

La empresa fue fundada el 5 de abril de 1993 por Jensen Huang (CEO a partir de 2022),[16] un ingeniero eléctrico taiwanés que anteriormente fue director de CoreWare en LSI Logic y diseñador de microprocesadores en AMD; Chris Malachowsky, ingeniero que trabajó en Sun Microsystems; y Curtis Priem, quien anteriormente fue ingeniero senior y diseñador de chips gráficos en IBM y Sun Microsystems.

En 1993, los tres cofundadores creían que la dirección correcta para la próxima ola de computación era la computación acelerada o basada en gráficos porque podía resolver problemas que la computación de propósito general no podía. También observaron que los videojuegos eran simultáneamente uno de los problemas más desafiantes desde el punto de vista computacional y tendrían un volumen de ventas increíblemente alto. Los videojuegos se convirtieron en el volante de inercia de la empresa para llegar a grandes mercados y financiar una enorme investigación y desarrollo para resolver problemas informáticos masivos. Con 40,000 dólares en el banco, nació la compañía.[19] Posteriormente, la compañía recibió 20 millones de dólares en fondos de capital de riesgo de Sequoia Capital y otros. NVIDIA inicialmente no tenía nombre y los cofundadores nombraron todos sus archivos NV, como "Next Version". La necesidad de incorporar la empresa llevó a los cofundadores a revisar todas las palabras con esas dos letras, llevándolos a "iNVIDIA", la palabra latina para "eNVIDIA".[19] NVIDIA se hizo pública el 22 de enero de 1999.

### Lanzamientos y adquisiciones

El lanzamiento de RIVA TNT en 1998 solidificó la reputación de NVIDIA en el desarrollo de adaptadores de gráficos competentes. A finales de 1999, NVIDIA lanzó la GeForce 256 (NV10), sobre todo introduciendo la transformación e iluminación (T&L) integradas en el hardware 3D de nivel de consumidor. Funcionando a 120 MHz y con canalizaciones de cuatro píxeles, implementó aceleración de video avanzada, compensación de movimiento y combinación alfa de subimagen de hardware. La GeForce superó a los productos existentes por un amplio margen.

Debido al éxito de sus productos, NVIDIA ganó el contrato para desarrollar el hardware de gráficos para la consola de juegos Xbox de Microsoft, lo que le valió a NVIDIA un anticipo de \$200 millones de dólares. Sin embargo, el proyecto alejó a muchos de sus mejores ingenieros de otros proyectos. A corto plazo, esto no importó, y la GeForce2 GTS se envió en el verano de 2000. En diciembre de 2000, NVIDIA llegó a un acuerdo para adquirir los activos intelectuales de su otrora rival 3dfx, un pionero en tecnología de gráficos 3D de consumo líder en el mercado, desde mediados de la década de 1990 hasta 2000. El proceso de adquisición finalizó en abril de 2002.

En julio de 2002, NVIDIA adquirió Exluna por una suma no revelada. Exluna creó herramientas de representación de software y el personal se fusionó con el proyecto Cg. En agosto de 2003, NVIDIA adquirió MediaQ por aproximadamente 70 millones de dólares. El 22 de abril de 2004, NVIDIA adquirió iReady, también proveedor de soluciones de descarga de TCP/IP e iSCSI de alto rendimiento.[28] En diciembre de 2004, se anunció que NVIDIA ayudaría a Sony con el diseño del procesador de gráficos (RSX) en la consola de juegos PlayStation 3. El 14 de diciembre de 2005, NVIDIA adquirió ULI Electronics, que en ese momento suministraba piezas de puente sur de terceros para chipsets a ATI, el competidor de NVIDIA.[29] En marzo de 2006, NVIDIA adquirió Hybrid Graphics.[30] En diciembre de 2006, NVIDIA, junto con su principal rival en la industria gráfica AMD (que había adquirido ATI), recibió citaciones del Departamento de Justicia de Estados Unidos con respecto a posibles violaciones antimonopolio en la industria de tarjetas gráficas.

