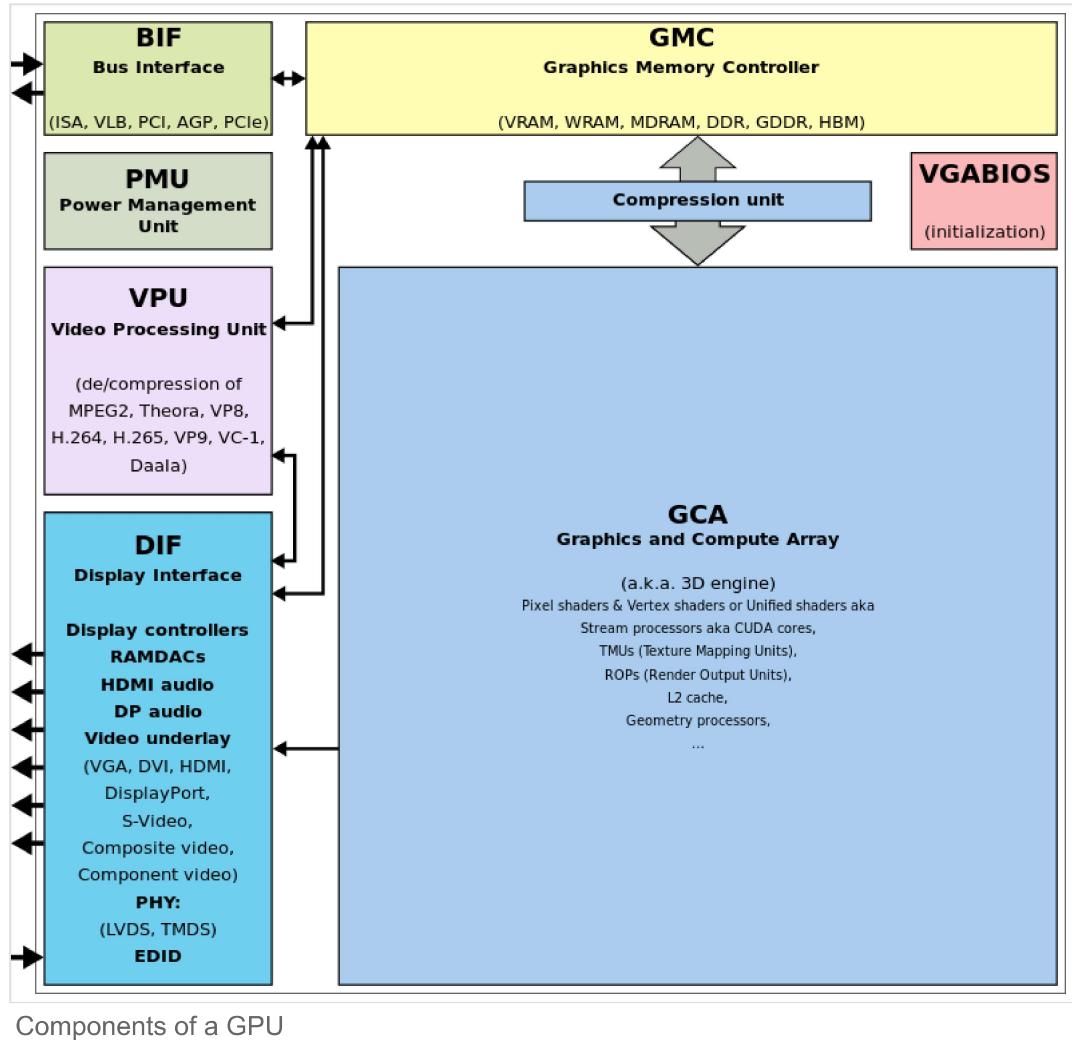


WIKIPEDIA
The Free Encyclopedia

Graphics processing unit



Components of a GPU

A **graphics processing unit (GPU)** is a specialized electronic circuit initially designed to accelerate computer graphics and image processing (either on a video card or embedded on motherboards, mobile phones, personal computers, workstations, and game consoles). After their initial design, GPUs were found to be useful for non-graphic calculations involving embarrassingly parallel problems due to their parallel structure. Other non-graphical uses include the training of neural networks and cryptocurrency mining.

History

1970s

Arcade system boards have used specialized graphics circuits since the 1970s. In early video game hardware, RAM for frame buffers was expensive, so video chips composited data together as the display was being scanned out on the monitor.^[1]

A specialized barrel shifter circuit helped the CPU animate the framebuffer graphics for various 1970s arcade video games from Midway and Taito, such as Gun Fight (1975), Sea Wolf (1976), and Space Invaders (1978).^[2] The Namco Galaxian arcade system in 1979 used specialized graphics

hardware that supported RGB color, multi-colored sprites, and tilemap backgrounds.^[3] The Galaxian hardware was widely used during the golden age of arcade video games, by game companies such as Namco, Centuri, Gremlin, Irem, Konami, Midway, Nichibutsu, Sega, and Taito.^[4]

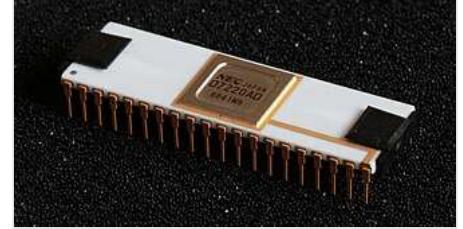
The Atari 2600 in 1977 used a video shifter called the Television Interface Adaptor.^[5] Atari 8-bit computers (1979) had ANTIC, a video processor which interpreted instructions describing a "display list"—the way the scan lines map to specific bitmapped or character modes and where the memory is stored (so there did not need to be a contiguous frame buffer).^[6] 6502 machine code subroutines could be triggered on scan lines by setting a bit on a display list instruction.^[7] ANTIC also supported smooth vertical and horizontal scrolling independent of the CPU.^[8]



Atari ANTIC microprocessor on an Atari 130XE motherboard

1980s

The NEC μPD7220 was the first implementation of a personal computer graphics display processor as a single large-scale integration (LSI) integrated circuit chip. This enabled the design of low-cost, high-performance video graphics cards such as those from Number Nine Visual Technology. It became the best-known GPU until the mid-1980s.^[9] It was the first fully integrated VLSI (very large-scale integration) metal–oxide–semiconductor (NMOS) graphics display processor for PCs, supported up to 1024×1024 resolution, and laid the foundations for the emerging PC graphics market. It was used in a number of graphics cards and was licensed for clones such as the Intel 82720, the first of Intel's graphics processing units.^[10] The Williams Electronics arcade games *Robotron 2084*, *Joust*, *Sinistar*, and *Bubbles*, all released in 1982, contain custom blitter chips for operating on 16-color bitmaps.^{[11][12]}



NEC μPD7220A

In 1984, Hitachi released ARTC HD63484, the first major CMOS graphics processor for personal computers. The ARTC could display up to 4K resolution when in monochrome mode. It was used in a number of graphics cards and terminals during the late 1980s.^[13] In 1985, the Amiga was released with a custom graphics chip including a blitter for bitmap manipulation, line drawing, and area fill. It also included a coprocessor with its own simple instruction set, that was capable of manipulating graphics hardware registers in sync with the video beam (e.g. for per-scanline palette switches, sprite multiplexing, and hardware windowing), or driving the blitter. In 1986, Texas Instruments released the TMS34010, the first fully programmable graphics processor.^[14] It could run general-purpose code, but it had a graphics-oriented instruction set. During 1990–1992, this chip became the basis of the Texas Instruments Graphics Architecture ("TIGA") Windows accelerator cards.

In 1987, the IBM 8514 graphics system was released. It was one of the first video cards for IBM PC compatibles to implement fixed-function 2D primitives in electronic hardware. Sharp's X68000, released in 1987, used a custom graphics chipset^[15] with a 65,536 color palette and hardware support for sprites, scrolling, and multiple playfields.^[16] It served as a development machine for Capcom's CP System arcade board. Fujitsu's FM Towns computer, released in 1989, had support for a 16,777,216 color palette.^[17] In 1988, the first dedicated polygonal 3D graphics boards were introduced in arcades with the Namco System 21^[18] and Taito Air System.^[19]

IBM introduced its proprietary Video Graphics Array (VGA) display standard in 1987, with a maximum resolution of 640×480 pixels. In November 1988, NEC Home Electronics announced its creation of the Video Electronics Standards Association (VESA) to develop and promote a Super VGA (SVGA) computer display standard as a successor to VGA. Super VGA enabled graphics display resolutions up to 800×600 pixels, a 36% increase.^[20]



The IBM 8514 Micro Channel adapter, with memory add-on

1990s

In 1991, S3 Graphics introduced the S3 86C911, which its designers named after the Porsche 911 as an indication of the performance increase it promised.^[21] The 86C911 spawned a host of imitators: by 1995, all major PC graphics chip makers had added 2D acceleration support to their chips.^[22] Fixed-function Windows accelerators surpassed expensive general-purpose graphics coprocessors in Windows performance, and such coprocessors faded from the PC market.

Throughout the 1990s, 2D GUI acceleration evolved. As manufacturing capabilities improved, so did the level of integration of graphics chips. Additional application programming interfaces (APIs) arrived for a variety of tasks, such as Microsoft's WinG graphics library for Windows 3.x, and their later DirectDraw interface for hardware acceleration of 2D games in Windows 95 and later.



VGA section on the motherboard in IBM PS/55



Tseng Labs ET4000/W32p



S3 Graphics ViRGE



Voodoo3 2000 AGP card

The term "GPU" was coined by Sony in reference to the 32-bit Sony GPU (designed by Toshiba) in the PlayStation video game console, released in 1994.^[31]

In the PC world, notable failed attempts for low-cost 3D graphics chips included the S3 ViRGE, ATI Rage, and Matrox Mystique. These chips were essentially previous-generation 2D accelerators with 3D features bolted on. Many were pin-compatible with the earlier-generation chips for ease of implementation and minimal cost. Initially, performance 3D graphics were possible only with discrete boards dedicated to accelerating 3D functions (and lacking 2D graphical user interface (GUI) acceleration entirely) such as the PowerVR and the 3dfx Voodoo. However, as manufacturing technology continued to progress, video, 2D GUI acceleration, and 3D functionality were all integrated into one chip. Rendition's Verite chipsets were among the first to do this well. In 1997, Rendition collaborated with Hercules and Fujitsu on a "Thriller Conspiracy" project which combined a Fujitsu FXG-1 Pinolite geometry processor with a Vérité V2200 core to create a graphics card with a full T&L engine years before Nvidia's GeForce 256; This card, designed to reduce the load placed upon the system's CPU, never made it to market. NVIDIA RIVA 128 was one of the first consumer-facing GPU integrated 3D processing unit and 2D processing unit on a chip.

