

La Tarjeta Grafica

Qué es una tarjeta gráfica

Una tarjeta gráfica es un componente de un ordenador que puede estar integrado dentro de la placa base, el procesador o ser externo y cuya función es la renderizar las imágenes en la pantalla y ofrecer una visualización de alta calidad, procesando y ejecutando datos gráficos mediante técnicas, características y funciones gráficas avanzadas.

Qué son las tarjetas gráficas de ordenador

¿Y quienes necesitan estas tarjetas gráficas?. En primer lugar, los usuarios que utilizan el ordenador para jugar videojuegos. Luego, los que usan aplicaciones de edición de vídeo como Adobe Premiere o DaVinci Resolve, como así para editar fotografías, como puede ser Adobe Photoshop. Por último, los usuarios que trabajan con complejos programas informáticos de inteligencia artificial. Para realizar cualquiera de estas tres actividades, una buena tarjeta gráfica es imprescindible.

Diferencia entre una tarjeta gráfica y una GPU

Hay que diferenciar la GPU (unidad de procesamiento gráfico) de lo que es una tarjeta gráfica, porque lo primero es un chip, mientras que lo segundo es un producto terminado.

Las siglas GPU se refieren a Graphics Processing Unit o unidad de procesamiento gráficos y se trata de un chip encargado para acelerar la representación de gráficos 3D. Sin embargo, este concepto ha evolucionado mucho debido al avance tecnológico en desarrollo y en arquitecturas. La GPU realiza multitud de operaciones relacionadas con los datos 2D y 3D, como es decodificar y renderizar objetos animados y vídeos. El grado de sofisticación con el que cuenta una GPU hoy en día es brutal, agregando más y más tareas.



Diferencias entre tarjeta gráfica y GPU

Antes, la GPU no era más que un chip que se encargaba de acelerar gráficos 3D; en pleno siglo XXI, la GPU renderiza, codifica, calcula el rebote de los rayos del Ray Tracing, optimiza el rendimiento con tecnologías de reescalado y algunas hacen uso de Machine Learning e Inteligencia Artificial.

Ahora que tenemos claramente identificado qué es una GPU, vamos a explicaros qué es una tarjeta gráfica como producto completo compuesto por:

- PCB, que es donde van soldados todos los componentes.
- Módulos de memoria (GRAM).
- Condensadores y demás componentes.
- Tuberías de calor.
- Ventiladores (o blower) y carcasa.

La GPU es un micro-componente muy complejo con diferentes núcleos encargados de procesar la información. Y son precisamente la cantidad y la capacidad de estos lo que determinará la potencia de la tarjeta gráfica. Al contrario que con la CPU (procesador) la GPU tiene muchos núcleos con frecuencias bajas encargado cada uno de procesar un número muy grande de píxeles.

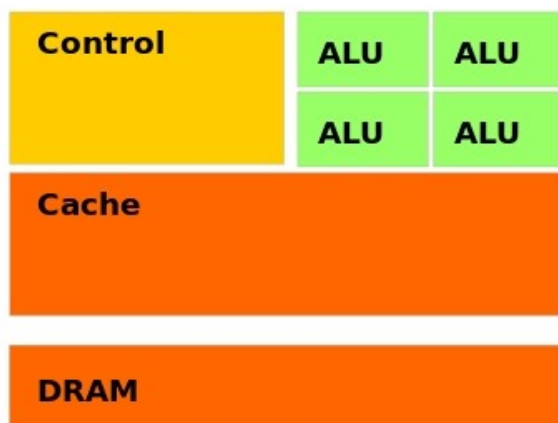
Otro de los componentes que te habíamos mencionado antes es la GRAM (memoria gráfica de acceso aleatorio) y su función dentro de una tarjeta gráfica es la de almacenar y transportar información. Esta información pasa el RAMDAC, que viene a ser conversor de señal digital a señal analógica necesario para que el monitor de nuestro ordenador pueda interpretar la información que se le está enviando.

En otras palabras, la tarjeta gráfica hace dos cosas: procesa los vértices, su ordenamiento espacial, su rotación, y qué segmentos de estos serán visibles gráficamente para transmitir luego esta información al monitor en un formato que pueda representar.