Forbes nombró a NVIDIA su Compañía del Año para 2007, citando los logros que logró durante dicho período, así como durante los cinco años anteriores. El 5 de enero de 2007, NVIDIA anunció que había completado la adquisición de PortalPlayer, Inc. En febrero de 2008, NVIDIA adquirió Ageia, desarrollador del motor de física PhysX y la unidad de procesamiento de física. NVIDIA anunció que planeaba integrar la tecnología PhysX en sus futuros productos de GPU.[34][35]

En julio de 2008, NVIDIA realizó una amortización de aproximadamente \$ 200 millones en sus ingresos del primer trimestre, luego de informar que ciertos chipsets móviles y GPU producidos por la compañía tenían "tasas de falla anormales" debido a defectos de fabricación. NVIDIA, sin embargo, no reveló los productos afectados. En septiembre de 2008, NVIDIA se convirtió en objeto de una demanda colectiva por los defectos, alegando que las GPU defectuosas se habían incorporado a ciertos modelos de portátiles fabricados por Apple Inc., Dell y HP. En septiembre de 2010, NVIDIA llegó a un acuerdo en el que reembolsaría a los propietarios de las computadoras portátiles afectadas las reparaciones o, en algunos casos, el reemplazo.[36][37] El 10 de enero de 2011, NVIDIA firmó un acuerdo de licencia cruzada de \$ 1.5 mil millones por seis años con Intel, poniendo fin a todos los litigios entre las dos compañías.

En noviembre de 2011, después de presentarlo inicialmente en el Mobile World Congress, NVIDIA lanzó su sistema Tegra 3 ARM en un chip para dispositivos móviles. NVIDIA afirmó que el chip presentaba la primera CPU móvil de cuatro núcleos. En mayo de 2011, se anunció que NVIDIA había acordado adquirir Icera, una empresa de fabricación de chips de banda base en el Reino Unido, por 367 millones de dólares. En enero de 2013, NVIDIA presentó Tegra 4, así como NVIDIA Shield, una consola de juegos portátil basada en Android que funciona con el nuevo sistema en chip. El 29 de julio de 2013, NVIDIA anunció que adquirió PGI de STMicroelectronics.

En 2014, NVIDIA transfirió los juegos de Valve Portal y Half Life 2 a su tableta NVIDIA Shield como Lightspeed Studio. Desde 2014, NVIDIA ha diversificado su negocio centrándose en tres mercados: juegos, electrónica automotriz y dispositivos móviles.

El 6 de mayo de 2016, NVIDIA presentó las primeras GPU de la serie GeForce 10, las GTX 1080 y 1070, basadas en la nueva microarquitectura Pascal de la empresa. NVIDIA afirmó que ambos modelos superaron a su modelo Titan X basado en Maxwell; los modelos incorporan memoria GDDR5X y GDDR5 respectivamente, y utilizan un proceso de fabricación de 16 nm. La arquitectura también es compatible con una nueva función de hardware conocida como multiproyección simultánea (SMP), que está diseñada para mejorar la calidad del renderizado de realidad virtual y multimonitor.[46][47][48] Las computadoras portátiles que incluyen estas GPU y son lo suficientemente delgadas (a fines de 2017, menos de 0,8 pulgadas (20 mm)) han sido designadas como que cumplen con el estándar de diseño "Max-Q" de NVIDIA .

En julio de 2016, NVIDIA llegó a un acuerdo por una demanda por publicidad falsa con respecto a su modelo GTX 970, ya que los modelos no podían usar todos los 4 GB de RAM anunciados debido a las limitaciones del diseño de su hardware.[50] En mayo de 2017, NVIDIA anunció una asociación con Toyota que utilizará la plataforma de inteligencia artificial de la serie Drive PX de NVIDIA para sus vehículos autónomos.[51] En julio de 2017, NVIDIA y el gigante de búsqueda chino Baidu anunciaron una asociación de inteligencia artificial de gran alcance que incluye computación en la nube, conducción autónoma, dispositivos de consumo y el marco de trabajo de inteligencia artificial de código abierto de Baidu, PaddlePaddle. Baidu reveló que Drive PX 2 AI de NVIDIA será la base de su plataforma de vehículos autónomos.

