Broadcom Unveils Industry's Highest Performance Fabric for AI Networks

April 18, 2023

Connected by Broadcom, New Jericho3-AI Provides High-Performance Ethernet for a 32,000 GPU Cluster

SAN JOSE, Calif., April 18, 2023 (GLOBE NEWSWIRE) -- Broadcom Inc. (NASDAQ:AVGO) announced today that it has delivered Jericho3-AI, enabling the industry's highest performance fabric for artificial intelligence (AI) networks. Jericho3-AI revolutionizes AI networking with best-in-class capabilities such as perfect load balancing, congestion-free operation, ultra-high radix, and Zero-Impact Failover, all culminating in significantly shorter job completion times for any AI workload.

Jericho3-AI comes to market at a time when the use of AI is skyrocketing. According to a new forecast from International Data Corporation (IDC), global spending on AI will reach $154 billion in 2023 and $300 billion or more by 2026. The 2023 figure represents an increase of 26.9 percent over 2022.[1]

Fabrics based on Jericho3-AI will help network operators handle the ever-expanding workloads AI demands will present.

"The benchmark for AI networking is reducing the time and effort it takes to complete the training and inference of large-scale AI models," said Ram Velaga, senior vice president and general manager, Core Switching Group, Broadcom. "Jericho3-AI delivers significant reduction in job completion time compared to any other alternative in the market."

AI workloads have unique characteristics such as a low number of large, long-lived flows, all starting concurrently upon completion of an AI computation cycle. The Jericho3-AI fabric provides the highest performance for these workloads with unique functionality designed specifically for AI workloads:

- **Perfect load balancing** equally sprays traffic over all links of the fabric, ensuring maximum network utilization under the highest network loads.
- **Congestion-free operation** with end-to-end traffic scheduling ensures no flow collisions and no jitter.
- **Ultra-high radix** uniquely allows the Jericho3-AI fabric to scale connectivity to 32,000 GPUs, each with 800Gbps, in a single cluster.
- **Zero-Impact Failover** functionality ensures sub-10ns automatic path convergence, resulting in no impact to job completion time.

Leveraging this unique functionality, the Jericho3-AI fabric provides at least 10 percent shorter job completion times versus alternative networking solutions for key AI benchmarks such as All-to-All. This performance improvement has a multiplicative effect on decreasing the cost of running AI workloads since it implies that expensive AI accelerators are used 10 percent more efficiently. The network, in effect, pays for itself.

The Jericho3-AI fabric offers 26 petabits per second of Ethernet bandwidth, almost four times the bandwidth of the previous generation, while simultaneously delivering 40 percent lower power per gigabit.

"Cloud operators will upgrade their AI infrastructure to address the massive growth in bandwidth, driven by a new generation of high-capacity GPUs and the emergence of large language models," said Bob Wheeler, principal analyst at Wheeler's Network **.** "Jericho3-AI offers a high-bandwidth, low-latency and low-power choice for networks connecting tens of thousands of GPUs, revolutionizing the economics of building and maintaining AI clusters for this exciting new era."

A unique feature of the Jericho3-AI fabric is that it provides the highest performance while also enabling the lowest total cost of ownership. This is achieved by virtue of attributes such as long-reach SerDes, distributed buffering, and advanced telemetry, all provided using industry-standard Ethernet. These factors provide a high degree of flexibility in network architectures and deployment options with the largest ecosystem of hardware and software providers.

Availability

To learn more about Jericho3-AI here. Jericho3-AI is now available to qualified customers. For more information on all Broadcom products please visit www.broadcom.com.

About Broadcom
Broadcom Inc. (NASDAQ: AVGO), a Delaware corporation headquartered in San Jose, CA, is a global technology leader that designs, develops and supplies a broad range of semiconductor and infrastructure software solutions. Broadcom's category-leading product portfolio serves critical markets including data center, networking, enterprise software, broadband, wireless, storage and industrial. Our solutions include data center networking and storage, enterprise, mainframe and cybersecurity software focused on automation, monitoring and security, smartphone components, telecoms and factory automation. For more information, go to https://www.broadcom.com.

Broadcom, the pulse logo and Connecting Everything are among the trademarks of Broadcom. The term "Broadcom" refers to Broadcom Inc., and/or its subsidiaries. Other trademarks are the property of their respective owners.

Press Contact:
Jon Piazza
Corporate Communications
press.relations@broadcom.com

Telephone: +1 408 433 7924

Industry Quotes:

Michael KT Lee, Senior Vice President, R&D, Accton
“With the availability of Jericho3-AI, Accton can create the building blocks for the next generation AI clusters, pushing the scale boundaries once again, significantly increasing the radix and throughput of AI fabric and drastically improving the power efficiency for mega-scale GPU clusters.”

Anshul Sadana, Chief Operating Officer & Senior Vice President, Arista
“The Arista 7800R3 running EOS is the premier AI spine supporting exponential growth of AI workloads and network traffic at petabit scale. We are excited to collaborate with Broadcom and drive the AI revolution using the innovative Ethernet-based cluster approach demonstrating groundbreaking performance in reducing the runtime of AI training and inference workloads.”

Gavin Cato, Head & Chief Technology Officer, Hardware Platform Solutions, Celestica
“The Jericho3-AI fabric architecture introduces a new way to build large-scale networks. Celestica is excited to see the next-generation building blocks of the merchant silicon-based architecture and the adoption of advanced AI training and inference clusters in support of open technologies.”

Ido Susan, Chief Executive Officer, DriveNets
“DriveNets has successfully partnered with Broadcom to deploy the world’s largest Core Routers based on Distributed Disaggregated Chassis (DDC) architecture. The Jericho3-AI fabric raises the bar for AI cluster scales. The unique properties of this fabric’s, perfect load-balancing and Zero-Impact Failover combined with DriveNets’ DDC creates an industry-leading utilization for AI training and inference jobs. DriveNets NOS running on a Jericho J3 platform gives both hyperscalers and service providers a unique, unparalleled technological solution for future high-scale and AI networks.”

Vincent Ho, CEO, Ufispace
“Ufispace is dedicated to creating hardware required to support the success of Distributed Disaggregated Chassis architecture in Service-Provider Core and Edge networks. We are excited to leverage the Jericho3-AI fabric to expand solutions for high performance AI clusters, enabling the training of large AI models with the shortest job completion times.”

[1] IDC, Worldwide Artificial Intelligence Spending Guide, March 2023

Source: Broadcom Inc.