



Draaiboek explainervideo
Jort & Isabel

Introductie onderwerp: NVIDIA

Onze explainervideo gaat over NVIDIA: een toonaangevende producent van computerhardware, gespecialiseerd in videokaarten en chipsets voor zogenaamde 'motherboards'

NVIDIA's technologieën en producten hebben de manier waarop we digitale inhoud creëren, consumeren en verwerken aanzienlijk veranderd

Centrale vraag:

Wat is NVIDIA en hoe heeft dit bedrijf invloed op jou als consument?

Onderdelen en scenes

Introductie door middel van gamingvoorbeeld

- Scene 1: montage 2D games (duidelijk voorbeeld: 'Doom')
- Scene 2: Introductie GeForce 256
- Scene 3: 's werelds eerste GPU (cubemap)

Wat is een GPU?

- Scene 4: het visualiseren van een 3D-beeld (a.d.h.v. 4 stages: 60 triangles, 600, 6000, etc)
- Scene 5: parallelle verwerking
- Scene 6: visuele voorbeelden 3D-rendering, videobewerking, zelfrijdende autos en wetenschappelijke simulaties

NVIDIA en de evolutie van GPU-technologie

- Scene 7: over het bedrijf zelf
- Scene 8:
- Scene 9: wat maakt NVIDIA's chip zo bijzonder?
- Scene 10: training
- Scene 11: inference
- Scene 12:
- Scene 13:

De toekomst van NVIDIA

- Scene 14:
- Scene 15:

Introductie door middel van gamingvoorbeeld

SCENE 1



Beeld: fade-in van verschillende 2D games

Inclusief achtergrondgeluid

<https://www.youtube.com/watch?v=MnqLJpgg7jc> (stukje gameplay)

Voice Over: geen

SCENE 1



Beeld: fade-in vergelijking 'Doom' uit 1993 en uit 2020 (<https://www.youtube.com/watch?v=4GbmBVTBKXw>)
'Doom' kwam uit in 1993 en was een van de eerste games die in 3D werden uitgebracht

Voice Over: wat heeft ervoor gezorgd dat dit spel door de jaren heen op deze manier is ontwikkeld? Heeft te maken met een bepaald soort chip: de GPU. Voorheen waren spellen alleen maar 2D, zoals Pacman en de eerste versies van Sonic the Hedgehog

'Doom' introduceerde 3D gaming aan de wereld, hierna heeft deze industrie snel tempo ontwikkeld. Tegenwoordig is 3D de standaard: op dit moment zien we vooral een sterke ontwikkeling in de wereld van AI.

'Doom' kon 3D worden door de introductie van de GPU chip: en speciaal type microprocessor dat is ontworpen om complexe grafische berekeningen uit te voeren. NVIDIA is de belangrijkste producent van deze chips.

SCENE 2

Beeld: GeForce 256

Klein in beeld:

- Fab by: TSMC.
- Transistors in processor: 17 million
- Year: 1999
- VRAM: 32 MB
- AGP 4x interface
- 120 Mhz Core



Voice Over:

Dit is de GeForce 256, een zeer speciale chip die in 1999 door NVIDIA werd gemaakt voor grafische computers

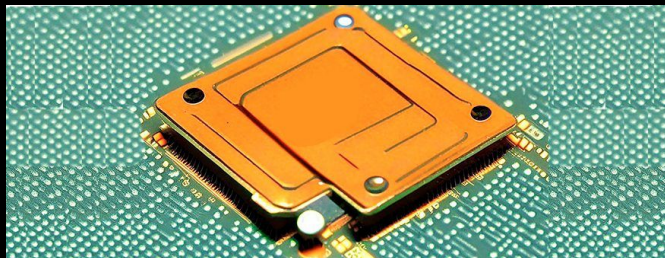
SCENE 3

Voice Over:

Staat nu bekend als 's wereld eerste GPU: Graphics Processing Unit: microprocessor die is ontworpen om complexe berekeningen uit te voeren die verband houden met grafische weergave

De belangrijkste functie van een GPU is het versnellen van de weergave van visuele inhoud, zoals 2D- en 3D-afbeeldingen, video's en animaties. Dit omvat taken zoals het renderen van polygonen, het toepassen van texturen, het uitvoeren van belichtingsberekeningen en het uitvoeren van post-processing effecten.