NVIDIA lanzó oficialmente Titan V el 7 de diciembre de 2017.

NVIDIA lanzó oficialmente la NVIDIA Quadro GV100 el 27 de marzo de 2018. NVIDIA lanzó oficialmente las GPU RTX 2080 el 27 de septiembre de 2018. En 2018, Google anunció que las tarjetas gráficas Tesla P4 de NVIDIA se integrarían en la inteligencia artificial del servicio Google Cloud.

En mayo de 2018, en el foro de usuarios de NVIDIA, se inició un hilo[57] solicitando a la empresa que actualizara a los usuarios cuando lanzarían controladores web para sus tarjetas instaladas en máquinas Mac Pro heredadas hasta mediados de 2012 5,1 con macOS Mojave sistema operativo 10.14. Se requieren controladores web para habilitar la aceleración de gráficos y las capacidades de múltiples monitores de pantalla de la GPU. En su sitio web de información de actualización de Mojave, Apple declaró que macOS Mojave se ejecutaría en máquinas heredadas con tarjetas gráficas 'compatibles con Metal'[58] y enumeraban GPU compatibles con Metal, incluidas algunas fabricadas por NVIDIA. Sin embargo, esta lista no

incluía tarjetas compatibles con Metal que actualmente funcionan en macOS High Sierra usando controladores web desarrollados por NVIDIA. En septiembre, NVIDIA respondió: "Apple controla completamente los controladores para Mac OS. Pero si Apple lo permite, nuestros ingenieros están listos y ansiosos por ayudar a Apple a ofrecer excelentes controladores para Mac OS 10.14 (Mojave)".[60] En octubre, NVIDIA siguió este con otro anuncio público, "Apple controla completamente los controladores para Mac OS. Desafortunadamente, NVIDIA actualmente no puede lanzar un controlador a menos que sea aprobado por Apple".[61] lo que sugiere una posible ruptura entre las dos compañías.[62] En enero de 2019, sin señales de que los controladores web fueran habilitados, Apple Insider intervino en la controversia con la afirmación de que la administración de Apple "no quiere compatibilidad con NVIDIA en macOS".[63] El mes siguiente, Apple Insider continuó con otra afirmación de que el soporte de NVIDIA se abandonó debido a "problemas relacionales en el pasado",[64] y que Apple estaba desarrollando su propia tecnología de GPU.[65] Sin los controladores web de NVIDIA aprobados por Apple, los usuarios de Apple deben reemplazar sus tarjetas NVIDIA con una marca compatible de la competencia, como AMD Radeon de la lista recomendada por Apple.

El 11 de marzo de 2019, NVIDIA anunció un acuerdo para comprar Mellanox Technologies por \$6900 millones[67] para expandir sustancialmente su presencia en el mercado de la computación de alto rendimiento. En mayo de 2019, NVIDIA anunció nuevas computadoras portátiles RTX Studio. Los creadores dicen que la nueva computadora portátil será siete veces más rápida que una MacBook Pro de gama alta con un Core i9 y gráficos Radeon Pro Vega 20 de AMD en aplicaciones como Maya y RedCine-X Pro.[68] En agosto de 2019, NVIDIA anunció Minecraft RTX, un parche oficial desarrollado por NVIDIA para el juego Minecraft que agrega el trazado de rayos DXR en tiempo real exclusivamente a la versión de Windows 10 del juego. Todo el juego es, en palabras de NVIDIA, "reacondicionado" con trazado de ruta, lo que afecta drásticamente la forma en que funcionan la luz, los reflejos y las sombras dentro del motor.

En mayo de 2020, los principales científicos de NVIDIA desarrollaron un ventilador de código abierto para abordar la escasez resultante de la pandemia mundial de coronavirus.[70] El 14 de mayo de 2020, NVIDIA anunció oficialmente su microarquitectura de GPU Ampere y el acelerador de GPU NVIDIA A100.[71][72] En julio de 2020, se informó que NVIDIA estaba en conversaciones con SoftBank para comprar Arm, un diseñador de chips con sede en el Reino Unido, por \$32 mil millones.

