

# GPU vs CPU



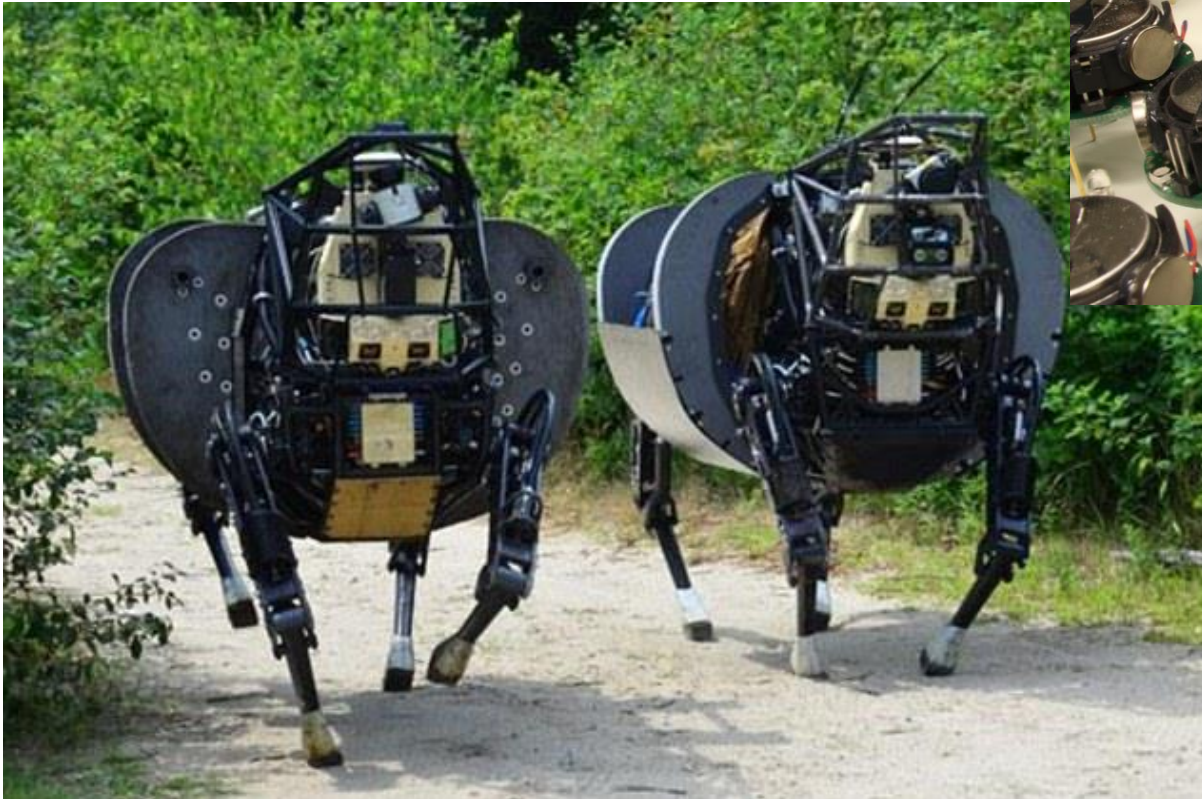
Michael Gopshtein

 @michael\_gop

 migocpp

EYE  
ALL IS ILLUMINATED

# The mental model



# Translation errors

Intel(R) Core(TM) i7-6820HQ CPU @ 2.70GHz, 2712 Mhz, 4 Core(s), 8 Logical Processor(s)

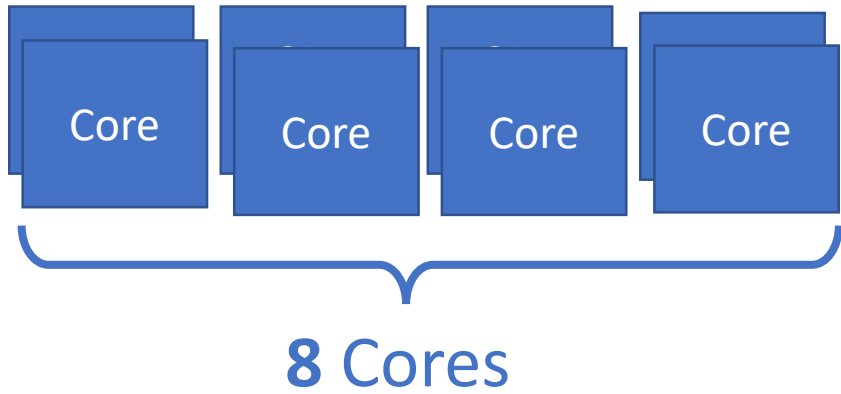
Wow!!