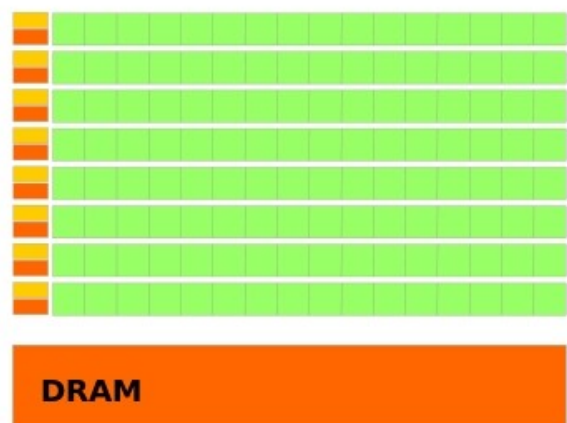
Arquitectura de la GPU

Existen varias compañías que diseñan GPUs, como Intel, AMD y NVIDIA para el ámbito del PC y el HPC (High Performance Computing), así como también han diseñado algunas para el sector de los dispositivos móviles, e incluso otras como Imagination Technologies (PowerVR Technologies) para ARM, Apple Silicon, Qualcomm (sus Adreno surgieron de una compra de la división móvil de ATI/AMD), etc.

Todos ellos tienen arquitecturas muy diferentes en cada GPU. Por este motivo, no se pueden comparar entre ellas a la ligera, ni tampoco los FLOPS (Floating point operations per second, que significa “operaciones de coma flotante por segundo”) se deben usar como unidad de referencia para comparativas. Y esto es debido a la heterogeneidad de unidades de procesamiento que existe, dando resultados muy diferentes.



CPU



GPU

Las unidades computacionales o núcleos de una GPU son mucho más pequeñas y simples que las de una CPU, sin embargo, al ser tan numerosas, la potencia de cálculo es bastante impresionante, especialmente en coma flotante. De hecho, el núcleo de una GPU lo que más hace es sumar y multiplicar (MAD), y multiplicar-sumar (FMA). Otros núcleos algo más complicados, como las unidades tensoriales o Tensor units, pueden hacer tareas algo más complejas que eso, como las operaciones tensor para acelerar IA (Inteligencia Artificial). Otros núcleos complejos son los de trazado de rayos, destinados a aportar mayor realismo en la iluminación.

las operaciones matemáticas que realizan, sumar y multiplicar (MAD) es una operación fundamental, mientras que sumar-multiplicar (FMA) es una versión optimizada que combina estas dos operaciones en una sola. FMA es más eficiente y se utiliza extensivamente en procesamiento de gráficos porque permite realizar más operaciones por unidad de tiempo.

Explicación detallada:

- **MAD (Multiplicar y Sumar):** Esta operación consiste en multiplicar dos números y luego sumar el resultado a un tercer número. Es una operación básica que se utiliza en muchas aplicaciones, incluyendo el procesamiento de gráficos.
- **FMA (Multiplicar-Sumar Fusionada):** En FMA, la multiplicación y la suma se realizan en una sola instrucción, con un solo redondeo. Esto permite que las operaciones se realicen de manera más eficiente, lo que se traduce en un mayor rendimiento en aplicaciones que requieren un alto nivel de cálculo, como los juegos y el procesamiento de imágenes.
- **Ejemplo:** Considera la operación $(a * b) + c$. En MAD, la multiplicación $a * b$ se realiza, el resultado se redondea y luego se suma c . En FMA, la operación se realiza en un solo paso, manteniendo la precisión máxima antes de redondear el resultado final.

Tipos de tarjetas gráficas

Por regla general, las mejores tarjetas AMD ofrecen un rendimiento similar al de las mejores tarjetas Nvidia. Por lo que será más cuestión de precios, de stock y de gustos personales.

Pero eso no es todo y es que, en función del fabricante de la tarjeta gráfica, podemos encontrar diferentes modelos en función a dos características principales:

1. Por refrigeración
 1. Activa: hacen uso de ventiladores.
 2. Pasiva: hacen uso de aletas de aluminio para expulsar el calor por el método de convección.

