

Groq GroqChip 1

Tensor Streaming Processor

PCIe

2019

14nm

OVERVIEW

—	—	—	215 W
FP32 TFLOPS	Memory On-die SRAM	Memory Bandwidth	TDP

PERFORMANCE METRICS

Precision	Peak (dense)	Bits	Type
FP64	—	—	Standard
FP32	—	32	Standard
TF32	—	—	Standard
BF16	—	—	Standard
FP16	188.0	16	Standard
FP8	—	—	Standard
FP6	—	—	Standard
FP4	—	—	Standard
INT8	750.0	8	Standard

MEMORY SPECIFICATIONS

Capacity	—
Type	On-die SRAM
Bandwidth	—
Interface Width	—

POWER & EFFICIENCY

TDP	215 W
Max power	300 W
FP32 dense TFLOPS per W of TDP	—
FP16 dense TFLOPS per W of TDP	0.87 TFLOPS/W

Per W of TDP figures divide dense peak by accelerator TDP.
Board power only: excludes host CPUs, networking, cooling and facility overhead.

HARDWARE DETAILS

Vendor: Groq	Architecture: Tensor Streaming Processor	Process Node: 14nm
Form Factor: PCIe	Launch Year: 2019	

COMPUTE ENGINE

Base Clock: 900 MHz

INTERCONNECT

GPU-to-GPU: RealScale	Links: 16	PCIe: 4.0 x16
-----------------------	-----------	---------------

CHIP DESIGN

Transistors: 26.8 billion	Die Size: 725 mm ²
---------------------------	-------------------------------

ENTERPRISE FEATURES

ECC: Yes	Sparsity: No
----------	--------------