，并且省下了完整的帧写入，这样可以大幅改善AA的性能。基本上，Accuview应该可使AA看起来更好看，并且跑的更顺。

　　GeForce4可以把nVIDIA特有的「Quincunx」-AA执行像一般2x-AA一样快。除此之外，nVIDIA增加了额外的模式，叫做「4xS」。这个模式应该看起来会比4x AA模式好很多，因为增加了50%的子像素覆盖范围(不过这个模式只支持Direct3D游戏，不支持OpenGL)。通过采用新的取样方式并优化执行过程，GeForce4 Ti在这方面的速度是GeForce3的2倍，它保证了高分辨率下的运行速度，取得了性能和画面品质之间的平衡。

**第五代产品：GeForce FX**

　　发布于2002年11月18（核心代号NV3X）

　　新一代的DX9.0技术，可提供比7.0和8.0更逼真的游戏特效，使得游戏更为生动！GeForce5200 作为NVIDIA第一款支持DX9.0的廉价显示卡，表现的效果并不理想，但凭借着DX9.0的名气，在取代GeForce4MX的地位上取得成功。市场购买量非常大，但性能却还不如GeForce4 TI 系列，以后出的中端GeForce5600 5700，也并不成功。就算是GeForce5600U在DX9.0方面，也不尽人意。其高端的GeForce5800，5900，5950，其温度高，功耗大，温度和功耗并对不起它本身的性能。与对手ATI的R9600，9700，9800的性能相差甚远，NVIDIA在GeForce5系列彻底输给对手ATI。GeForce5950U具有噪音王之称！GeForce5200具有GeForceMX440复刻版之称！

　　DX9.0+ OpenGL AGP8X 桥接PCI-E（PCX系列

　　支持HCT、CineFX1.0、CineFX2.0、UltraShadow等技术。

**第六代产品：GeForce 6**

　　DX9.0C OpenGL1.5 AGP8X/PCI-E

　　首款支持DX9C与SM3.0的产品。

　　NV40将NVIDIA特有的CineFX技术提升到了3.0版本，支持Pixel Shader 3.0 Vertex Shader 3.0、实时位移映射和Tone映射，最多支持16纹理/通道，支持16bit和32bit FP格式、sRGB纹理，支持DirectX和S3TC压缩、支持32bpp、64bpp和128bpp渲染以及更多的新视觉效果。

　　GeForce 6加入了HDR?High-Precision Dynamic-Range，高精确度动态范围　OpenEXR技术、支持FP滤波、纹理、混合。改进的Intellisample 3.0，支持16×AA，改善了压缩性能；支持HCT?高分辨率压缩　，它是一种全新的色彩、纹理和Z buffer在所有模式中的无损压缩算法，并具有高分辨率高频率和快速Z buffer清除功能。采用了UltraShadow Ⅱ技术，在大量阴影化应用游戏如Doom Ⅲ中，与上一代GPU相比可提高4倍性能。

　　另外geometry instancing也是很实用的功能之一，它可以让开发者批量处理大量琐碎的运算和小模件，把他们组成一个大的模块并有效的由总线传输到图像里面――这个功能对于即时战略游戏特别有用。通过geometry instancing的功能，程序只需要发送较少的绘图指令就可以批量处理所有的数据，从而大幅度减轻处理器的负担。

　　GeForce 6还增强了温度监控和管理功能，增强了显示和视频输出功能，如NVIDIA Digital Vibrance Control ?DVC　3.0和NVIDIA数字振动控制3.0。在API方面，NV40支持最新的DirectX 9.0C和OpenGL 1.5。

　　另外，GeForce 6还增加了渲染管线和CineFX3.0技术。渲染管线作为引擎最重要的功能之一，当3D模型制作完毕之后，美工会按照不同的面把材质贴图赋予模型，这相当于为骨骼蒙上皮肤，最后再通过渲染引擎把模型、动画、光影、特效等所有效果实时计算出来并展示在屏幕上。渲染管线就相当于处理这一系列工作的工作者。渲染管线的快慢直接影响图形最终生成的快慢，虽然GPU可以在一秒钟内处理相当多的三角形，但需要依赖渲染管线来进行贴图等即时处理。所以它的强大与否直接决定着最终的输出质量。

