

¿QUIÉN ES ESTE CUATE?

Enrique Rivera





The house of the goblins

GPU in the Cloud



¿QUÉ SABEMOS DE GPUs?

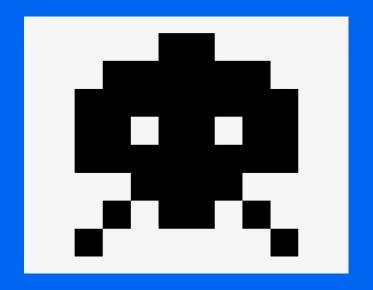


¿QUÉ HACEN LAS GPUs?

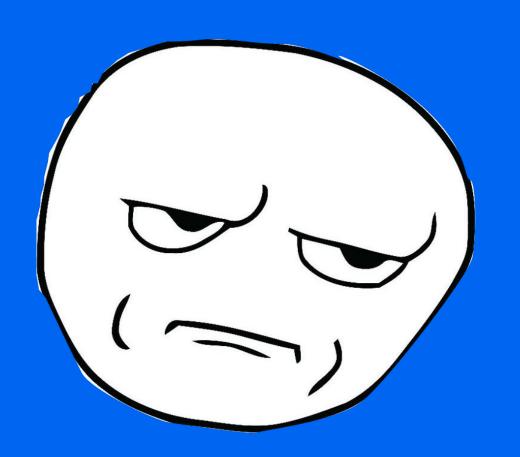






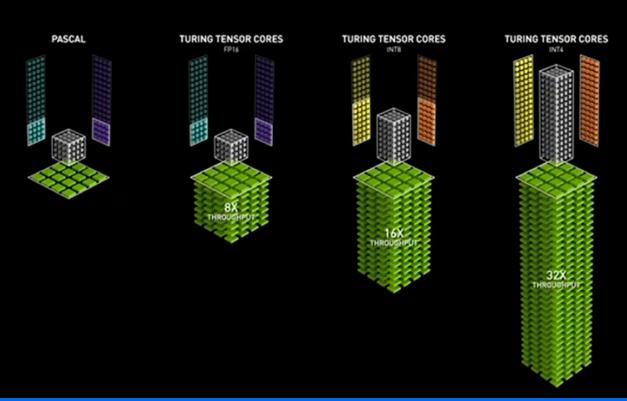






Compute-intensive Portions of the application Rest of Sequential CPU Code **GPU CPU** Optimized for Optimized for Parallel throughput Sequential Execution

T4 powered by multi-precision Turing Tensor Cores





Machine Learning Training



Machine Learning Inference



High Performance Computing



Visualization

ML Training on GPUs = Creation of your Al

Insights



Innovation



Automation







¿GPUs en la nube?



Google Cloud



NIDIA

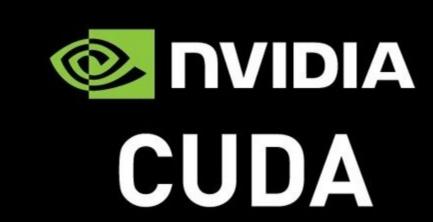
NVIDIA GPU availability on GCP

Goog	le C	loud	GPU	Type

VM Configuration Options

NVIDIA GPUs	Training	Inference	Compute	3D	Memory (GB)	TFLOPs (single-precision)	GPUs	vCPU	System Memory
V100 w/ NVLINK					16 / 128 max	14.0	1, 2, 4, 8	1-96	1-624 GB
P100	•			•	16 / 64 max	9.3	1, 2, 4	1-96	1-624 GB
K80	•				12 / 96 max	8.7	1, 2, 4, 8	1-64	1-416 GB
Т4	•		•		16 / 64 max	8.1	1, 2, 4	1-96	1-624 GB
P4			•		8 / 32 max	5.5	1, 2, 4	1-96	Activar W. 24 GBactivar W

¿CÓMO LO HACES?



Genial hora de trabajar en CLOUD

¡GRACIAS!





Kike Rivera