3. Líquida: utilizan un bloque de agua que lleva conectado 2 tubos dirigidos a un radiador con uno o varios ventiladores para expulsar el calor fuera de la caja.
2. Por tamaño

Qué marcas de tarjetas gráficas existen

A pesar de que la mitad de las tarjetas gráficas sean de Nvidia y la otra mitad sean de AMD, hay una gran cantidad de empresas que se dedican a fabricar y a comercializar dichas tarjetas gráficas. Estas son las mas conocidas:

- Asus
- MSI
- Gigabyte
- EVGA
- Zotac
- Sapphire
- PowerColor
- PNY
- Inno3D
- XFX
- Gainward
- Asrock

Qué modelos de tarjetas son los más vendidos

A estas alturas seguramente tengas ya ganas de lanzarte a mirar tarjetas gráficas para comprar alguna. Para que no te hagas problemas buscando y buscando entre los cientos de modelos, vamos a hacerte algunas recomendaciones. Desde modelos de gama alta para gamers apasionados o otros con mejor relación calidad-precio que igual sirven para jugar y también para tareas de diseño gráfico y programación.

Componentes Claves

GPU : La principal pieza de una placa de vídeo es la GPU, la unidad de procesamiento gráfico. Se refiere a un procesador dedicado, justamente, al procesamiento de los gráficos. Lo que hace es aligerar la carga de trabajo del procesador central, acción que la convierte en la parte más importante de toda tarjeta gráfica y es la que determina el rendimiento de la misma.

GPC : (Graphics Processing Cluster) se refiere a un grupo de unidades de procesamiento gráfico dentro de una GPU. Un GPC contiene los elementos clave para

el procesamiento de gráficos, como motores de renderizado y multiprocesadores de transmisión.

GRAM: Luego, encontramos la GRAM, la memoria gráfica de acceso aleatorio, que son chips de memoria que almacenan y transportan información. En este sentido, identificamos dos tipos de memorias gráficas. Por una parte, la dedicada, memoria más eficiente cuando la tarjeta o la GPU disponen exclusivamente para sí este tipo de memorias. Por otro lado, está la compartida, utilizada cuando se emplea memoria en detrimento de la memoria de acceso aleatorio.

A su vez, la potencia de la memoria gráfica es el resultado de tres aspectos: capacidad, interfaz de memoria y frecuencia de memoria.

1. La capacidad determina el número máximo de datos y texturas que puede procesar. Es un detalle muy relevante a tener en cuenta para resoluciones superiores a 1440p y monitores múltiples, donde cada imagen toma mucho más espacio.
2. La interfaz de memoria, o bus de datos, resulta de la multiplicación del ancho de bits de cada chip por su número de unidades. Es un detalle clave para determinar el ancho de banda, es decir, la cantidad de datos que puede transferir en cierto tiempo.
3. Por último, la frecuencia de memoria es, valga la redundancia, la frecuencia a la que la memoria puede transportar los datos ya procesados. Es así que se complementa con la interfaz y, entre ambas características, se determina el ancho de banda de la memoria.

TPC: Clúster de procesamiento de texturas

SM: (Streaming Multiprocessor) se refiere a una unidad de procesamiento que se encuentra dentro de la tarjeta gráfica.