640 >> 8



Items	Details
Quadro M2000M	Driver version: 388.19 Direct3D API version: 12 Direct3D feature level: 11_0 CUDA Cores: 640 Graphics clock: 1098 MHz

# SIMD Perspective

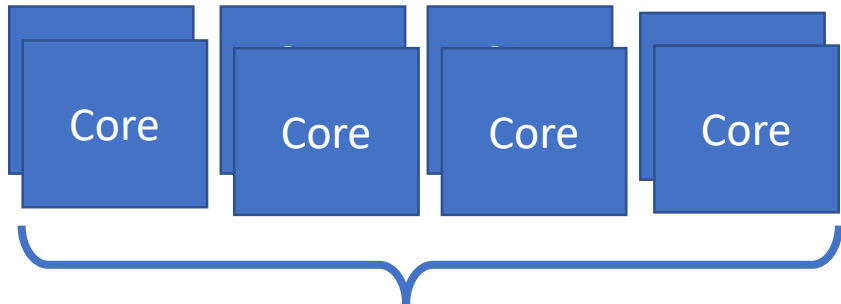


**2.7Ghz**

**8 SIMD** [AVX2 256-bit]



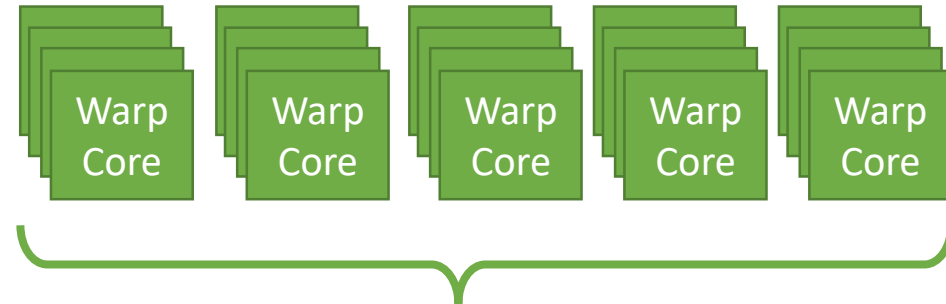
# SIMD Perspective



8 Cores

2.7Ghz

8 SIMD [AVX2 256-bit]



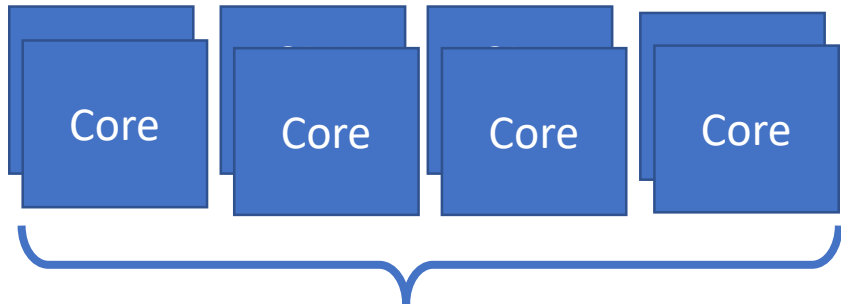
