Arm Neoverse N1 CPU

Accelerating the transformation to a scalable cloud to edge infrastructure

Today's world is more connected than ever, and the volume and diversity of connected devices will only continue to grow. These devices will increasingly deliver data to the cloud or be acted upon by artificial intelligence at the edge of the network—often in real-time.

To support this explosion of traffic, the network and compute infrastructure must evolve from the cloud all the way to the edge. The new world of cloud-to-edge compute requires high performance secure systems and architectures.

Arm Neoverse N1 CPU

The Arm Neoverse N1 CPU delivers the performance, features, and scalability needed to accelerate the transformation to a scalable cloud-to-edge infrastructure.

Revolutionary compute performance

N1 CPU is truly revolutionary, delivering industry-leading socket performance at half the power, with server-class thread performance. Compared to Arm's Cortex-A72 processor deployed in various infrastructure applications including servers, N1 CPU delivers:

- Up to 2.5X performance on data center workloads
- 5X better machine learning vector performance
- 30% better performance per watt in same technology node

Features specific to infrastructure

The Neoverse N1 CPU delivers a host of architectural and implementation features targeted at the infrastructure market:

- Large-system core scalability with atomics, cache stashing and I-cache coherency, atomics and cache stashing improves high core count scalability
- Efficient virtualization for Type1 and Type2 hypervisors
- State of the art RAS with data poisoning and error injection
- Intelligent power and thread performance management at runtime
- Statistical profiling to enable software optimization

Integrated platform designed for extreme range of scale and diversity of compute

The Arm Neoverse N1 Platform comprises the N1 CPU and supporting system IP connected via a coherent mesh interconnect. The platform is optimized for low-latency and bandwidth efficiency, to deliver extreme core scalability from sub-35W 8-core systems to 128+ cores in servers. A diverse ecosystem can further deliver integrations using CCIX to maintain coherency across multiple sockets and chiplets, providing differentiation with on/off-chip accelerators. N1's world-class performance and power efficiency, paired with innovative memory architectures, can drive integrations beyond 128 cores to realize new levels of parallelism and performance. Seamless heterogeneous compute systems are made possible through Armv8.2-A architectural compatibility with the Neoverse E1 CPU, and a shared Arm NN framework with Arm's Project Trillium ML accelerators to fuel the next-generation of solutions targeting 5G deployment, data analytics and machine learning.

Specifications and Features

- 64kB L1 instruction and data cache
- Up to 1MB large private L2 cache
- Up to 128MB of shared system level cache through a low-latency direct-connect interface to the CMN-600 mesh
- Microarchitecture for large footprint, branch-heavy workloads
- I-cache coherency to enable a broader range of server workloads
- Double the vector and crypto compute bandwidth over previous generation
- No-compromise, full-frequency, sustainable compute efficiency managed at runtime
- Server-class virtualization, RAS and code profiling