# Unigen Corporation Achieves Industry First with Introduction of New AI Module

"Biscotti" E1.S AI Modules Demonstrate Breakthrough Video Processing Performance In An Air-Cooled Environment

NEWARK, CALIFORNIA, UNITED STATES, June 3, 2024 /EINPresswire.com/ -- Unigen Corporation has set a new benchmark with its Biscotti Artificial Intelligence (AI) E1.S Module. When integrated with an AMD Genoa server running Network Optix's latest Enterprise Video Operating System (EVOS) software with AI, Biscotti will allow big box stores, warehouses,

Biscotti AI E1.S Module

smart cities, transportation systems, and factories to gather hundreds of video streams from IP security cameras and process them in a single server, either on the premises or in a co-location data center.
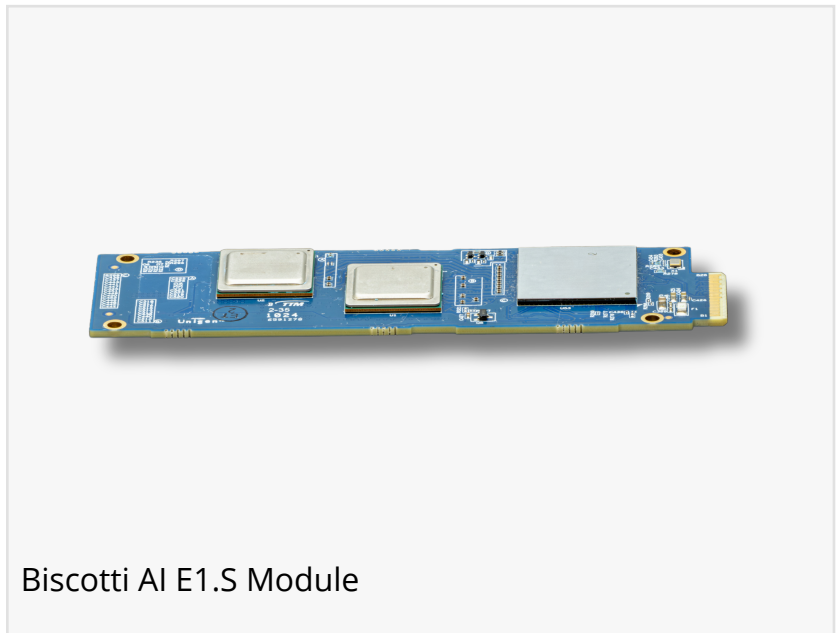
For the first time ever, an air-cooled AI server has processed 64 streams of standard model YOLOv4 in 720p resolution at 25 frames per second with 80 class object detection. In plain language, this means decoding and processing live AI video from 64 IP cameras while utilizing less than 50% of the CPU or AI. This has been accomplished all without the expense, weight and power use of liquid cooling.

"Unigen's new AI server technology running our Nx Enterprise Video Operating System scales video AI solutions beyond what was previously attainable." said James Cox, VP Business Development of Network Optix. "By running our AI manager software on new TPU technology on OCP standard modules in high-density high-performance servers, Unigen has significantly reduced the overall cost-per-channel."

"Unigen's innovative approach to integrating our Hailo-8 AI accelerator has changed how we approach the server market," said Orr Danon, CEO of Hailo. "By using an E1.S to deliver a full 52 TOPS at a low power consumption, we enable servers to run cooler and deliver faster AI

processing."

"AIC is proud to have collaborated with Unigen to be the first server OEM to deliver this innovative AI architecture," said Michael Liang, CEO of AIC. "Our EB202-CB-UG server can support 8 Unigen Biscottis with a total power consumption of under 500 watts, delivering 21,500 frames per second on the resnet_v1_50 neural network benchmark, all at a low TCO."

"Unigen has addressed the global need to reduce the power footprint of AI inference data centers." said Jennifer Cooke, analyst at IDC. "The Biscotti architecture is a compelling offering for organizations that require high-performance systems yet are conscious of the need to operate in a manner consistent with their corporate environmental sustainability goals."

High Performance, Low Power: The Biscotti E1.S AI Module provides 52 TOPS from as little as 10 Watts. By integrating two Hailo-8 Edge AI processors, each featuring up to 26 tera-operations per second (TOPS), Biscotti provides exceptional performance in the realm of edge processor modules. The advanced architecture harnesses the core properties of neural networks, allowing edge devices to run deep learning applications at full scale more efficiently, effectively, and sustainably than other AI chips and solutions. By targeting an E1.S standard form factor, it becomes feasible to power both AI processors, resulting in performance that excels in power efficiency.

Plug-and-Play for Servers and Edge Devices: Biscotti can be inserted directly into E1.S slots, typically used by SSDs, to instantly enhance server configurations with AI capabilities. It supports multiple parallel Neural Networks from a large array of camera inputs, or can be integrated into a single Large Language Model (LLM) array to solve complex AI cases. With significantly lower power than GPU modules or Add-In-Cards, a solution using Biscotti can change the game for a data center's power envelope.

Neural Network Models & Application Support: The integrated Hailo AI processors on Biscotti have a robust software suite that supports state-of-the-art deep learning models and applications. Additionally, it is equipped with a comprehensive dataflow compiler that enables customers to port their neural network models easily and quickly. Biscotti supports various AI frameworks, including TensorFlow, TensorFlow Lite, Keras, PyTorch, and ONNX, making it ideal for edge neural networks today and generative AI in the future.

Unigen is excited to announce its participation at ComputeX in Taipei, taking place from June 4 to 7, 2024. Attendees are cordially invited to visit the AIC (N0806) and Network Optix (N0314) booths, where Unigen will showcase the Biscotti AI E1.S Modules in action.

About Unigen
Unigen, founded in 1991, is an established global leader in the design and manufacture of original and custom SSD, DRAM, NVDIMM modules and Enterprise IO solutions. Headquartered in Newark, California, the company operates state of the art manufacturing facilities (ISO-

9001/14001/13485 and IATF 16949) in the Silicon Valley Bay Area of California and near Hanoi Vietnam, along with five additional engineering and support facilities located around the globe. Unigen markets its products to both enterprise and client OEMs worldwide focused on embedded, industrial, networking, server, telecommunications, imaging, automotive and medical device industries. Unigen also offers best in class electronics manufacturing services (EMS), including new product introduction and volume production, supply chain management, assembly & test, TaaS (Test-as-a-Service) and post-sales support. Learn more about Unigen's products and services at unigen.com.

Jeff Chang
Unigen Corporation
jchang@unigen.com
Visit us on social media:
LinkedIn

This press release can be viewed online at: https://www.einpresswire.com/article/716929057