20 Cores

1.1Ghz

32 SIMD [*warp*]



# SIMD Perspective

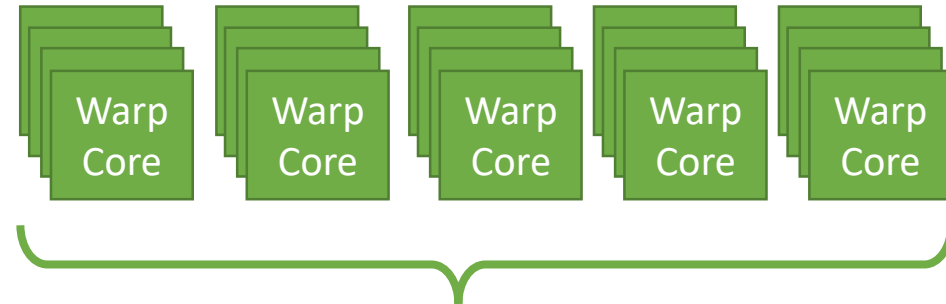


8 Cores

2.7Ghz

8 SIMD [AVX2 256-bit]

**170 GFLOPS (SIMD)**



20 Cores

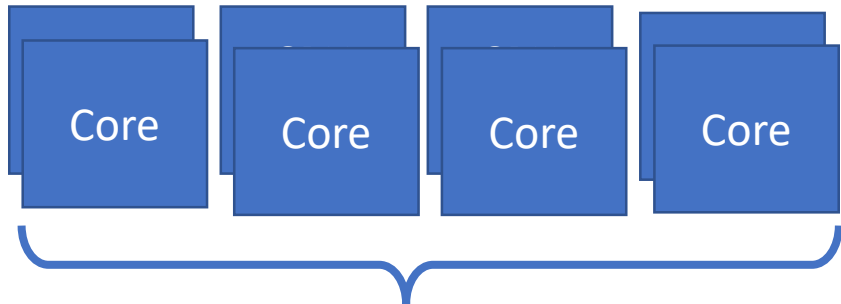
1.1Ghz

32 SIMD [*warp*]

**700 GFLOPS (SIMD)**



# SIMD Perspective



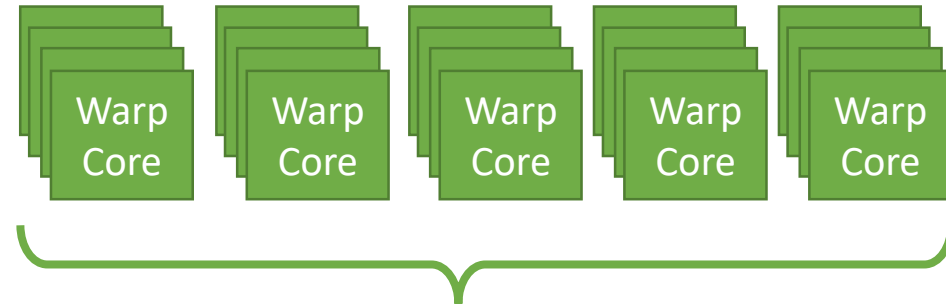
8 Cores

2.7Ghz

8 SIMD [AVX2 256-bit]