In de loop der jaren zijn GPU's steeds krachtiger geworden en hebben ze hun toepassingsgebied uitgebreid naar gebieden buiten grafische weergave. Tegenwoordig worden GPU's veel gebruikt in toepassingen zoals machine learning, data-analyse, wetenschappelijke simulaties en cryptocurrency mining vanwege hun vermogen om enorme aantallen berekeningen parallel uit te voeren.



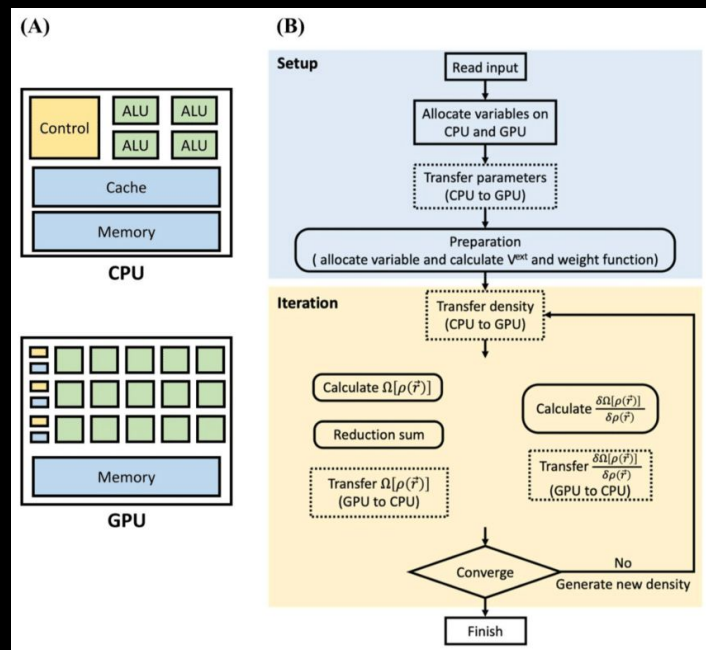
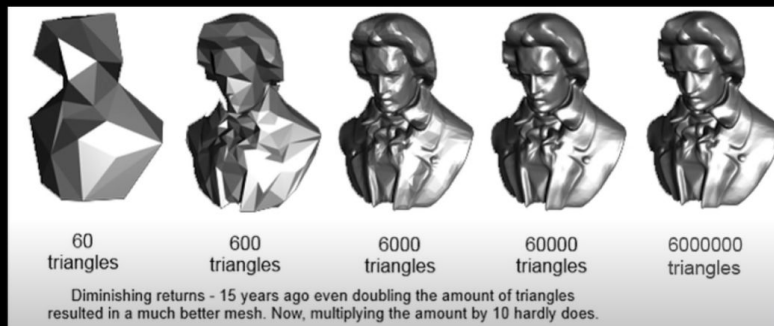
Wat is een GPU?

SCENE 4

Voice Over:

Het visualiseren van een 3D-beeld

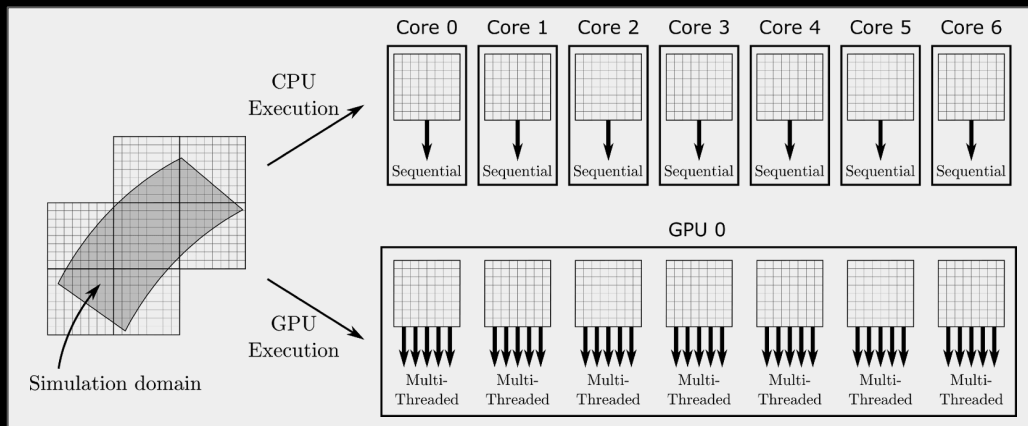
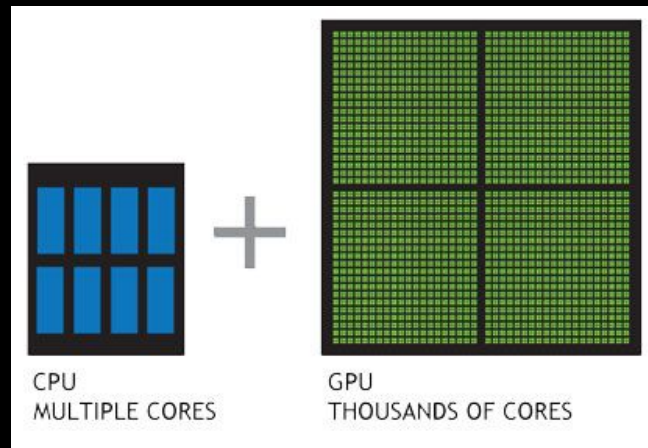
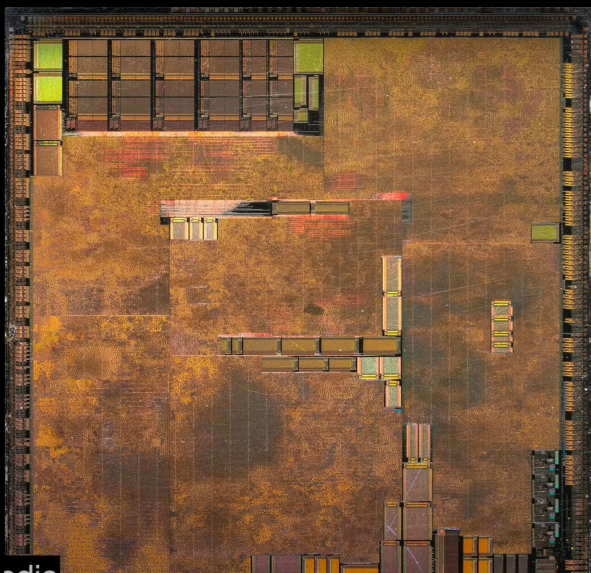
Beeld: object zien renderen aan de hand van driehoeken (zoals hieronder geïllustreerd)



SCENE 5

Voice Over:

Parallele verwerking



SCENE 6

Voice Over:

Dit betekent dat een GPU niet alleen voor games essentieel is maar ook voor andere taken die veel parallele berekeningen vereisen, zoals 3D-rendering, videobewerking, zelfrijdende autos en wetenschappelijke simulaties

Beeld:

Visuele voorbeelden 3D-rendering, videobewerking, zelfrijdende autos en wetenschappelijke simulaties



NVIDIA en de evolutie van GPU-technologie

SCENE 7

Voice Over:

Dit heeft het bedrijf NVIDIA uit Santa Clara, California als beste doorged. Door hun focus te verleggen van gamers naar een nieuwe industrie. NVIDIA: het ontwerpen, ontwikkelen en verkopen van grafische verwerkingseenheden (GPU's) en aanverwante technologieën.

Game bedrijf, maar nu AI bedrijf geworden

Large Language Models... Ookwel AI

In deze "AI goldrush" heeft NVIDIA de lucratieve positie als enige de kaarten (pickhouwelen) te verkopen die voor de grote AI bedrijven onmisbaar zijn.