Tex unit: Unidad de textura

Tex cache: Caché de textura

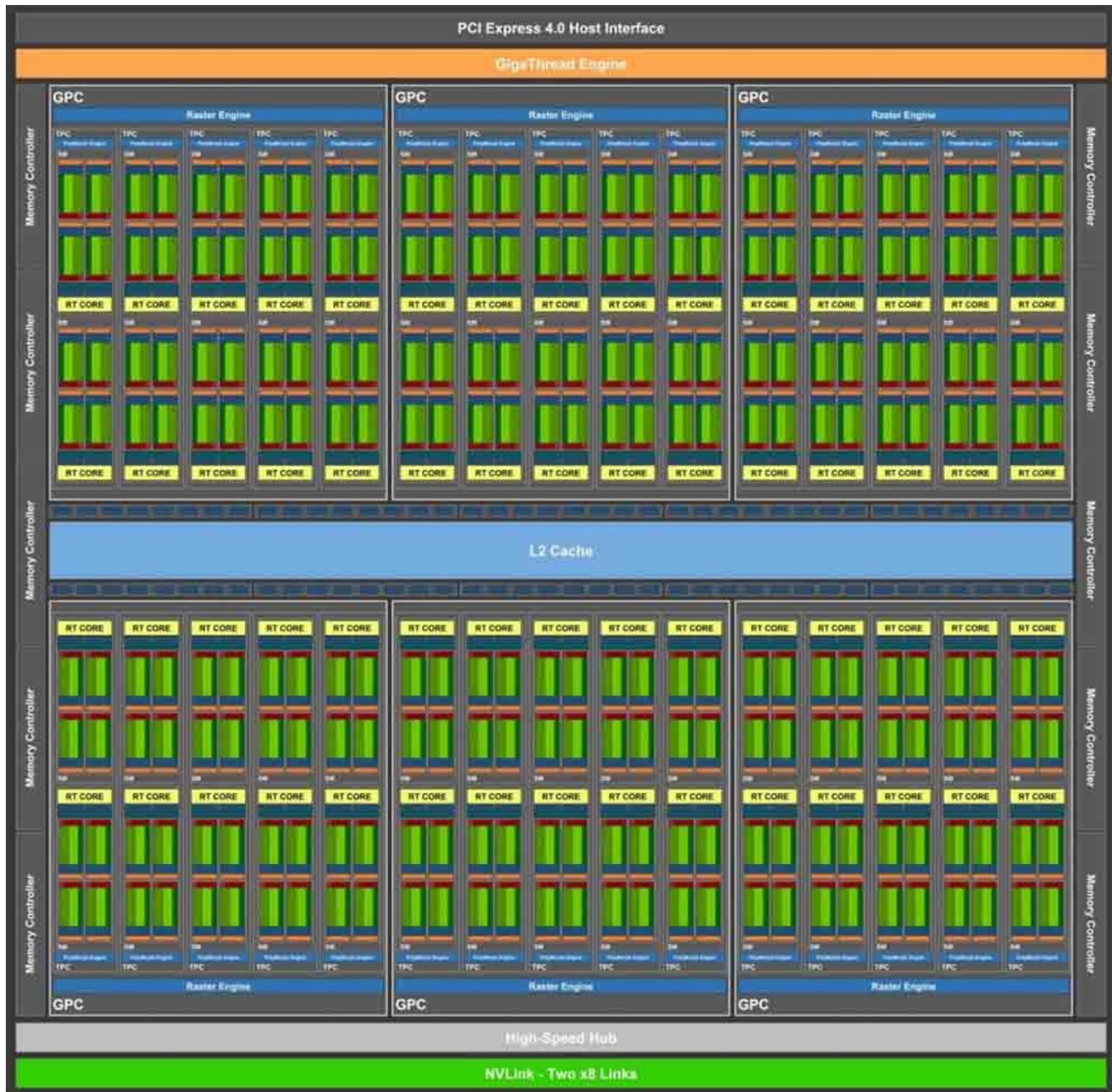
SP: Procesador escalar.

Las texturas son imágenes digitales que se aplican a modelos 3D para añadir detalle y realismo. El proceso de aplicar texturas se conoce como "mapeo de texturas" y se utiliza en gráficos 3D para hacer que los objetos parezcan más realistas. Las texturas se almacenan en la memoria de la GPU y se utilizan para mejorar el aspecto de los modelos poligonales.

El Proceso de Mapeo de Texturas:

1. Creación de Texturas: Las texturas pueden ser creadas digitalmente o provenir de imágenes existentes.
2. Almacenamiento en Memoria: Las texturas se almacenan en la memoria de la GPU, que es una memoria de acceso rápido diseñada para este tipo de datos.
3. Aplicación a Modelos 3D: Las texturas se aplican a los modelos 3D mediante el mapeo de texturas, donde se asignan las coordenadas de la textura a las coordenadas del modelo.