OpenGL appeared in the early '90s as a professional graphics API, but originally suffered from performance issues which allowed the Glide API to become a dominant force on the PC in the late '90s.^[32] These issues were quickly overcome and the Glide API fell by the wayside. Software implementations of OpenGL were common during this time, although the influence of OpenGL eventually led to widespread hardware support. DirectX became popular among Windows game developers during the late '90s. Unlike OpenGL, Microsoft insisted on providing strict one-to-one support of hardware. The approach made DirectX less popular as a standalone graphics API initially, since many GPUs provided their own specific features, which existing OpenGL applications were already able to benefit from, leaving DirectX often one generation behind. (See: Comparison of OpenGL and Direct3D.)

Microsoft began to work more closely with hardware developers and started to target the releases of DirectX to coincide with those of the supporting graphics hardware. Direct3D 5.0 was the first version of the API to gain widespread adoption in the gaming market, and it competed directly with more-hardware-specific, often proprietary, graphics libraries, while OpenGL maintained a strong following. Direct3D 7.0 introduced support for hardware-accelerated transform and lighting (T&L) for Direct3D; OpenGL had this capability from its inception. 3D accelerator cards moved beyond being simple rasterizers to add another significant hardware stage to the 3D rendering pipeline. The Nvidia GeForce 256 (also known as NV10) was the first consumer-level card with hardware-accelerated T&L; professional 3D cards already had this capability. Hardware transform and lighting—existing features of OpenGL—came to consumer-level hardware in the '90s and set the precedent for later pixel shader and vertex shader units which were far more flexible and programmable.

2000s

Nvidia was first to produce a chip capable of programmable shading: the GeForce 3. Each pixel could now be processed by a short program that could include additional image textures as inputs, and each geometric vertex could likewise be processed by a short program before it was projected onto the screen. Used in the Xbox console, this chip competed with the one in the PlayStation 2, which used a custom vector unit for hardware accelerated vertex processing (commonly referred to as VU0/VU1). The earliest incarnations of shader execution engines used in Xbox were not general purpose and could not execute arbitrary pixel code. Vertices and pixels were processed by different units which had their own resources, with pixel shaders having tighter constraints (because they

execute at higher frequencies than vertices). Pixel shading engines were actually more akin to a highly customizable function block and did not really "run" a program. Many of these disparities between vertex and pixel shading were not addressed until the [Unified Shader Model](#).

In October 2002, with the introduction of the [ATI Radeon 9700](#) (also known as R300), the world's first [Direct3D 9.0](#) accelerator, pixel and vertex shaders could implement [looping](#) and lengthy floating point math, and were quickly becoming as flexible as CPUs, yet orders of magnitude faster for image-array operations. Pixel shading is often used for [bump mapping](#), which adds texture to make an object look shiny, dull, rough, or even round or extruded.^[33]

With the introduction of the [Nvidia GeForce 8 series](#) and new generic stream processing units, GPUs became more generalized computing devices. [Parallel](#) GPUs are making computational inroads against the CPU, and a subfield of research, dubbed GPU computing or [GPGPU](#) for *general purpose computing on GPU*, has found applications in fields as diverse as [machine learning](#),^[34] [oil exploration](#), [scientific image processing](#), [linear algebra](#),^[35] [statistics](#),^[36] [3D reconstruction](#), and [stock options pricing](#). [GPGPU](#) was the precursor to what is now called a compute shader (e.g. CUDA, OpenCL, DirectCompute) and actually abused the hardware to a degree by treating the data passed to algorithms as texture maps and executing algorithms by drawing a triangle or quad with an appropriate pixel shader. This entails some overheads since units like the [scan converter](#) are involved where they are not needed (nor are triangle manipulations even a concern—except to invoke the pixel shader).

Nvidia's [CUDA](#) platform, first introduced in 2007,^[37] was the earliest widely adopted programming model for GPU computing. [OpenCL](#) is an open standard defined by the [Khronos Group](#) that allows for the development of code for both GPUs and CPUs with an emphasis on portability.^[38] OpenCL solutions are supported by Intel, AMD, Nvidia, and ARM, and according to a report in 2011 by Evans Data, OpenCL had become the second most popular HPC tool.^[39]

2010s

In 2010, Nvidia partnered with [Audi](#) to power their cars' dashboards, using the [Tegra](#) GPU to provide increased functionality to cars' navigation and entertainment systems.^[40] Advances in GPU technology in cars helped advance [self-driving technology](#).^[41] AMD's [Radeon HD 6000 Series](#) cards were released in 2010, and in 2011 AMD released its 6000M Series discrete GPUs for mobile devices.^[42] The Kepler line of graphics cards by Nvidia were released in 2012 and were used in the Nvidia's 600 and 700 series cards. A feature in this GPU microarchitecture included GPU boost, a technology that adjusts the clock-speed of a video card to increase or decrease it according to its power draw.^[43] The Kepler microarchitecture was manufactured [on the 28 nm process](#).

The [PS4](#) and [Xbox One](#) were released in 2013; they both use GPUs based on [AMD's Radeon HD 7850 and 7790](#).^[44] Nvidia's Kepler line of GPUs was followed by the [Maxwell](#) line, manufactured on the same process. Nvidia's 28 nm chips were manufactured by [TSMC](#) in Taiwan using the 28 nm process. Compared to the 40 nm technology from the past, this manufacturing process allowed a 20 percent boost in performance while drawing less power.^{[45][46]} [Virtual reality headsets](#) have high system requirements; manufacturers recommended the [GTX 970](#) and the [R9 290X](#) or better at the time of their release.^{[47][48]} Cards based on the [Pascal](#) microarchitecture were released in 2016. The [GeForce 10 series](#) of cards are of this generation of graphics cards. They are made using the 16 nm manufacturing process which improves upon previous microarchitectures.^[49] Nvidia released one non-consumer card under the new [Volta](#) architecture, the [Titan V](#). Changes from the [Titan XP](#), Pascal's high-end card, include an increase in the number of CUDA cores, the addition of tensor cores, and [HBM2](#). Tensor cores are designed for deep

learning, while high-bandwidth memory is on-die, stacked, lower-clocked memory that offers an extremely wide memory bus. To emphasize that the Titan V is not a gaming card, Nvidia removed the "GeForce GTX" suffix it adds to consumer gaming cards.

In 2018, Nvidia launched the RTX 20 series GPUs that added ray-tracing cores to GPUs, improving their performance on lighting effects.^[50] Polaris 11 and Polaris 10 GPUs from AMD are fabricated by a 14 nm process. Their release resulted in a substantial increase in the performance per watt of AMD video cards.^[51] AMD also released the Vega GPU series for the high end market as a competitor to Nvidia's high end Pascal cards, also featuring HBM2 like the Titan V.

In 2019, AMD released the successor to their Graphics Core Next (GCN) microarchitecture/instruction set. Dubbed RDNA, the first product featuring it was the Radeon RX 5000 series of video cards.^[52]

The company announced that the successor to the RDNA microarchitecture would be incremental (aka a refresh). AMD unveiled the Radeon RX 6000 series, its RDNA 2 graphics cards with support for hardware-accelerated ray tracing.^[53] The product series, launched in late 2020, consisted of the RX 6800, RX 6800 XT, and RX 6900 XT.^{[54][55]} The RX 6700 XT, which is based on Navi 22, was launched in early 2021.^[56]

The PlayStation 5 and Xbox Series X and Series S were released in 2020; they both use GPUs based on the RDNA 2 microarchitecture with incremental improvements and different GPU configurations in each system's implementation.^{[57][58][59]}

Intel first entered the GPU market in the late 1990s, but produced lackluster 3D accelerators compared to the competition at the time. Rather than attempting to compete with the high-end manufacturers Nvidia and ATI/AMD, they began integrating Intel Graphics Technology GPUs into motherboard chipsets, beginning with the Intel 810 for the Pentium III, and later into CPUs. They began with the Intel Atom 'Pineview' laptop processor in 2009, continuing in 2010 with desktop processors in the first generation of the Intel Core line and with contemporary Pentiums and Celerons. This resulted in a large nominal market share, as the majority of computers with an Intel CPU also featured this embedded graphics processor. These generally lagged behind discrete processors in performance. Intel re-entered the discrete GPU market in 2022 with its Arc series, which competed with the then-current GeForce 30 series and Radeon 6000 series cards at competitive prices.

2020s

In the 2020s, GPUs have been increasingly used for calculations involving embarrassingly parallel problems, such as training of neural networks on enormous datasets that are needed for large language models. Specialized processing cores on some modern workstation's GPUs are dedicated for deep learning since they have significant FLOPS performance increases, using 4×4 matrix multiplication and division, resulting in hardware performance up to 128 TFLOPS in some applications.^[60] These tensor cores are expected to appear in consumer cards, as well.^[61]

GPU companies

Many companies have produced GPUs under a number of brand names. In 2009, Intel, Nvidia, and AMD/ATI were the market share leaders, with 49.4%, 27.8%, and 20.6% market share respectively. In addition, Matrox^[62] produces GPUs. Modern smartphones use mostly Adreno GPUs from Qualcomm, PowerVR GPUs from Imagination Technologies, and Mali GPUs from ARM.