Beeld:

Marktwaaarde (animatie)



SCENE 8

Voice Over:

Dit kwam niet uit de lucht vallen

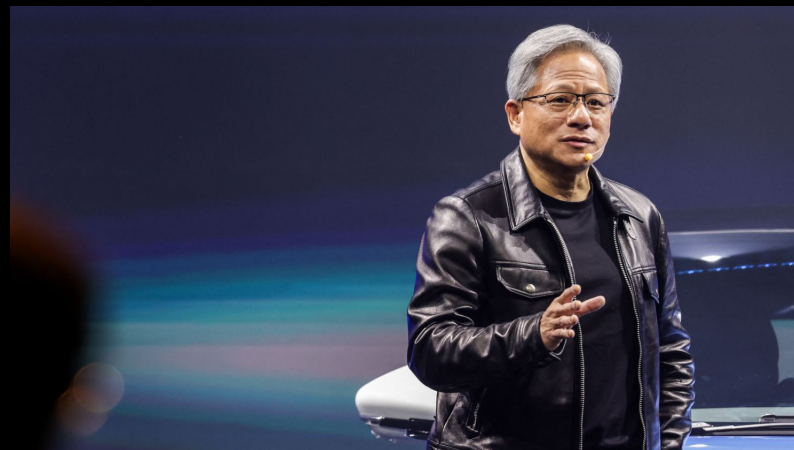
Het bedrijf onder leiding van Jenssen Huang (CEO), heeft jarenlang hun prestigieuze kaarten gratis weggeven aan talloze onderzoekers

CEO Jensen Huang, Chris Malachowsky, Curtis Priem

Begonnen met crypto-minen, nu vooral actief in onderzoeksvelden als Artificiële Intelligentie.

Beeld:

Videobeelden van Jenssen Huang in interviews en tijdens talks/presentaties

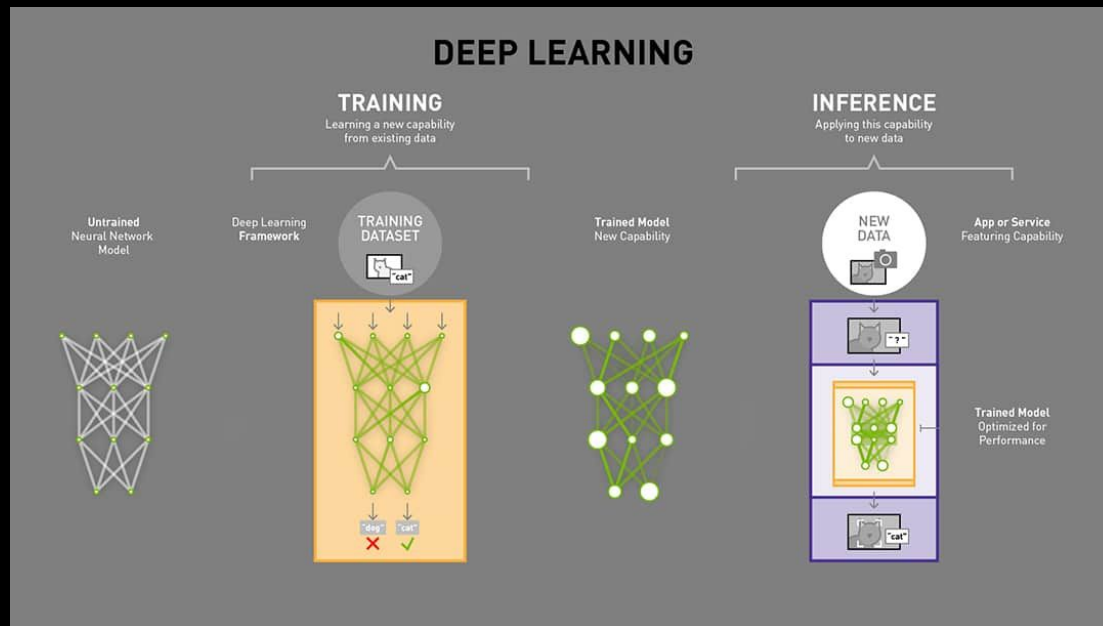


SCENE 9

Met als resultaat dat GPU's van NVIDIA cruciaal zijn voor de twee belangrijke stappen van AI...