　　GeForce 6采用了一个完全不同于GeForce FX的着色器，尤其是对Shader Model 3.0（着色器模型3.0）的支持是目前独有的功能。Shader Model 3带来了三方面的改进：大大丰富了的编程模型，这样你可以编写更长的程序，可以进行更佳的程序流控制。因此从开发者的角度来看，他们在Shader Model 3上可以做更多他们感兴趣的事情，可以做过去在Shader Model 2上不能实现的功能，或者在Shader Model 3上更有效的去实现某些功能。例如复杂的人物动画片制作，在Shader Model 3的程序中当你需要略过某个角色或场景时，你可以分支或跳过一些暂时无用的代码，这样可以达到更快及更佳的效果，而在Shader Model 2你就必须去执行这些代码。由此来分析，Shader Model顶点纹理取码（vertex texture fetch）是相当值得关注的一项技术。顶点纹理取码允许应用程序直接从显存中提取纹理信息来作顶点处理，这种技术可以用在包括实时位移贴图（displacement mapping）等方面使用，通过这个功能你可以在顶点着色器3.0中实现各顶点的位移工作。

　　从象素的角度来说，GeForce 6提供了一个更丰富的编程环境，你可以编写非常非常长的程序，它可以提供比Shader Model 2更多的指令数。你将得到一个真实的程序流控制模型，它可以提供循环及分支选项，就等同一个真实的编程环境。Shader Model 3拥有FP32（32位浮点）的精确度，你不可能得到任何因不够精度计算下产生的对象，现在所有低于FP32的运算都只能定为不够精确。随着程序的复杂化，FP24在精度方面的不足会越来越明显。Pixel Shader 3.0的指令槽数目是512，通过循环和分支选项，最多可以执行65000条指令。

　　在处理方面，GeForce 6采用了顶点处理器，和R420不同在NV40的顶点处理器中是一个32位的浮点向量ALU、32位的浮点标量ALU，以及一个顶点材质阴影单元（这个材质单元还有一个专用的cache），相比较X800的128位的向量处理器似乎少了一点，但是在GeForce 6的顶点处理器中最至关重要的就是他的顶点材质处理器，这是ATi所没有的，这个处理单元是GeForce 6所使用的Vertex shader 3.0技术所必需的，这项技术已经给大家展示出了它的魅力。

　　和R420不同，NV40的16条像素管线仍然并行分布，而不是上面我们看到R420那样将16条管线分成了4组，并规定了各自的功能。这样的优点是可以非常的灵活组合，NV40支持8x2以及16x1以及32x0三种管线模式，即一个GPU周期同时处理两块像素，每块材质可以有8条管线；或者同时只处理一块材质但是其中使用了全部16条管线；当处理Z轴数据时它不需要分两次独立处理，也就成为事实上的32管线。NV40的像素处理器的另一个特点就是使用了超标量技术，我们知道这项技术最好的应用典范便是Intel的CPU。这项技术使得NV40在处理像素时每个GPU周期可以处理最多达到8个像素点，效率比不使用这项技术时理论上最多可提高50%的性能。

**第七代产品：GeForce 7**

　　DX9.0C OpenGL2.0，DX9.0C OpenGL1.5（7100GS） AGP8X/PCI-E

　　支持CineFX 4.0Engine高精度渲染引擎、IntlliSample智能反锯齿系统、Pixel Shader渲染指令集，完美支持Microsoft DirectX 9.0 Shader Model 3.0 和OpenGL 2.0 optimizations引擎等技术特性，同时执行效率较GeForce 6系列有大幅提升。

　　G70核心提升至IntelliSample 4.0版本，除了提升至16X各项异性过滤达128 Taps采样外，更加有了两种全新的反锯齿模式，TSAA (透明动态超级采样)及TMAA(透明动态多采样)，旧有的反锯齿技术无法对幼长的物件如栅栏、树叶、植物等物品于Direct3D下产生作用，但新的TSAA及TMAA模式则能解决以上问题，这也是上代GeForce 6所不能比拟。