Computational functions

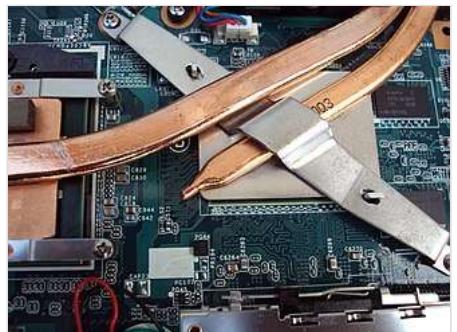
Modern GPUs use most of their transistors to do calculations related to 3D computer graphics. In addition to the 3D hardware, today's GPUs include basic 2D acceleration and framebuffer capabilities (usually with a VGA compatibility mode). Newer cards such as AMD/ATI HD5000–HD7000 lack dedicated 2D acceleration; it is emulated by 3D hardware. GPUs were initially used to accelerate the memory-intensive work of texture mapping and rendering polygons. Later, units were added to accelerate geometric calculations such as the rotation and translation of vertices into different coordinate systems. Recent developments in GPUs include support for programmable shaders which can manipulate vertices and textures with many of the same operations that are supported by CPUs, oversampling and interpolation techniques to reduce aliasing, and very high-precision color spaces.

Several factors of GPU construction affect the performance of the card for real-time rendering, such as the size of the connector pathways in the semiconductor device fabrication, the clock signal frequency, and the number and size of various on-chip memory caches. Performance is also affected by the number of streaming multiprocessors (SM) for NVidia GPUs, or compute units (CU) for AMD GPUs, or Xe cores for Intel discrete GPUs, which describe the number of core on-silicon processor units within the GPU chip that perform the core calculations, typically working in parallel with other SM/CUs on the GPU. GPU performance is typically measured in floating point operations per second (FLOPS); GPUs in the 2010s and 2020s typically deliver performance measured in teraflops (TFLOPS). This is an estimated performance measure, as other factors can affect the actual display rate.^[63]

GPU accelerated video decoding and encoding

Most GPUs made since 1995 support the YUV color space and hardware overlays, important for digital video playback, and many GPUs made since 2000 also support MPEG primitives such as motion compensation and iDCT. This hardware-accelerated video decoding, in which portions of the video decoding process and video post-processing are offloaded to the GPU hardware, is commonly referred to as "GPU accelerated video decoding", "GPU assisted video decoding", "GPU hardware accelerated video decoding", or "GPU hardware assisted video decoding".

Recent graphics cards decode high-definition video on the card, offloading the central processing unit. The most common APIs for GPU accelerated video decoding are DxVA for Microsoft Windows operating systems and VDPAU, VAAPI, XvMC, and XvBA for Linux-based and UNIX-like operating systems. All except XvMC are capable of decoding videos encoded with MPEG-1, MPEG-2, MPEG-4 ASP (MPEG-4 Part 2), MPEG-4 AVC (H.264 / DivX 6), VC-1, WMV3/WMV9, Xvid / OpenDivX (DivX 4), and DivX 5 codecs, while XvMC is only capable of decoding MPEG-1 and MPEG-2.



The ATI HD5470 GPU (above, with copper heatpipe attached) features UVD 2.1 which enables it to decode AVC and VC-1 video formats.

There are several dedicated hardware video decoding and encoding solutions.

Video decoding processes that can be accelerated

Video decoding processes that can be accelerated by modern GPU hardware are:

- Motion compensation (mocomp)
- Inverse discrete cosine transform (iDCT)
 - Inverse telecine 3:2 and 2:2 pull-down correction
- Inverse modified discrete cosine transform (iMDCT)
- In-loop deblocking filter
- Intra-frame prediction
- Inverse quantization (IQ)
- Variable-length decoding (VLD), more commonly known as slice-level acceleration
- Spatial-temporal deinterlacing and automatic interlace/progressive source detection
- Bitstream processing (Context-adaptive variable-length coding/Context-adaptive binary arithmetic coding) and perfect pixel positioning

These operations also have applications in video editing, encoding, and transcoding.

2D graphics APIs

An earlier GPU may support one or more 2D graphics API for 2D acceleration, such as GDI and DirectDraw.^[64]

3D graphics APIs

A GPU can support one or more 3D graphics API, such as DirectX, Metal, OpenGL, OpenGL ES, Vulkan.

GPU forms

Terminology

In the 1970s, the term "GPU" originally stood for *graphics processor unit* and described a programmable processing unit working independently from the CPU that was responsible for graphics manipulation and output.^{[65][66]} In 1994, Sony used the term (now standing for *graphics processing unit*) in reference to the PlayStation console's Toshiba-designed Sony GPU.^[31] The term was popularized by Nvidia in 1999, who marketed the GeForce 256 as "the world's first GPU".^[67] It was presented as a "single-chip processor with integrated transform, lighting, triangle setup/clipping, and rendering engines".^[68] Rival ATI Technologies coined the term "**visual processing unit**" or **VPU** with the release of the Radeon 9700 in 2002.^[69] The AMD Alveo MA35D features dual VPU's, each using the 5 nm process in 2023.^[70]

In personal computers, there are two main forms of GPUs. Each has many synonyms:^[71]

- Dedicated graphics also called *discrete graphics*.

- *Integrated graphics* also called *shared graphics solutions*, *integrated graphics processors* (IGP), or *unified memory architecture* (UMA).

Usage-specific GPU

Most GPUs are designed for a specific use, real-time 3D graphics, or other mass calculations:

1. Gaming

- [GeForce GTX, RTX](#)
- [Nvidia Titan](#)
- [Radeon HD, R5, R7, R9, RX, Vega and Navi series](#)
- [Radeon VII](#)
- [Intel Arc](#)

2. Cloud Gaming

- [Nvidia GRID](#)
- [Radeon Sky](#)

3. Workstation

- [Nvidia Quadro](#)
- [Nvidia RTX](#)
- [AMD FirePro](#)
- [AMD Radeon Pro](#)
- [Intel Arc Pro](#)

4. Cloud Workstation

- [Nvidia Tesla](#)
- [AMD FireStream](#)

5. Artificial Intelligence training and Cloud

- [Nvidia Tesla](#)
- [AMD Radeon Instinct](#)

6. Automated/Driverless car

- [Nvidia Drive PX](#)

Dedicated graphics processing unit

Dedicated graphics processing units are not necessarily removable, nor does it necessarily interface with the motherboard in a standard fashion. The term "dedicated" refers to the fact that *graphics cards* have *RAM* that is dedicated to the card's use, not to the fact that *most* dedicated GPUs are removable. This RAM is usually specially selected for the expected serial workload of the graphics card (see [GDDR](#)). Sometimes systems with dedicated *discrete* GPUs were called "DIS" systems as opposed to "UMA" systems (see next section).^[72] Dedicated GPUs for portable computers are most commonly interfaced through a non-standard and often proprietary slot due to size and weight constraints. Such ports may still be considered PCIe or AGP in terms of their logical host interface, even if they are not physically interchangeable with their counterparts.

Graphics cards with dedicated GPUs typically interface with the *motherboard* by means of an *expansion slot* such as *PCI Express* (PCIe) or *Accelerated Graphics Port* (AGP). They can usually be replaced or upgraded with relative ease, assuming the motherboard is capable of supporting the

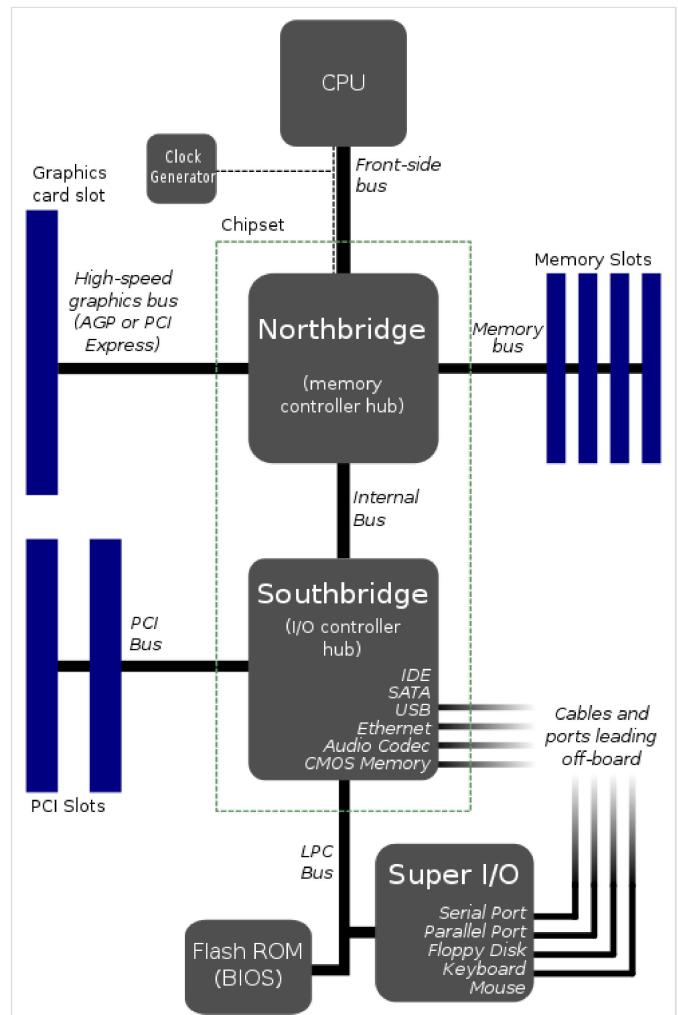
upgrade. A few graphics cards still use Peripheral Component Interconnect (PCI) slots, but their bandwidth is so limited that they are generally used only when a PCIe or AGP slot is not available.

Technologies such as [SLI](#) and [NVLink](#) by Nvidia and [CrossFire](#) by AMD allow multiple GPUs to draw images simultaneously for a single screen, increasing the processing power available for graphics. These technologies, however, are increasingly uncommon; most games do not fully use multiple GPUs, as most users cannot afford them.^{[73][74][75]} Multiple GPUs are still used on supercomputers (like in Summit), on workstations to accelerate video (processing multiple videos at once)^{[76][77][78]} and 3D rendering,^[79] for VFX,^[80] GPGPU workloads and for simulations,^[81] and in AI to expedite training, as is the case with Nvidia's lineup of DGX workstations and servers, Tesla GPUs, and Intel's Ponte Vecchio GPUs.