El 1 de septiembre de 2020, NVIDIA anunció oficialmente la serie GeForce 30 basada en la nueva microarquitectura Ampere de la empresa.

El 13 de septiembre de 2020, se anunció que NVIDIA compraría Arm a SoftBank Group por \$ 40 mil millones, sujeto al escrutinio habitual, y este último retendría una participación del 10% en NVIDIA. En octubre de 2020, NVIDIA anunció su plan para

construir la computadora más poderosa en Cambridge, Inglaterra. Nombrada Cambridge-1, la computadora empleará IA para respaldar la investigación de atención médica, y se espera que finalice a fines de 2020, a un costo de aproximadamente £ 40 millones. Según Jensen Huang, "La supercomputadora Cambridge-1 servirá como un centro de innovación para el Reino Unido y promoverá el trabajo innovador que realizan los investigadores de la nación en el descubrimiento de fármacos y atención médica crítica".

También en octubre de 2020, junto con el lanzamiento de NVIDIA RTX A6000, NVIDIA anunció que retirará su marca de GPU para estaciones de trabajo Quadro, cambiando el nombre del producto a NVIDIA RTX para futuros productos y la fabricación se basará en la arquitectura NVIDIA Ampere.

En agosto de 2021, la adquisición propuesta de Arm Holdings se estancó después de que la Autoridad de Mercados y Competencia del Reino Unido planteara "preocupaciones importantes sobre la competencia".[78] En octubre de 2021, la Comisión Europea abrió una investigación de competencia sobre la adquisición. La Comisión declaró que la adquisición de NVIDIA podría restringir el acceso de los competidores a los productos de Arm y proporcionar a NVIDIA demasiada información interna sobre sus competidores debido a sus acuerdos con Arm. SoftBank (la empresa matriz de Arm) y NVIDIA anunciaron a principios de febrero de 2022 que "habían acordado no seguir adelante con la transacción 'debido a importantes desafíos regulatorios'". Está previsto que la investigación finalice el 15 de marzo de 2022.[80][81] Ese mismo mes, NVIDIA supuestamente se vio comprometida por un ciberataque. El ataque coincidió con la invasión rusa de Ucrania en 2022, aunque no hay indicios de que el ataque proviniera de Rusia o de piratas informáticos rusos.

En marzo de 2022, el director ejecutivo de NVIDIA, Jensen Huang, mencionó que están abiertos a que Intel fabrique sus chips en el futuro. Esta fue la primera vez que la empresa mencionó que trabajaría junto con los próximos servicios de fundición de Intel.

En abril de 2022, se informó que NVIDIA planeaba abrir un nuevo centro de investigación en Ereván, Armenia.

En septiembre de 2022, NVIDIA anunció su chip de grado automotriz de última generación, Drive Thor.

Siguiendo las regulaciones del Departamento de Comercio de Estados Unidos que impusieron un embargo a las exportaciones a China de microchips avanzados, que entraron en vigencia en octubre de 2022, NVIDIA vio cómo se agregó su chip de centro de datos a la lista de control de exportaciones. El mes siguiente, la empresa presentó un nuevo chip avanzado en China, denominado GPU A800, que cumplía con las normas de control de exportaciones.

En septiembre de 2023, Getty Images anunció que se asociaría con NVIDIA para lanzar Generative AI by Getty Images, una nueva herramienta que permite a las personas crear imágenes utilizando la biblioteca de fotografías con licencia de Getty. Getty utilizará el modelo Edify de NVIDIA, que está disponible en la biblioteca de modelos de IA generativa Picasso de NVIDIA.

En octubre de 2023, se informó que NVIDIA había comenzado silenciosamente a diseñar unidades de procesamiento central (CPU) basadas en ARM para el sistema operativo Windows de Microsoft con el objetivo de comenzar a venderlas en 2025.