**第八代产品：GeForce 8**

　　DX10.0 OpenGL 2.0 PCI-E/PCI-E2.0

　　2006年11月8日（北京时间9日），nVIDIA公布了新的显卡型号--Geforce 8系列，揭开了DirectX 10的大幕，GeForce 8800 GTX与GTS天价问世，其后续推出的旗舰级显卡GeForce 8800Ultra成为了史上最强悍的显卡。

　　GeForce 8采用前所未有的设计，统一Shader架构（Unified Shader）带来强劲的性能。G80完全硬件支持DirectX10的各项先进特性，例如Geometry Shaders、stream out、Improved instancing和Shader Modle4.0，支持这些特性使得Geforce8800 GPU拥有极高性能。所有的DirectX9、OpenGL和先前的DirectX程序和游戏在GeForce 8 GPU的Unified Shader设计上都有高性能的演出。

第九代产品 ：GeForce 9

　　DX10.0 OpenGL 2.0 PCI-E2.0

　　GeForce 8大部分产品的工艺与频率改进版

第十代产品：GeForce+定位+型号

　　（比如最新的GeForce GTX580）

　　DX10.0 OpenGL 2.0 PCI-E2.0

　　定位字母解释（从高至低排列）：

　　Ultra：Nvidia的旗舰级产品，为本系列中的最强者。

　　GTX:其性能介于GTS和Ultra之间。一般为Nvidia首发的高端产品。

　　GTS:始终为Nvidia第三强的产品（GeForce 2 GTS与GeForce 8800GTS均可展示出来），性能处于GTX之下，与GT相比有些模糊。

　　GT:频率提升版本"GeForce Technoloty"的缩写,频率和管线都较LE GS SE XT有较大的提升

　　GS:相比GT，渲染管线或者显存位宽的缩减，频率一般在GT之下，并且由于规格限制，性能在GT之下。

　　LE:"Limit Edition"的缩写，表示限制版本，代表某一产系列中的入门级产品，主要是频率和规格均与标准版本相比有一定的下降。

　　SE:在Nvidia卡中不常出现，与LE相似。

　　XT:"Cost Down"表示降频率版本,将标准版的频率降低，部分产品削减了管线。

　　早期命名：

　　MX:GeForce2/4系列，低端产品

　　TI：GeForce2/3/4系列，中高端产品

　　NV曾经还有PRO(GF2) ZT(FX5900) GSO(9600)

累计出货10亿颗！NVIDIA GeForce创造历史[回目录](http://www.techcn.com.cn/index.php?doc-view-153889.html" \l "section)

NVIDIA、台积电今天联合宣布，前者设计、后者制造的GeForce GPU图形芯片历史累计出货量已经突破10亿颗大关。1999年8月，NVIDIA发布了图形芯片领域开天辟地的产品GeForce 256，也由此提出了全新的GPU概念。此后经过将近十二年的发展，跨越五个DirectX时代(DX7-11)，GeForce终于迎来了这个里程碑的历史性时刻，**平均每天出货24多万颗，每小时就超过1万颗。**

NVIDIA老大黄仁勋表示：“自从十多年前发明GPU以来，NVIDIA以技术产业无可匹敌的速度推动着这些处理器的创新。感谢与台积电的合作关系，这些设备的复杂度已经提高了1000多倍，推动了PC、掌上设备、工作站、数据中心各个领域的全面进步。”

台积电CEO兼董事长张忠谋也不吝赞美之词：“NVIDIA的成就有力地证明了代工厂与无工厂企业的紧密合作给消费电子产业、半导体产业以及双方公司股东所带来的好处。NVIDIA堪称创新的典范，我们也高度重视NVIDIA、台积电之间的长期合作关系。”



参考文献[回目录](http://www.techcn.com.cn/index.php?doc-view-153889.html" \l "section)

http://baike.baidu.com/view/403319.htm  
http://www.nvidia.cn/object/geforce\_family\_cn.html