**170 GFLOPS (SIMD)**

**21.3 GFLOPS (single)**



20 Cores

1.1Ghz

32 SIMD [*warp*]

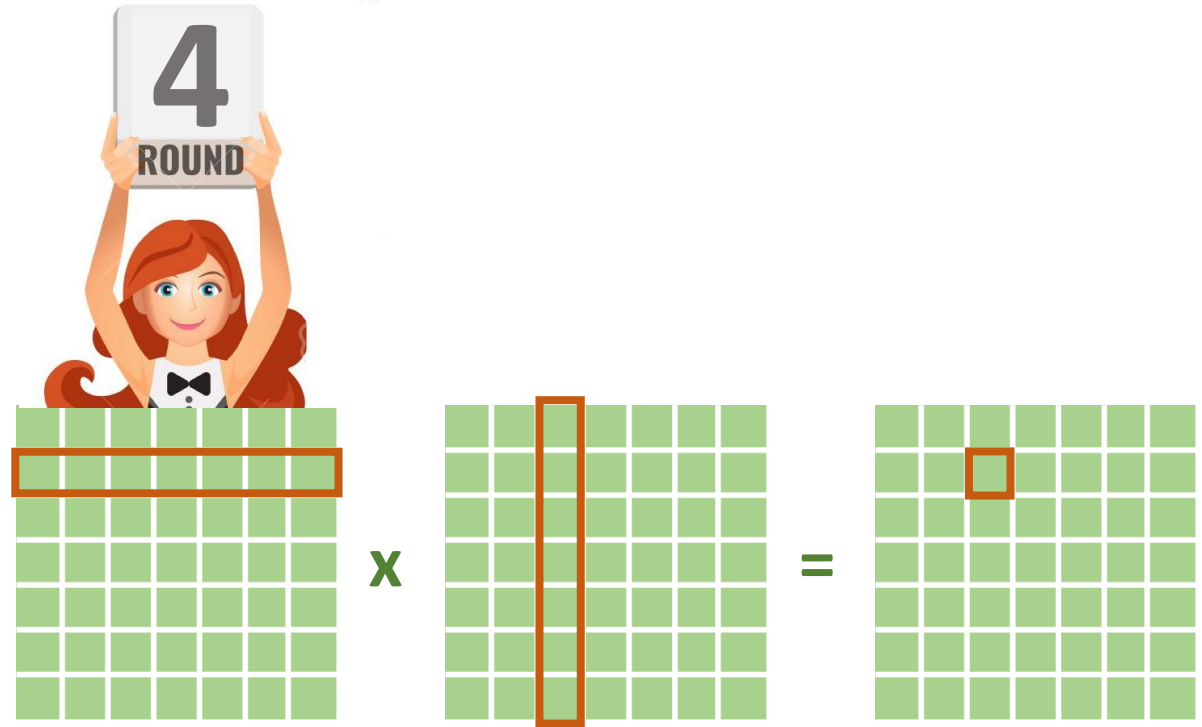
**700 GFLOPS (SIMD)**

**21.9 GFLOPS (single)**



# Test Case

Multiply 2 matrices,  $N \times N$   
 $2N^3$  float operations

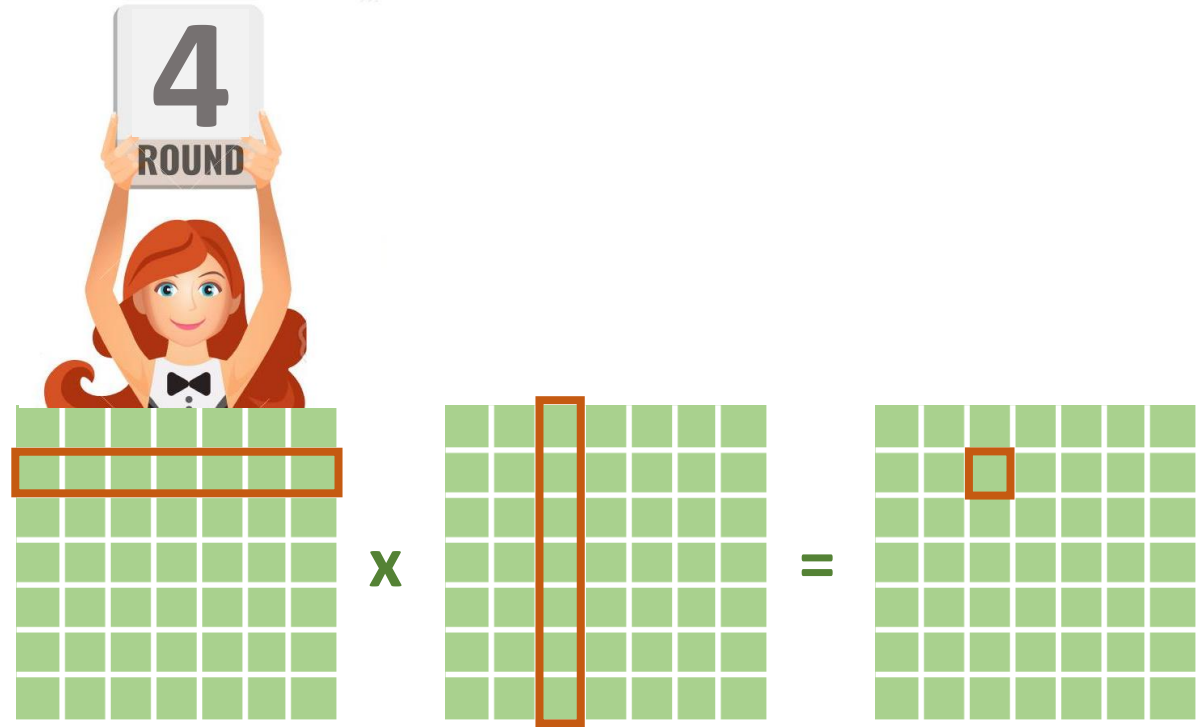


for $N=1024$	Theoretical
<b>CPU</b>	12.2 msec
<b>GPU</b>	3 msec