Integrated graphics processing unit

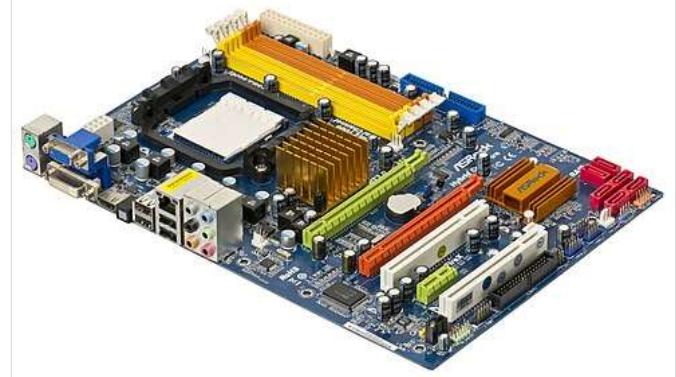
Integrated graphics processing unit (IGPU), integrated graphics, shared graphics solutions, integrated graphics processors (IGP), or unified memory architecture (UMA) use a portion of a computer's system RAM rather than dedicated graphics memory. IGPs can be integrated onto a motherboard as part of its [northbridge](#) chipset,^[82] or on the same die ([integrated circuit](#)) with the CPU (like [AMD APU](#) or [Intel HD Graphics](#)). On certain motherboards,^[83] AMD's IGPs can use dedicated sideport memory: a separate fixed block of high performance memory that is dedicated for use by the GPU. As of early 2007 computers with integrated graphics account for about 90% of all PC shipments.^[84] They are less costly to implement than dedicated graphics processing, but tend to be less capable. Historically, integrated processing was considered unfit for 3D games or graphically intensive programs but could run less intensive programs such as Adobe Flash. Examples of such IGPs would be offerings from SiS and VIA circa 2004.^[85] However, modern integrated graphics processors such as [AMD Accelerated Processing Unit](#) and [Intel Graphics Technology \(HD, UHD, Iris, Iris Pro, Iris Plus, and Xe-LP\)](#) can handle 2D graphics or low-stress 3D graphics.

Since GPU computations are memory-intensive, integrated processing may compete with the CPU for relatively slow system RAM, as it has minimal or no dedicated video memory. IGPs use system memory with bandwidth up to a current maximum of 128 GB/s, whereas a discrete graphics card may have a bandwidth of more than 1000 GB/s between its VRAM and GPU core. This [memory bus](#) bandwidth can limit the performance of the GPU, though [multi-channel memory](#) can mitigate this deficiency.^[86] Older integrated graphics chipsets lacked hardware transform and lighting, but newer ones include it.^{[87][88]}



The position of an integrated GPU in a northbridge/southbridge system layout

On systems with "Unified Memory Architecture" (UMA), including modern AMD processors with integrated graphics,^[89] modern Intel processors with integrated graphics,^[90] Apple processors, the PS5 and Xbox Series (among others), the CPU cores and the GPU block share the same pool of RAM and memory address space. This allows the system to dynamically allocate memory between the CPU cores and the GPU block based on memory needs (without needing a large static split of the RAM) and thanks to zero copy transfers, removes the need for either copying data over a bus (computing) between physically separate RAM pools or copying between separate address spaces on a single physical pool of RAM, allowing more efficient transfer of data.



An ASRock motherboard with integrated graphics, which has HDMI, VGA and DVI-out ports

Hybrid graphics processing

Hybrid GPUs compete with integrated graphics in the low-end desktop and notebook markets. The most common implementations of this are ATI's HyperMemory and Nvidia's TurboCache.

Hybrid graphics cards are somewhat more expensive than integrated graphics, but much less expensive than dedicated graphics cards. They share memory with the system and have a small dedicated memory cache, to make up for the high latency of the system RAM. Technologies within PCI Express make this possible. While these solutions are sometimes advertised as having as much as 768 MB of RAM, this refers to how much can be shared with the system memory.

Stream processing and general purpose GPUs (GPGPU)

It is common to use a general purpose graphics processing unit (GPGPU) as a modified form of stream processor (or a vector processor), running compute kernels. This turns the massive computational power of a modern graphics accelerator's shader pipeline into general-purpose computing power. In certain applications requiring massive vector operations, this can yield several orders of magnitude higher performance than a conventional CPU. The two largest discrete (see "Dedicated graphics processing unit" above) GPU designers, AMD and Nvidia, are pursuing this approach with an array of applications. Both Nvidia and AMD teamed with Stanford University to create a GPU-based client for the Folding@home distributed computing project for protein folding calculations. In certain circumstances, the GPU calculates forty times faster than the CPUs traditionally used by such applications.^{[91][92]}

GPGPUs can be used for many types of embarrassingly parallel tasks including ray tracing. They are generally suited to high-throughput computations that exhibit data-parallelism to exploit the wide vector width SIMD architecture of the GPU.

GPU-based high performance computers play a significant role in large-scale modelling. Three of the ten most powerful supercomputers in the world take advantage of GPU acceleration.^[93]

GPUs support API extensions to the C programming language such as OpenCL and OpenMP. Furthermore, each GPU vendor introduced its own API which only works with their cards: AMD APP SDK from AMD, and CUDA from Nvidia. These allow functions called compute kernels to run on the GPU's stream processors. This makes it possible for C programs to take advantage of a

GPU's ability to operate on large buffers in parallel, while still using the CPU when appropriate. CUDA was the first API to allow CPU-based applications to directly access the resources of a GPU for more general purpose computing without the limitations of using a graphics API.

Since 2005 there has been interest in using the performance offered by GPUs for evolutionary computation in general, and for accelerating the fitness evaluation in genetic programming in particular. Most approaches compile linear or tree programs on the host PC and transfer the executable to the GPU to be run. Typically a performance advantage is only obtained by running the single active program simultaneously on many example problems in parallel, using the GPU's SIMD architecture.^[94] However, substantial acceleration can also be obtained by not compiling the programs, and instead transferring them to the GPU, to be interpreted there.^[95] Acceleration can then be obtained by either interpreting multiple programs simultaneously, simultaneously running multiple example problems, or combinations of both. A modern GPU can simultaneously interpret hundreds of thousands of very small programs.

External GPU (eGPU)

An external GPU is a graphics processor located outside of the housing of the computer, similar to a large external hard drive. External graphics processors are sometimes used with laptop computers. Laptops might have a substantial amount of RAM and a sufficiently powerful central processing unit (CPU), but often lack a powerful graphics processor, and instead have a less powerful but more energy-efficient on-board graphics chip. On-board graphics chips are often not powerful enough for playing video games, or for other graphically intensive tasks, such as editing video or 3D animation/rendering.

Therefore, it is desirable to attach a GPU to some external bus of a notebook. PCI Express is the only bus used for this purpose. The port may be, for example, an ExpressCard or mPCIe port (PCIe $\times 1$, up to 5 or 2.5 Gbit/s respectively), a Thunderbolt 1, 2, or 3 port (PCIe $\times 4$, up to 10, 20, or 40 Gbit/s respectively), or an OCuLink port. Those ports are only available on certain notebook systems.^{[96][97]} eGPU enclosures include their own power supply (PSU), because powerful GPUs can consume hundreds of watts.^[98]

Official vendor support for external GPUs has gained traction. A milestone was Apple's decision to support external GPUs with MacOS High Sierra 10.13.4.^[99] Several major hardware vendors (HP, Razer) released Thunderbolt 3 eGPU enclosures.^{[100][101]} This support fuels eGPU implementations by enthusiasts.^[102]

Sales

In 2013, 438.3 million GPUs were shipped globally and the forecast for 2014 was 414.2 million. However, by the third quarter of 2022, shipments of integrated GPUs totaled around 75.5 million units, down 19% year-over-year.^[103] ^[104]

See also

- Texture mapping unit (TMU)
- Render output unit (ROP)
- Brute force attack
- Computer hardware
- Computer monitor
- GPU cache

- [GPU virtualization](#)
- [Manycore processor](#)
- [Physics processing unit \(PPU\)](#)
- [Tensor processing unit \(TPU\)](#)
- [Ray-tracing hardware](#)
- [Software rendering](#)
- [Vision processing unit \(VPU\)](#)
- [Vector processor](#)
- [Video card](#)
- [Video display controller](#)
- [Video game console](#)
- [AI accelerator](#)
- [GPU Vector Processor internal features](#)

Hardware

- [List of AMD graphics processing units](#)
- [List of Nvidia graphics processing units](#)
- [List of Intel graphics processing units](#)
- [Intel GMA](#)
- [Larrabee](#)
- [Nvidia PureVideo](#) – the bit-stream technology from [Nvidia](#) used in their graphics chips to accelerate video decoding on hardware GPU with DXVA.
- [SoC](#)
- [UVD \(Unified Video Decoder\)](#) – the video decoding bit-stream technology from [ATI](#) to support hardware (GPU) decode with DXVA

APIs

- [OpenGL API](#)
- [DirectX Video Acceleration \(DxVA\) API](#) for [Microsoft Windows](#) operating-system.
- [Mantle \(API\)](#)
- [Vulkan \(API\)](#)
- [Video Acceleration API \(VA API\)](#)
- [VDPAU \(Video Decode and Presentation API for Unix\)](#)
- [X-Video Bitstream Acceleration \(XvBA\)](#), the X11 equivalent of DXVA for MPEG-2, H.264, and VC-1
- [X-Video Motion Compensation](#) – the X11 equivalent for MPEG-2 video codec only

Applications

- [GPU cluster](#)
- [Mathematica](#) – includes built-in support for CUDA and OpenCL GPU execution
- [Molecular modeling on GPU](#)
- [Deeplearning4j](#) – open-source, distributed deep learning for Java