- <https://listado.mercadolibre.com.co/computacion/componentes-pc/tarjetas/tarjetas-video/>



NVIDIA Arquitectura

Qué características tiene una tarjeta gráfica

Vamos a hablar de cuestiones más técnicas que guardan relación con el rendimiento y las características que tendrás que tener en cuenta a la hora de comprar una tarjeta gráfica.

Arquitectura

En toda presentación de AMD o NVIDIA, se hacen esfuerzos en enseñar el diseño de la arquitectura que van a usar todas las tarjetas gráficas de una generación. Es tan importante este diseño en el papel de la GPU que NVIDIA diferencia sus generaciones a través del nombre de la arquitectura (Maxwell, Pascal, Turing, Ampere, Ada Lovelace...).

AMD empezó a hacer lo mismo en 2018 con el nacimiento de RDNA, y posteriormente en 2020 con RDNA 2.

Arquitectura tarjeta gráfica

De hecho, sigue habiendo el mismo pensamiento en AMD, a pesar de tener que renunciar a Machine Learning y la Inteligencia Artificial para optimizar los cálculos.

Intel anunció que llegarían las Arc Alchemist, tarjetas gráficas de escritorio y portátil, para el año 2022. Ha enseñado varios detalles de la arquitectura, apareciendo en escena los Xe Cores, y parece que sería un complejo compuesto por motores Vector y Matrix, así como los famosos XMX Cores, que serían los encargados del Intel XeSS (alternativa a FSR o DLSS).

¿Son muy importantes? Sí cuando comparamos las GPUs de una misma marca y generación, lo cual nos advierte del pronóstico de rendimiento.