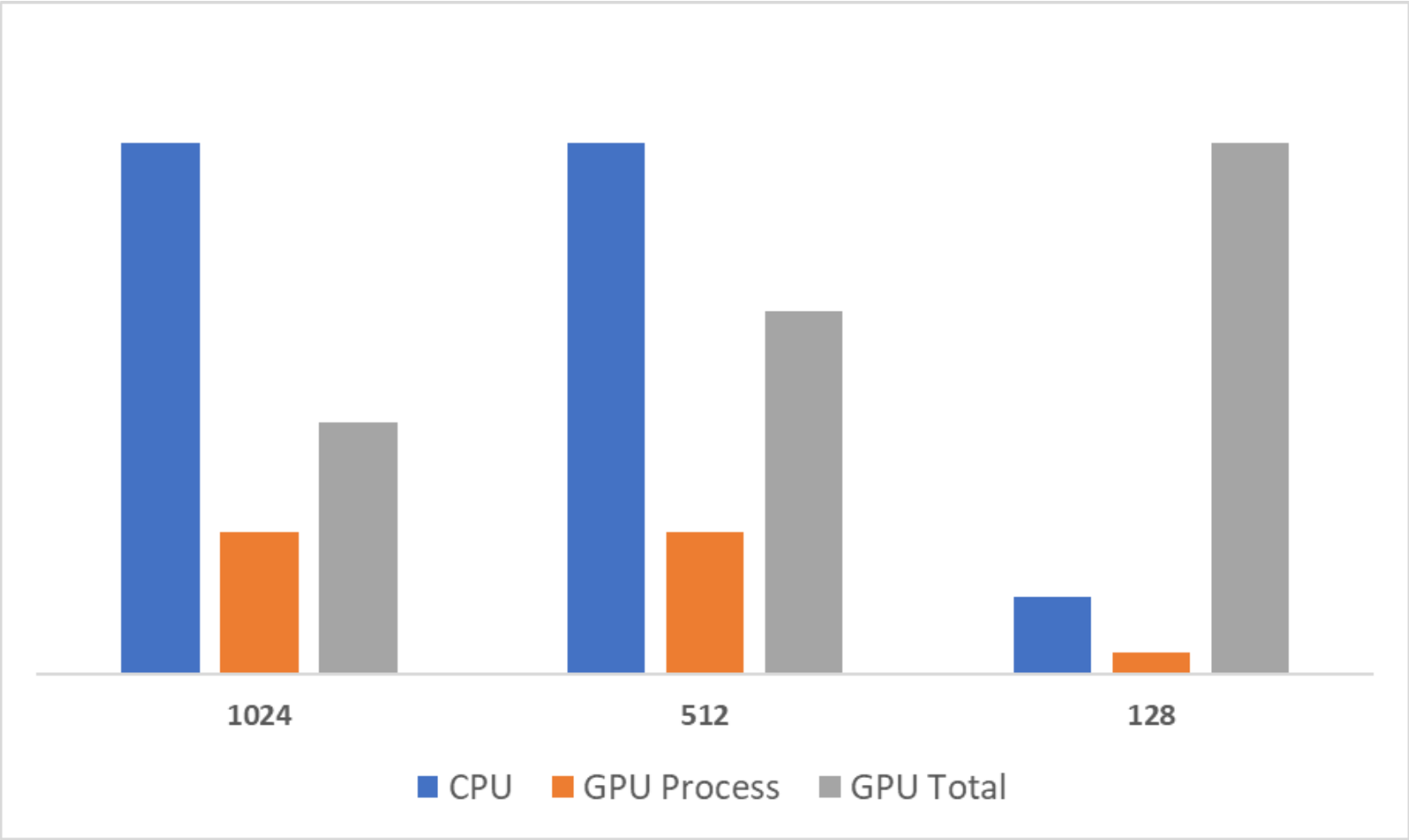
# Test Case

Multiply 2 matrices, NxN  
 $2N^3$  float operations



for N=1024	Theoretical	Measured	With Data Copy
<b>CPU</b> (Eigen + Intel MKL)	12.2 msec	<b>12</b> msec	---
<b>GPU</b> (cublas)	3 msec	<b>3.2</b> msec	5.7 msec

# Test Case



Interpolated

**LIVE**



**CORE C++**

# CPU VS GPU

**June 28th**

**MICHAEL GOPSHTAIN | COMPARING APPLES TO WATERMELONS**