References

1. Hague, James (September 10, 2013). "Why Do Dedicated Game Consoles Exist?" (<https://web.archive.org/web/20150504042057/https://prog21.dadgum.com/181.html>). *Programming in the 21st Century*. Archived from the original (<https://prog21.dadgum.com/181.html>) on May 4, 2015. Retrieved November 11, 2015.
2. "mame/8080bw.c at master 路 mamedev/mame 路 GitHub" (<https://archive.today/20141121125813/https://github.com/mamedev/mame/tree/master/src/mame/drivers/8080bw.c>). GitHub. Archived from the original (<https://github.com/mamedev/mame/tree/master/src/mame/drivers/8080bw.c>) on 2014-11-21.
 - "mame/mw8080bw.c at master 路 mamedev/mame 路 GitHub" (<https://archive.today/20141121125812/https://github.com/mamedev/mame/tree/master/src/mame/drivers/mw8080bw.c>). GitHub. Archived from the original (<https://github.com/mamedev/mame/tree/master/src/mame/drivers/mw8080bw.c>) on 2014-11-21.
 - "Arcade/SpaceInvaders – Computer Archeology" (<https://web.archive.org/web/20140913080838/https://www.computerarcheology.com/wiki/wiki/Arcade/SpaceInvaders>). computerarcheology.com. Archived from the original (<https://www.computerarcheology.com/wiki/wiki/Arcade/SpaceInvaders>) on 2014-09-13.
3. "mame/galaxian.c at master 路 mamedev/mame 路 GitHub" (<https://archive.today/20141121114447/https://github.com/mamedev/mame/tree/master/src/mame/video/galaxian.c>). GitHub. Archived from the original (<https://github.com/mamedev/mame/tree/master/src/mame/video/galaxian.c>) on 2014-11-21.
4. "mame/galaxian.c at master 路 mamedev/mame 路 GitHub" (<https://archive.today/20141121114529/https://github.com/mamedev/mame/tree/master/src/mame/drivers/galaxian.c>). GitHub. Archived from the original (<https://github.com/mamedev/mame/tree/master/src/mame/drivers/galaxian.c>) on 2014-11-21.
 - "MAME – src/mame/drivers/galdrvr.c" (<https://web.archive.org/web/20140103070737/http://mamedev.org/source/src/mame/drivers/galdrvr.c.html>). Archived from the original (<http://mamedev.org/source/src/mame/drivers/galdrvr.c.html>) on 3 January 2014.
5. Springmann, Alessondra. "Atari 2600 Teardown: What's Inside Your Old Console?" (<https://www.washingtonpost.com/wp-dyn/content/article/2010/09/01/AR2010090103543.html>). The Washington Post. Archived (<https://web.archive.org/web/20150714082924/https://www.washingtonpost.com/wp-dyn/content/article/2010/09/01/AR2010090103543.html>) from the original on July 14, 2015. Retrieved July 14, 2015.
6. "What are the 6502, ANTIC, CTIA/GTIA, POKEY, and FREDDIE chips?" (<https://web.archive.org/web/20160305010645/https://www.atari8.com/node/31>). Atari8.com. Archived from the original (<https://www.atari8.com/node/31>) on 2016-03-05.
7. Wiegers, Karl E. (April 1984). "Atari Display List Interrupts" (https://www.atarimagazines.com/compute/issue47/153_1_Atari_Display_List_Interrupts.php). Compute! (47): 161. Archived (https://web.archive.org/web/20160304035625/https://www.atarimagazines.com/compute/issue47/153_1_Atari_Display_List_Interrupts.php) from the original on 2016-03-04.
8. Wiegers, Karl E. (December 1985). "Atari Fine Scrolling" (https://www.atarimagazines.com/compute/issue67/338_1_Atari_Fine_Scrolling.php). Compute! (67): 110. Archived (https://web.archive.org/web/20060216181611/https://www.atarimagazines.com/compute/issue67/338_1_Atari_Fine_Scrolling.php) from the original on 2006-02-16.
9. Hopgood, F. Robert A.; Hubbald, Roger J.; Duce, David A., eds. (1986). *Advances in Computer Graphics II* (https://books.google.com/books?id=2j4hTAqxJ_sC&pg=PA169). Springer. p. 169. ISBN 9783540169109. "Perhaps the best known one is the NEC 7220."
10. Anderson, Marian (2018-07-18). "Famous Graphics Chips: NEC μPD7220 Graphics Display Controller" (<https://www.computer.org/publications/tech-news/chasing-pixels/famous-graphics-chips/>). IEEE Computer Society. Retrieved 2023-10-17.
11. Riddle, Sean. "Blitter Information" (<https://seanriddle.com/blitter.html>). Archived (<https://web.archive.org/web/20151222155908/https://seanriddle.com/blitter.html>) from the original on 2015-12-22.

12. Wolf, Mark J. P. (June 2012). *Before the Crash: Early Video Game History* (<https://books.google.com/books?id=oK3D4i5ldKgC>). Wayne State University Press. p. 185. **ISBN 978-0814337226**.
13. Anderson, Marian (2018-10-07). "GPU History: Hitachi ARTC HD63484" (<https://www.computer.org/publications/tech-news/chasing-pixels/gpu-history-hitachi-artc-hd63484/>). IEEE Computer Society. Retrieved 2023-10-17.
14. "Famous Graphics Chips: TI TMS34010 and VRAM. The first programmable graphics processor chip | IEEE Computer Society" (<https://www.computer.org/publications/tech-news/chasing-pixels/Famous-Graphics-Chips-IBMs-professional-graphics-the-PGC-and-8514A/Famous-Graphics-Chips-TI-TMS34010-and-VRAM>). 10 January 2019.
15. "X68000" (<https://nfgames.com/games/x68k/>). Archived (<https://web.archive.org/web/20140903010307/https://nfgames.com/games/x68k/>) from the original on 2014-09-03. Retrieved 2014-09-12.
16. "museum ~ Sharp X68000" (<https://web.archive.org/web/20150219114323/https://www.old-computers.com/museum/computer.asp?st=1&c=298>). Old-computers.com. Archived from the original (<https://www.old-computers.com/museum/computer.asp?st=1&c=298>) on 2015-02-19. Retrieved 2015-01-28.
17. "Hardcore Gaming 101: Retro Japanese Computers: Gaming's Final Frontier" (<https://www.hardcoregaming101.net/JPNcomputers/Japanesecomputers2.htm>). hardcoregaming101.net. Archived (<https://web.archive.org/web/20110113214919/https://www.hardcoregaming101.net/JPNcomputers/Japanesecomputers2.htm>) from the original on 2011-01-13.
18. "System 16 – Namco System 21 Hardware (Namco)" (<https://www.system16.com/hardware.php?id=536>). system16.com. Archived (<https://web.archive.org/web/20150518005344/https://system16.com/hardware.php?id=536>) from the original on 2015-05-18.
19. "System 16 – Taito Air System Hardware (Taito)" (<https://www.system16.com/hardware.php?id=656>). system16.com. Archived (<https://web.archive.org/web/20150316214310/https://system16.com/hardware.php?id=656>) from the original on 2015-03-16.
20. Brownstein, Mark (November 14, 1988). "NEC Forms Video Standards Group" (<https://books.google.com/books?id=wTsEAAAAMBAJ&pg=PT2>). *InfoWorld*. Vol. 10, no. 46. p. 3. **ISSN 0199-6649** (<https://www.worldcat.org/issn/0199-6649>). Retrieved May 27, 2016.
21. "S3 Video Boards" (<https://books.google.com/books?id=XIEAAAAMBAJ&q=S3+86C911&pg=PA62>). *InfoWorld*. 14 (20): 62. May 18, 1992. Archived (https://web.archive.org/web/20171122191720/https://books.google.com/books?id=XIEAAAAMBAJ&pg=PA62&dq=S3+86C911&hl=en&sa=X&ved=0CBwQ6AEwAGoVChMI85bTh7XYxglVyyeUCh3Z_wYp#v=onepage&q=S3%2086C911&f=false) from the original on November 22, 2017. Retrieved July 13, 2015.
22. "What the numbers mean" (<https://books.google.com/books?id=4RN8nH8oZ2QC&pg=128>). *PC Magazine*. 12: 128. 23 February 1993. Archived (<https://web.archive.org/web/20170411172653/https://books.google.com/books?id=4RN8nH8oZ2QC&pg=128>) from the original on 11 April 2017. Retrieved 29 March 2016.
 - Singer, Graham. "The History of the Modern Graphics Processor" (<https://www.techspot.com/article/650-history-of-the-gpu/>). Techspot. Archived (<https://web.archive.org/web/20160329140009/https://www.techspot.com/article/650-history-of-the-gpu/>) from the original on 29 March 2016. Retrieved 29 March 2016.
23. "System 16 – Namco Magic Edge Hornet Simulator Hardware (Namco)" (<https://www.system16.com/hardware.php?id=832>). system16.com. Archived (<https://web.archive.org/web/20140912000953/https://www.system16.com/hardware.php?id=832>) from the original on 2014-09-12.
24. "MAME – src/mame/video/model2.c" (<https://web.archive.org/web/20130104200822/https://mamedev.org/source/src/mame/video/model2.c.html>). Archived from the original (<https://mamedev.org/source/src/mame/video/model2.c.html>) on 4 January 2013.
25. "System 16 – Sega Model 2 Hardware (Sega)" (<https://www.system16.com/hardware.php?id=713>). system16.com. Archived (<https://web.archive.org/web/20101221001009/https://system16.com/hardware.php?id=713>) from the original on 2010-12-21.

