



Arteris Interconnect IP Deployed in NeuReality Inference Server for Generative AI and Large Language Model Applications

October 10, 2023

FlexNoC network-on-chip IP seamlessly provides connectivity across the NR1 chip within the inference server to efficiently meet high-density, low-latency AI performance needs at a minimal total cost of ownership.

CAMPBELL, Calif., Oct. 10, 2023 (GLOBE NEWSWIRE) -- Arteris, Inc. (Nasdaq: AIP), a leading provider of system IP which accelerates system-on-chip (SoC) creation, today announced that NeuReality has deployed Arteris FlexNoC interconnect IP as part of the NR1 network addressable inference server-on-a-chip to deliver high-performance, disruptive cost and power consumption improvements for machine and deep learning compute in its AI inference products. This integration is architected in an 8-hierarchy NoC with an aggregated bandwidth of 4.5TB/sec, meeting low latency requirements for running AI applications at scale and lower cost. The NeuReality inference server targets Generative AI, Large Language Models (LLMs) and other AI workloads.

"The new era of Generative AI with LLMs requires large-scale computing that is faster, easier, and less expensive. We created a category of microprocessors for today's AI-centric data centers supporting sustainability," said Moshe Tanach, co-founder and CEO of NeuReality. "Arteris has earned a notable reputation in the market which together with their AI-ready network-on-chip technology were determining factors in our decision to adopt their FlexNoC IP for our AI server. This IP enabled us to successfully address AI performance requirements, scalability, high density, and low latency, all with a minimal total cost of ownership."

NeuReality's innovative NR1 server-on-a-chip, is the first Network Addressable Processing Unit (NAPU), which is a workflow-optimized hardware device with specialized processing units, native network and virtualization capabilities. It provides native AI-over-fabric networking, including full AI pipeline offload and hardware-based AI hypervisor capabilities. The ability to offload CPUs, GPUs and even deep learning accelerators to multiple NR1 chips is what makes it possible for NeuReality's inference server to effectively deliver up to 10 times the performance with less power consumption and at a fraction of the cost in its inference server.

"Developing inference platforms for advanced AI and machine learning applications, such as Generative AI, is a complex process that requires a deep understanding of both software and hardware, along with state-of-art connected chip development," said K. Charles Janac, president and CEO of Arteris. "We are thrilled to be working with NeuReality, and deploying Arteris IP to provide AI connectivity, supporting their vision of cost-effective, high-performance AI at scale."

About Arteris

Arteris is a leading provider of system IP for the acceleration of system-on-chip (SoC) development across today's electronic systems. Arteris network-on-chip (NoC) interconnect IP and SoC integration automation technology enable higher product performance with lower power consumption and faster time to market, delivering better SoC economics so its customers can focus on dreaming up what comes next. Learn more at [arteris.com](https://www.arteris.com).

About NeuReality

The mission of NeuReality is to make AI easy – both in its deployment and use in the data center. By taking a systems-level approach, the team of industry experts serve AI inference holistically, determine pain points, and deliver purpose-built, affordable solutions that democratize AI adoption for organizations large and small, in technology and non-technology businesses. The revolutionary combination of AI technology, business model, and people accelerates the possibilities of AI. Learn more at [neureality.ai](https://www.neureality.ai).

© 2004-2023 Arteris, Inc. All rights reserved worldwide. Arteris, Arteris IP, the Arteris IP logo, and the other Arteris marks found at <https://www.arteris.com/trademarks> are trademarks or registered trademarks of Arteris, Inc. or its subsidiaries. All other trademarks are the property of their respective owners.

Media Contact: Gina Jacobs Arteris +1 408 560 3044 newsroom@arteris.com