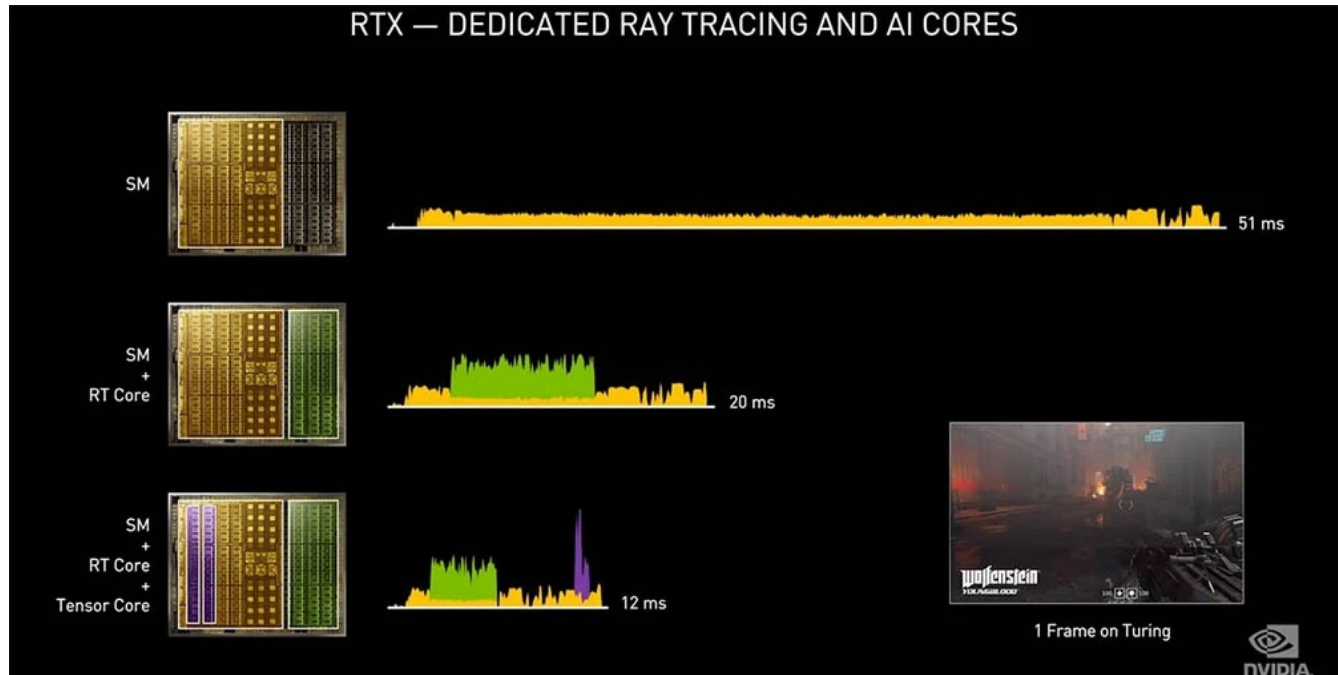
Frecuencia: base, game y boost

En casi todas las tarjetas gráficas encontraremos 3 frecuencias distintas:

1. **Base**, que es el reloj a la que funciona la GPU en reposo.
2. **Game**, el reloj cuando jugamos o seleccionamos un preset de overclock determinado.
3. **Boost**, la frecuencia máxima teórica que podemos disfrutar en la GPU.

La frecuencia GPU se mide en MHz e indica lo rápido que los núcleos de una GPU pueden ser. A mayor refrigeración, mayor capacidad de tener una frecuencia más

alta (más calor, requiere mejor refrigeración), por ello las tarjetas gráficas con 3 ventiladores o con bloques de agua manejan frecuencias más altas.



La velocidad del reloj es el número de veces que la GPU oscila por segundo. Por ejemplo, si tiene 2000 MHz, significa que se cicla 2 millones de veces por segundo.

Hay que diferenciar el Core Clock del Memory clock, siendo la segunda la velocidad de la memoria VRAM o memoria de vídeo. Cuando decimos VRAM nos referimos a la cantidad de memoria GDDR que tiene la GPU, siendo muy habitual encontrar 4 GB, 6 GB, 8 GB, 10-12 GB, 16 GB y 24 o más GB.

Tanto el Core clock, como el Memory clock se pueden ajustar en MSI Afterburner para hacer overclock, pero aconsejo revisar las coberturas de garantía del producto antes de modificar. Esto te va ayudar para descartar modelos y saber qué tarjeta gráfica comprar en caso de querer overclock.

Ancho de banda y velocidad de memoria

Diferente al Memory clock, tenemos la velocidad de memoria, que es medida en Gbps (Gigabits por segundo). Esto significa la velocidad de interfaz de la línea de datos del chip de memoria, que suele ser de 14 o 16 Gbps, aunque AMD lanzó las RX 6X50 con 18 Gbps.

Este valor es parte de la especificación de la memoria GDDR de turno (GDDR5, GDDR6, GDDR6X, etc.). Por este motivo, la tarjetas graficas RTX 3090 Ti lleva memoria GDDR6X y su velocidad de memoria es de 21 Gbps.

Por otro lado, está el ancho de banda de la memoria, que es medido en GB/segundo (o GB/s), diferenciado del ancho de banda de la interfaz (medido en bit). Con estos datos no vas a saber si una GPU es más potente que otra, pero son muy útiles para saber la potencia teórica del modelo.

Conectores, alimentación y salidas de vídeo

Estas características sí que suponen ciertas diferencias entre tarjetas gráficas de cara al uso de las mismas y a su aprovechamiento. Primero, vamos a diferenciar conectores y salidas:

En una tarjeta gráfica te vas a encontrar:

- Conectores PCIe para alimentarla de energía directamente desde la fuente de alimentación. Varían según modelos y marcas, encontrando 6+4, 8+2 e, incluso un conector entero de 12 pines, ¡consulta el modelo antes de comprar!
- Unos pines que deben aterrizar en la ranura PCI-Express.
- Posiblemente un switch o interruptor para elegir un perfil de rendimiento o silencioso (ASUS hace mucho esto).
- Salidas de vídeo:
 - HDMI 2.1.
 - DisplayPort 1.4.
 - USB, en el caso de las más potentes para realidad virtual.



<https://www.pccomponentes.com/como-elegir-tarjeta-grafica>
<https://www.profesionalreview.com/2022/05/22/arquitectura-gpu/>
<https://www.xataka.com/basics/que-son-los-teraflops-y-que-miden-exactamente>