26. "3D Graphics Processor Chip Set" (https://web.archive.org/web/20161011194640/https://www.hotchips.org/wp-content/uploads/hc_archives/hc07/3_Tue/HC7.S5/HC7.5.1.pdf) (PDF). Archived from the original (https://www.hotchips.org/wp-content/uploads/hc_archives/hc07/3_Tue/HC7.S5/HC7.5.1.pdf) (PDF) on 2016-10-11. Retrieved 2016-08-08.
- "3D-CG System with Video Texturing for Personal Computers" (<https://web.archive.org/web/20140906132147/https://www.fujitsu.com/downloads/MAG/vol33-2/paper08.pdf>) (PDF). Archived from the original (<https://www.fujitsu.com/downloads/MAG/vol33-2/paper08.pdf>) (PDF) on 2014-09-06. Retrieved 2016-08-08.
27. "Fujitsu Develops World's First Three Dimensional Geometry Processor" (<https://pr.fujitsu.com/jp/news/1997/Jul/2e.html>). *fujitsu.com*. Archived (<https://web.archive.org/web/20140912004938/https://pr.fujitsu.com/jp/news/1997/Jul/2e.html>) from the original on 2014-09-12.
28. "The Nintendo 64 is one of the greatest gaming devices of all time" (<https://xenol.kinja.com/the-nintendo-64-is-one-of-the-greatest-gaming-devices-o-1722364688>). *xenol*. Archived (<https://web.archive.org/web/20151118044413/https://xenol.kinja.com/the-nintendo-64-is-one-of-the-greatest-gaming-devices-o-1722364688>) from the original on 2015-11-18.
29. "Mitsubishi's 3DPro/2mp Chipset Sets New Records for Fastest 3D Graphics Accelerator for Windows NT Systems; 3DPro/2mp grabs Viewperf performance lead; other high-end benchmark tests clearly show that 3DPro's performance outdistances all Windows NT competitors" (<https://web.archive.org/web/20181115235023/https://www.thefreelibrary.com/Mitsubishi%27s+3DPro%2f2mp+Chipset+Sets+New+Records+for+Fastest+3D...-a019465188>). Archived from the original (<https://www.thefreelibrary.com/Mitsubishi%27s+3DPro/2mp+Chipset+Sets+New+Records+for+Fastest+3D...-a019465188>) on 2018-11-15. Retrieved 2022-02-18.
30. Vlask. "VGA Legacy MKIII – Diamond Fire GL 4000 (Mitsubishi 3DPro/2mp)" (<https://vgamuseum.info/index.php/component/k2/item/547-diamond-fire-gl-4000-mitsubishi-3dpro-2mp>). Archived (<https://web.archive.org/web/20151118114320/https://vgamuseum.info/index.php/component/k2/item/547-diamond-fire-gl-4000-mitsubishi-3dpro-2mp>) from the original on 2015-11-18.
31. "Is it Time to Rename the GPU? | IEEE Computer Society" (<https://www.computer.org/publications/tech-news/chasing-pixels/is-it-time-to-rename-the-gpu>). 17 July 2018.
32. **3dfx Glide API**
33. Dreijer, Søren. "Bump Mapping Using CG (3rd Edition)" (https://web.archive.org/web/20100120195901/http://www.blacksmith-studios.dk/projects/downloads/bumpmapping_using_cg.php). Archived from the original (https://www.blacksmith-studios.dk/projects/downloads/bumpmapping_using_cg.php) on 2010-01-20. Retrieved 2007-05-30.
34. Raina, Rajat; Madhavan, Anand; Ng, Andrew Y. (2009-06-14). "Large-scale deep unsupervised learning using graphics processors". *Proceedings of the 26th Annual International Conference on Machine Learning – ICML '09*. Di.acm.org. pp. 1–8. doi:10.1145/1553374.1553486 (<https://doi.org/10.1145/1553374.1553486>). ISBN 9781605585161. S2CID 392458 (<https://api.semanticscholar.org/CorpusID:392458>).
35. "Linear algebra operators for GPU implementation of numerical algorithms" (<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.1988&rep=rep1&type=pdf>), Kruger and Westermann, International Conference on Computer Graphics and Interactive Techniques, 2005
36. Liepe; et al. (2010). "ABC-SysBio—approximate Bayesian computation in Python with GPU support" (<https://web.archive.org/web/20151105144736/https://bioinformatics.oxfordjournals.org/content/26/14/1797.full>). *Bioinformatics*. 26 (14): 1797–1799. doi:10.1093/bioinformatics/btq278 (<https://doi.org/10.1093/bioinformatics/btq278>). PMC 2894518 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2894518>). PMID 20591907 (<https://pubmed.ncbi.nlm.nih.gov/20591907>). Archived from the original (<https://bioinformatics.oxfordjournals.org/content/26/14/1797.full>) on 2015-11-05. Retrieved 2010-10-15.

37. Sanders, Jason; Kandrot, Edward (2010-07-19). *CUDA by Example: An Introduction to General-Purpose GPU Programming, Portable Documents* (<https://books.google.com/books?id=49OmnOmTEtQC>). Addison-Wesley Professional. ISBN 9780132180139. Archived (<https://web.archive.org/web/20170412034641/https://books.google.com/books?id=49OmnOmTEtQC>) from the original on 2017-04-12.
38. "OpenCL – The open standard for parallel programming of heterogeneous systems" (<https://www.khronos.org/opencl/>). khronos.org. Archived (<https://web.archive.org/web/20110809103233/https://www.khronos.org/opencl/>) from the original on 2011-08-09.
39. Handy, Alex (2011-09-28). "AMD helps OpenCL gain ground in HPC space" (<https://sdtimes.com/amd/amd-helps-opencl-gain-ground-in-hpc-space/>). SD Times. Retrieved 2023-06-04.
40. Teglet, Traian (8 January 2010). "NVIDIA Tegra Inside Every Audi 2010 Vehicle" (<https://news.softpedia.com/news/NVIDIA-Tegra-Inside-Every-Audi-2010-Vehicle-131529.shtml>). Archived (<https://web.archive.org/web/20161004185422/https://news.softpedia.com/news/NVIDIA-Tegra-Inside-Every-Audi-2010-Vehicle-131529.shtml>) from the original on 2016-10-04. Retrieved 2016-08-03.
41. "School's in session – Nvidia's driverless system learns by watching" (<https://www.digitaltrends.com/cars/nvidia-gpu-driverless-car/>). 2016-04-30. Archived (<https://web.archive.org/web/2016051203712/https://www.digitaltrends.com/cars/nvidia-gpu-driverless-car/>) from the original on 2016-05-01. Retrieved 2016-08-03.
42. "AMD Radeon HD 6000M series – don't call it ATI!" (<https://www.cnet.com/news/amd-radeon-hd-6000m-series-dont-call-it-ati/>). CNET. Archived (<https://web.archive.org/web/20161011195008/https://www.cnet.com/news/amd-radeon-hd-6000m-series-dont-call-it-ati/>) from the original on 2016-10-11. Retrieved 2016-08-03.
43. "Nvidia GeForce GTX 680 2GB Review" (<https://www.bit-tech.net/hardware/2012/03/22/nvidia-geforce-gtx-680-2gb-review/4>). Archived (<https://web.archive.org/web/20160911210258/https://www.bit-tech.net/hardware/2012/03/22/nvidia-geforce-gtx-680-2gb-review/4>) from the original on 2016-09-11. Retrieved 2016-08-03.
44. "Xbox One vs. PlayStation 4: Which game console is best?" (<https://www.extremetech.com/gaming/156273-xbox-720-vs-ps4-vs-pc-how-the-hardware-specs-compare>). ExtremeTech. Retrieved 2019-05-13.
45. "Kepler TM GK110" (<https://www.nvidia.com/content/PDF/kepler/NVIDIA-Kepler-GK110-Architecture-Whitepaper.pdf>) (PDF). NVIDIA Corporation. 2012. Archived (<https://web.archive.org/web/20161011194501/https://www.nvidia.com/content/PDF/kepler/NVIDIA-Kepler-GK110-Architecture-Whitepaper.pdf>) (PDF) from the original on October 11, 2016. Retrieved August 3, 2016.
46. "Taiwan Semiconductor Manufacturing Company Limited" (<https://www.tsmc.com/english/dedicatedFoundry/technology/28nm.htm>). www.tsmc.com. Archived (<https://web.archive.org/web/20160810092021/https://www.tsmc.com/english/dedicatedFoundry/technology/28nm.htm>) from the original on 2016-08-10. Retrieved 2016-08-03.
47. "Building a PC for the HTC Vive" (<https://www.octopusrift.com/building-a-vive-pc/>). 2016-06-16. Archived (<https://web.archive.org/web/20160729024408/https://www.octopusrift.com/building-a-vive-pc/>) from the original on 2016-07-29. Retrieved 2016-08-03.
48. "VIVE Ready Computers" (<https://www.vive.com/ready/>). Vive. Archived (<https://web.archive.org/web/20160224182407/https://www.htcvive.com/us/product-optimized/>) from the original on 2016-02-24. Retrieved 2021-07-30.
49. "Nvidia's monstrous Pascal GPU is packed with cutting-edge tech and 15 billion transistors" (<https://www.pcworld.com/article/3052312/components-graphics/nvidias-monstrous-pascal-gpu-is-packed-with-cutting-edge-tech-and-15-billion-transistors.html>). 5 April 2016. Archived (<https://web.archive.org/web/20160731220844/https://www.pcworld.com/article/3052312/components-graphics/nvidias-monstrous-pascal-gpu-is-packed-with-cutting-edge-tech-and-15-billion-transistors.html>) from the original on 2016-07-31. Retrieved 2016-08-03.
50. Sarkar, Samit (20 August 2018). "Nvidia RTX 2070, RTX 2080, RTX 2080 Ti GPUs revealed: specs, price, release date" (<https://www.polygon.com/2018/8/20/17760038/nvidia-geforce-rtx-2080-ti-2070-specs-release-date-price-turing>). Polygon. Retrieved 11 September 2019.