Namelijk:

TRAINING & INFERENCE (groot in beeld)



SCENE 10

Training: In deze stap wordt het model getraind met behulp van een dataset die bestaat uit invoergegevens (kenmerken) en bijbehorende doeluitvoer (labels)

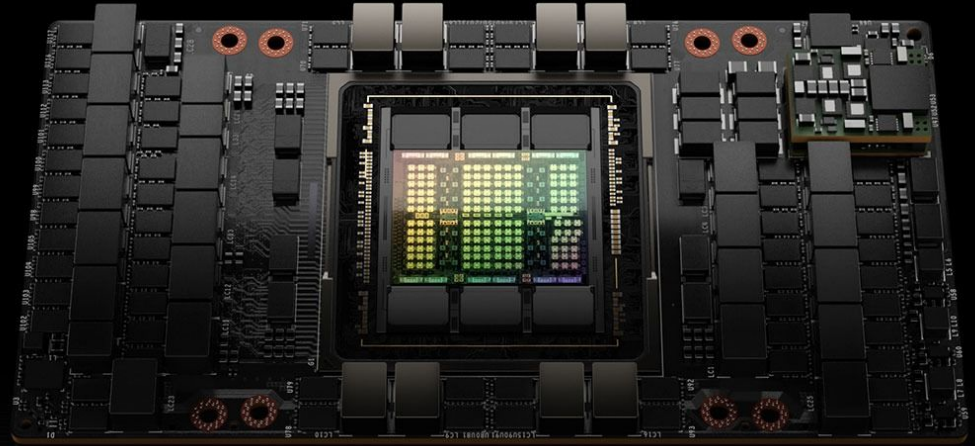
Je leert het model patronen en structuren aan in de gegevens, door middel van vele parallele matrix berekeningen.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix}$$

SCENE 11

Inference: nadat het model is getraind kan je het gebruiken maar ook hiervoor heb je een GPU nodig om het model te laten werken.

(H100 chip, en orders van 50.000-150.000stuks)



SCENE 12

Sinds dit de bouwblokken zijn voor alle soorten AI, heeft heel silicon valley massa orders gedaan voor de allerbeste chips van NVIDIA. De H-serie (H100 en H200)

Door het verkopen van deze chips is dit bedrijf het 3de meest waardevolle bedrijf op aarde, met zo'n 2 biljoen aan marktwaarde.

Nvidia on Track To Breach \$2 Trillion in Value



SCENE 13

Silicon valley en in het speciaal de grote klanten van deze GPU's: Microsoft, Google en Amazon moeten (gemiddeld) \$30,000 afrekenen voor 1 h100 kaart. Maar voor het trainen van een LLM heb je honderdduizenden van deze chips nodig die samenwerken. De chips zijn zelfs zo waardevol dat bedrijven ze gebruiken als onderpand voor leningen.



De toekomst van NVIDIA

SCENE 14

Het gevaar voor NVIDIA is, is dat deze bedrijven zelf hun eigen GPU's gaan ontwikkelen 'in-house' wat zal betekenen dat de grootste klanten wegvallen. Dit komt omdat er schattingen zijn dat de H100 chip ongeveer \$3,000 kost om te produceren.

Ook is hierin het bedrijf Nvidia erg uniek ze produceren deze chip namelijk helemaal niet zelf. Deze kosten worden gemaakt door de foundry TSMC in Taiwan.

Nvidia is namelijk verantwoordelijk voor het design/layout van deze chips. Hun taak is om elk met een vernieuwing in dit design te komen.

Door NVIDIA's voortdurende innovatie, hun overnames van kleinere, vergelijkbare bedrijven en de voortdurende vraag naar grafische kaarten voor gaming, datacenters, maar ook elektrische auto's zoals Tesla, is het bedrijf in zeer korte tijd miljarden waard geworden



SCENE 15

Hoe lang deze lucratieve positie van Nvidia nog zal duren weet niemand, er liggen veel concurrenten op de loer om van deze rush een graantje mee te pikken.

Maar anders kan het bedrijf altijd terugkeren voor een spelletje 'Doom'.

