SC_AS1

Zhigao XIE

February 2023

1 FLOPS

For AMD 6800HS, with 1 socket, 8 cores, Max. boost clock of 4.7GHz, supporting AVX2 instruction, has about 600GFLOPS.

2 Latency

The latency between CPU and memory is about 3.3×10^{-10} s. The CPU clock time for Intel i9-13900K is 5.8GHz, which means latency is about 1.7×10^{-10} s. It's twice faster than that data transmission between CPU and memory.

3 NPU, FPGA, TPU

Google Neural Network Processing Unit (NPU) NPU has a large number of processing cores, each of them are design to tackle special tasks, which make it much more powerful in parallel processing. High-bandwidth memory system of NPU allows it to quickly access and process large amounts of data beneficial to 'data-driven processing. The NPU has a high-speed interconnect system that connects the processing cores and memory system. This interconnect system allows the NPU to perform parallel processing, which is critical for deep learning algorithms.

Field-Programmable Gate Array (FPGA) FPGA is a type of integrated circuit that can be programmed after manufacturing to perform a specific set of tasks. The programmable Logic Block, I/O Block and Interconnect are the main components of FPGA and the Look Up Table and SRAM enables it's high performance.

Tensor Processing Unit (TPU) TPU is a custom-designed chip specifically built for machine learning and deep learning applications. The TPU is optimized for matrix operations and provides high performance for AI workloads. It has a high-bandwidth memory system and a large number of processing cores, which allow it to perform complex computations much faster than

a general-purpose processor. Additionally, the TPU has a specialized architecture designed to support machine learning, making it easy to implement deep learning models on the chip.

4 Floating point numbers

Floating-point numbers violate the axioms of real numbers because they represent real numbers with a limited number of bits, which can lead to approximations and rounding errors. Associativity of addition, Existence of multiplicative inverses and Distributive law are violated by floating point numbers