51. "AMD RX 480, 470 & 460 Polaris GPUs To Deliver The 'Most Revolutionary Jump In Performance' Yet" (<https://wccftech.com/amd-unveils-polaris-11-10-gpu/>). 2016-01-16. Archived (<https://web.archive.org/web/20160801225420/https://wccftech.com/amd-unveils-polaris-11-10-gpu/>) from the original on 2016-08-01. Retrieved 2016-08-03.
52. AMD press release: "AMD Announces Next-Generation Leadership Products at Computex 2019 Keynote" (<https://www.amd.com/en/press-releases/2019-05-26-amd-announces-next-generation-leadership-products-computex-2019-keynote>). AMD. Retrieved October 5, 2019.
53. "AMD to Introduce New Next-Gen RDNA GPUs in 2020, Not a Typical 'Refresh' of Navi" (<https://www.tomshardware.com/news/amds-navi-to-be-refreshed-with-next-gen-rdna-architecture-in-2020>). *Tom's Hardware*. 2020-01-29. Retrieved 2020-02-08.
 - Garreffa, Anthony (September 9, 2020). "AMD to reveal next-gen Big Navi RDNA 2 graphics cards on October 28" (<https://www.tweaktown.com/news/75066/amd-to-reveal-next-gen-big-navi-rdna-2-graphics-cards-on-october-28/index.html>). *TweakTown*. Retrieved September 9, 2020.
 - Lyles, Taylor (September 9, 2020). "AMD's next-generation Zen 3 CPUs and Radeon RX 6000 'Big Navi' GPU will be revealed next month" (<https://www.theverge.com/2020/9/9/21429127/amd-zn-3-cpu-big-navi-gpu-events-october>). *The Verge*. Retrieved September 10, 2020.
54. "AMD Teases Radeon RX 6000 Card Performance Numbers: Aiming For 3080?" (<https://www.anandtech.com/show/16150/amd-teases-radeon-rx-6000-card-performance-numbers-aiming-for-3080>). *AnandTech*. 2020-10-08. Retrieved 2020-10-25.
 - "AMD Announces Ryzen 'Zen 3' and Radeon 'RDNA2' Presentations for October: A New Journey Begins" (<https://www.anandtech.com/show/16077/amd-announces-ryzen-zn-3-and-radeon-rDNA2-presentations-for-october-a-new-journey-begins>). *AnandTech*. 2020-09-09. Retrieved 2020-10-25.
55. Judd, Will (October 28, 2020). "AMD unveils three Radeon 6000 graphics cards with ray tracing and RTX-beating performance" (<https://www.eurogamer.net/articles/digitalfoundry-2020-10-28-amd-unveils-three-radeon-6000-graphics-cards-with-ray-tracing-and-impressive-performance>). *Eurogamer*. Retrieved October 28, 2020.
56. Mujtaba, Hassan (2020-11-30). "AMD Radeon RX 6700 XT 'Navi 22 GPU' Custom Models Reportedly Boost Up To 2.95 GHz" (<https://wccftech.com/amd-radeon-rx-6700-xt-custom-models-boost-up-to-2-95-ghz-220w-tgp/>). *Wccftech*. Retrieved 2020-12-03.
 - Tyson, Mark (December 3, 2020). "AMD CEO keynote scheduled for CES 2020 on 12th January" (<https://hexus.net/tech/news/industry/147068-amd-ceo-keynote-scheduled-ces-2020-12th-january>). *HEXUS*. Retrieved 2020-12-03.
 - Cutress, Ian (January 12, 2021). "AMD to Launch Mid-Range RDNA 2 Desktop Graphics in First Half 2021" (<https://www.anandtech.com/show/16402/amd-to-launch-midrange-rDNA-2-desktop-graphics-in-first-half-2021>). *AnandTech*. Retrieved January 4, 2021.
57. Funk, Ben (December 12, 2020). "Sony PS5 Gets A Full Teardown Detailing Its RDNA 2 Guts And Glory" (<https://web.archive.org/web/20201212165054/https://hothardware.com/news/sony-playstation-5-independent-teardown-ifixit>). *Hot Hardware*. Archived from the original (<https://hothardware.com/news/sony-playstation-5-independent-teardown-ifixit>) on December 12, 2020. Retrieved January 3, 2021.
58. Gartenberg, Chaim (March 18, 2020). "Sony reveals full PS5 hardware specifications" (<https://www.theverge.com/2020/3/18/21183181/sony-ps5-playstation-5-specs-details-hardware-processor-8k-ray-tracing>). *The Verge*. Retrieved January 3, 2021.
59. Smith, Ryan. "Microsoft Drops More Xbox Series X Tech Specs: Zen 2 + RDNA 2, 12 TFLOPs GPU, HDMI 2.1, & a Custom SSD" (<https://www.anandtech.com/show/15546/microsoft-drops-more-xbox-series-x-tech-specs-zen-2-rDNA-2-12-tflops-gpu-hDMI-21-a-custom-ssd>). *AnandTech*. Retrieved 2020-03-19.

60. Smith, Ryan. "NVIDIA Volta Unveiled: GV100 GPU and Tesla V100 Accelerator Announced" (<https://www.anandtech.com/show/11367/nvidia-volta-unveiled-gv100-gpu-and-tesla-v100-accelerator-announced>). *AnandTech*. Retrieved 16 August 2018.
61. Hill, Brandon (11 August 2017). "AMD's Navi 7nm GPU Architecture to Reportedly Feature Dedicated AI Circuitry" (<https://web.archive.org/web/20180817023122/https://hothardware.com/news/amd-7nm-navi-gpu-ai-deep-learning>). *HotHardware*. Archived from the original (<https://hothardware.com/news/amd-7nm-navi-gpu-ai-deep-learning>) on 17 August 2018. Retrieved 16 August 2018.
62. "Matrox Graphics – Products – Graphics Cards" (https://www.matrox.com/graphics/en/products/graphics_cards). Matrox.com. Archived (https://web.archive.org/web/20140205042218/http://www.matrox.com/graphics/en/products/graphics_cards/) from the original on 2014-02-05. Retrieved 2014-01-21.
63. Hruska, Joel (February 10, 2021). "How Do Graphics Cards Work?" (<https://www.extremetech.com/gaming/269335-how-graphics-cards-work>). *Extreme Tech*. Retrieved July 17, 2021.
64. *CL-GD5446 64-bit VisualMedia Accelerator Preliminary Data Book* (<https://theretroweb.com/chip/documentation/cl-gd5446-smol-645654db1ae03286068933.pdf>) (PDF), Cirrus Logic, November 1996, retrieved 30 January 2024 – via The Datasheet Archive
65. Barron, E. T.; Glorioso, R. M. (September 1973). "A micro controlled peripheral processor". *Conference record of the 6th annual workshop on Microprogramming – MICRO 6*. pp. 122–128. doi:[10.1145/800203.806247](https://doi.org/10.1145/800203.806247) (<https://doi.org/10.1145%2F800203.806247>). ISBN 9781450377836. S2CID 36942876 (<https://api.semanticscholar.org/CorpusID:36942876>).
66. Levine, Ken (August 1978). "Core standard graphic package for the VGI 3400" (<https://dl.acm.org/doi/abs/10.1145/965139.807405>). *ACM SIGGRAPH Computer Graphics*. 12 (3): 298–300. doi:[10.1145/965139.807405](https://doi.org/10.1145/965139.807405) (<https://doi.org/10.1145%2F965139.807405>).
67. "NVIDIA Launches the World's First Graphics Processing Unit: GeForce 256" (https://www.nvidia.com/object/IO_20020111_5424.html). Nvidia. 31 August 1999. Archived (https://web.archive.org/web/20160412035751/https://www.nvidia.com/object/IO_20020111_5424.html) from the original on 12 April 2016. Retrieved 28 March 2016.
68. "Graphics Processing Unit (GPU)" (<https://www.nvidia.com/object/gpu.html>). Nvidia. 16 December 2009. Archived (<https://web.archive.org/web/20160408122443/https://www.nvidia.com/object/gpu.html>) from the original on 8 April 2016. Retrieved 29 March 2016.
69. Pabst, Thomas (18 July 2002). "ATI Takes Over 3D Technology Leadership With Radeon 9700" (<https://www.tomshardware.com/reviews/ati-takes-3d-technology-leadership-radeon-9700,491-3.html>). Tom's Hardware. Retrieved 29 March 2016.
70. Child, J. (6 April 2023). "AMD Rolls Out 5 nm ASIC-based Accelerator for the Interactive Streaming Era" (<https://www.allaboutcircuits.com/news/amd-rolls-out-5-nm-asic-based-accelerator-for-the-interactive-streaming-era>). EETech Media. Retrieved 24 December 2023.
71. "Help Me Choose: Video Cards" (<https://archive.today/20160909115302/http://www.dell.com/learn/us/en/19/help-me-choose/hmc-video-card-inspiron-lt?stp=1>). Dell. Archived from the original (<https://www.dell.com/learn/us/en/19/help-me-choose/hmc-video-card-inspiron-lt>) on 2016-09-09. Retrieved 2016-09-17.
72. "Nvidia Optimus documentation for Linux device driver" (<https://nouveau.freedesktop.org/Optimus.html>). freedesktop. 13 November 2023. Retrieved 24 December 2023.
73. Abazovic, F. (3 July 2015). "Crossfire and SLI market is just 300.000 units" (<https://www.fudzilla.com/news/graphics/38134-crossfire-and-sli-market-is-just-300-000-units>). fudzilla. Retrieved 24 December 2023.
74. "Is Multi-GPU Dead?" (<https://thetechaltar.com/is-multi-gpu-dead/>). 7 January 2018.
75. "Nvidia SLI and AMD CrossFire is dead – but should we mourn multi-GPU gaming? | TechRadar" (<https://www.techradar.com/amp/news/nvidia-sli-and-amd-crossfire-is-dead-but-should-we-mourn-multi-gpu-gaming>). 24 August 2019.
76. "NVIDIA FFmpeg Transcoding Guide" (<https://devblogs.nvidia.com/nvidia-ffmpeg-transcoding-guide/>). 24 July 2019.

77. "Hardware Selection and Configuration Guide DaVinci Resolve 15" (https://documents.blackmagicdesign.com/ConfigGuides/DaVinci_Resolve_15_Mac_Configuration_Guide.pdf) (PDF). BlackMagic Design. 2018. Retrieved 31 May 2022.
78. "Recommended System: Recommended Systems for DaVinci Resolve" (<https://www.pugetsystems.com/recommended/Recommended-Systems-for-DaVinci-Resolve-187/Hardware-Recommendations>). Puget Systems.
- "GPU Accelerated Rendering and Hardware Encoding" (<https://helpx.adobe.com/x-products/b/multi/gpu-acceleration-and-hardware-encoding.html>).
79. "V-Ray Next Multi-GPU Performance Scaling" (<https://www.pugetsystems.com/labs/articles/V-Ray-Next-Multi-GPU-Performance-Scaling-1559/>). 20 August 2019.
- "FAQ | GPU-accelerated 3D rendering software | Redshift" (<https://www.redshift3d.com/support/faq>).
 - "OctaneRender 2020™ Preview is here!" (<https://home.otoy.com/render/octane-render/faqs/>). Otoy.
 - "Exploring Performance with Autodesk's Arnold Renderer GPU Beta" (<https://techgage.com/article/autodesk-arnold-renderer-gpu-beta-performance/>). 8 April 2019.
 - "GPU Rendering – Blender Manual" (https://docs.blender.org/manual/en/latest/render/cycles/gpu_rendering.html).
80. "V-Ray for Nuke – Ray Traced Rendering for Compositors | Chaos Group" (<https://www.chaosgroup.com/vray/nuke>).
- "System Requirements | Nuke | Foundry" (<https://www.foundry.com/products/nuke/requirements>).
81. "What about multi-GPU support? – Folding@home" (<https://foldingathome.org/faqs/gpu2-common/frequently-asked-questions-common-ati-nvidia-gpu2-clients-2/multi-gpu-support/>).
82. "Evolution of Intel Graphics: I740 to Iris Pro" (<https://www.tomshardware.com/amp/picturestory/693-intel-graphics-evolution.html>). 4 February 2017.
83. "GA-890GPA-UD3H overview" (<https://web.archive.org/web/20150415095629/https://www.gigabyte.com/products/product-page.aspx?pid=3785#ov>). Archived from the original (<https://www.gigabyte.com/products/product-page.aspx?pid=3785#ov>) on 2015-04-15. Retrieved 2015-04-15.
84. Key, Gary. "AnandTech – µATX Part 2: Intel G33 Performance Review" (<https://www.anandtech.com/mb/showdoc.aspx?i=3111&p=23>). *anandtech.com*. Archived (<https://web.archive.org/web/20080531045027/https://www.anandtech.com/mb/showdoc.aspx?i=3111&p=23>) from the original on 2008-05-31.
85. Tscheblockov, Tim. "Xbit Labs: Roundup of 7 Contemporary Integrated Graphics Chipsets for Socket 478 and Socket A Platforms" (<https://web.archive.org/web/20070526124817/https://www.xbitlabs.com/articles/chipsets/display/int-chipsets-roundup.html>). Archived from the original (<https://www.xbitlabs.com/articles/chipsets/display/int-chipsets-roundup.html>) on 2007-05-26. Retrieved 2007-06-03.
86. Coelho, Rafael (18 January 2016). "Does dual-channel memory make difference on integrated video performance?" (<https://www.hardwaresecrets.com/dual-channel-memory-make-difference-in-integrated-video-performance/>). *Hardware Secrets*. Retrieved 4 January 2019.
87. Sanford, Bradley. "Integrated Graphics Solutions for Graphics-Intensive Applications" (https://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/Integrated_Graphics_Solutions_white_paper_rev61.pdf) (PDF). Archived (https://web.archive.org/web/20071128165723/https://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/Integrated_Graphics_Solutions_white_paper_rev61.pdf) (PDF) from the original on 2007-11-28. Retrieved 2007-09-02.

88. Sanford, Bradley. "Integrated Graphics Solutions for Graphics-Intensive Applications" (<https://www.techspot.com/news/46773-amd-announces-radeon-hd-7970-claims-fastest-gpu-title.html>). Archived (<https://web.archive.org/web/20120107200529/https://www.techspot.com/news/46773-amd-announces-radeon-hd-7970-claims-fastest-gpu-title.html>) from the original on 2012-01-07. Retrieved 2007-09-02.
89. Shimpi, Anand Lal. "AMD Outlines HSA Roadmap: Unified Memory for CPU/GPU in 2013, HSA GPUs in 2014" (<https://www.anandtech.com/show/5493/amd-outlines-hsa-roadmap-unified-memory-for-cpugpu-in-2013-hsa-gpus-in-2014>). www.anandtech.com. Retrieved 2024-01-08.
90. Lake, Adam T. "Getting the Most from OpenCL™ 1.2: How to Increase Performance by..." (<https://www.intel.com/content/www/us/en/developer/articles/training/getting-the-most-from-opencl-12-how-to-increase-performance-by-minimizing-buffer-copies-on-intel-processor-graphics.html>) Intel. Retrieved 2024-01-08.
91. Murph, Darren (29 September 2006). "Stanford University tailors Folding@home to GPUs" (<https://www.engadget.com/2006/09/29/stanford-university-tailors-folding-home-to-gpus/>). Archived (<https://web.archive.org/web/20071012000648/https://www.engadget.com/2006/09/29/stanford-university-tailors-folding-home-to-gpus/>) from the original on 2007-10-12. Retrieved 2007-10-04.
92. Houston, Mike. "Folding@Home – GPGPU" (<https://graphics.stanford.edu/~mhouston/>). Archived (<https://web.archive.org/web/20071027130116/https://graphics.stanford.edu/~mhouston/>) from the original on 2007-10-27. Retrieved 2007-10-04.
93. "Top500 List – June 2012 | TOP500 Supercomputer Sites" (<https://web.archive.org/web/20140113044747/https://www.top500.org/list/2012/06/100/>). Top500.org. Archived from the original (<https://www.top500.org/list/2012/06/100>) on 2014-01-13. Retrieved 2014-01-21.
94. Nickolls, John. "Stanford Lecture: Scalable Parallel Programming with CUDA on Manycore GPUs" (<https://www.youtube.com/watch?v=nIGnKPpOpbE>). YouTube. Archived (<https://web.archive.org/web/20161011195103/https://www.youtube.com/watch?v=nIGnKPpOpbE>) from the original on 2016-10-11.
- Harding, S.; Banzhaf, W. "Fast genetic programming on GPUs" (https://www.cs.bham.ac.uk/~wbl/biblio/gp-html/eurogp07_harding.html). Archived (https://web.archive.org/web/20080609231021/https://www.cs.bham.ac.uk/~wbl/biblio/gp-html/eurogp07_harding.html) from the original on 2008-06-09. Retrieved 2008-05-01.
95. Langdon, W.; Banzhaf, W. "A SIMD interpreter for Genetic Programming on GPU Graphics Cards" (https://www.cs.bham.ac.uk/~wbl/biblio/gp-html/langdon_2008_eurogp.html). Archived (https://web.archive.org/web/20080609231026/https://www.cs.bham.ac.uk/~wbl/biblio/gp-html/langdon_2008_eurogp.html) from the original on 2008-06-09. Retrieved 2008-05-01.
- V. Garcia and E. Debreuve and M. Barlaud. Fast k nearest neighbor search using GPU. In Proceedings of the CVPR Workshop on Computer Vision on GPU, Anchorage, Alaska, USA, June 2008.
96. "eGPU candidate system list" (<https://forum.techinferno.com/diy-e-gpu-projects/4109-egpu-candidate-system-list.html#post57511>). *Tech-Inferno Forums*. 15 July 2013.
97. Mohr, Neil. "How to make an external laptop graphics adaptor" (<https://www.techradar.com/news/computing-components/graphics-cards/how-to-make-an-external-laptop-graphics-adaptor-915616>). TechRadar. Archived (<https://web.archive.org/web/20170626052950/https://www.techradar.com/news/computing-components/graphics-cards/how-to-make-an-external-laptop-graphics-adaptor-915616>) from the original on 2017-06-26.
98. "Best External Graphics Card 2020 (EGPU) [The Complete Guide]" (<https://www.gamingscan.com/best-external-graphics-card/>). 16 March 2020.
99. "Use an external graphics processor with your Mac" (<https://support.apple.com/en-us/HT208544>). Apple Support. Retrieved 2018-12-11.
100. "OMEN Accelerator | HP® Official Site" (<https://www8.hp.com/us/en/campaigns/omenaccelerator/overview.html>). www8.hp.com. Retrieved 2018-12-11.

101. "Alienware Graphics Amplifier | Dell United States" (<https://www.dell.com/en-us/shop/alienware-graphics-amplifier/apd/452-bcfe/gaming>). Dell. Retrieved 2018-12-11.
- "Razer Core X – Thunderbolt™ 3 eGPU" (<https://www.razer.com/gaming-laptops/razer-core-x>). Razer. Retrieved 2018-12-11.
102. Box, ► Suggestions (2016-11-25). "Build Guides by users" (<https://egpu.io/build-guides/>). eGPU.io. Retrieved 2018-12-11.
103. "GPU Q3'22 biggest quarter-to-quarter drop since the 2009 recession" (<https://www.jonpeddie.com/news/q322-biggest-qtr-to-qtr-drop-since-the-2009-recession/>). Jon Peddie Research. 2022-11-20. Retrieved 2023-06-06.
104. "Graphics chips market is showing some life" (<https://www.tgdaily.com/enterprise/126591-graphics-chips-market-is-showing-some-life>). TG Daily. August 20, 2014. Archived (<https://web.archive.org/web/20140826115733/https://www.tgdaily.com/enterprise/126591-graphics-chips-market-is-showing-some-life>) from the original on August 26, 2014. Retrieved August 22, 2014.

Sources

- Peddie, Jon (1 January 2023). *The History of the GPU – New Developments* (<https://books.google.com/books?id=vfKkEAAAQBAJ>). Springer Nature. ISBN 978-3-03-114047-1. OCLC 1356877844 (<https://www.worldcat.org/oclc/1356877844>).

External links

- NVIDIA – What is GPU computing? (<https://www.nvidia.com/object/what-is-gpu-computing.html>)
 - The *GPU Gems* book series (<https://web.archive.org/web/20130605124431/https://developer.nvidia.com/content/gpu-gems>)
 - – A Graphics Hardware History (<https://titancity.com/articles/gfxcards.html>) Archived (<https://web.archive.org/web/20220331004621/http://titancity.com/articles/gfxcards.html>) 2022-03-31 at the Wayback Machine
 - [1] (<https://newstribuneworld.com/graphics-card-the-gpu-gpu-meaning>)
 - How GPUs work (https://web.archive.org/web/20161227132101/https://www.cs.virginia.edu/~gfx/papers/paper.php?paper_id=59)
 - GPU Caps Viewer – Video card information utility (https://www.ozone3d.net/gpu_caps_viewer/)
 - OpenGPU-GPU Architecture (In Chinese) (<https://www.opengpu.org/bbs/index.php>)
 - ARM Mali GPUs Overview (<https://web.archive.org/web/20130114054853/https://malideveloper.arm.com/learn-about-mali/about-mali/arm-mali-gpus/>)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Graphics_processing_unit&oldid=1202181477"

▪