Industry's First MCU-based Implementation of Glow Neural Network Compiler for Machine Learning at the Edge

July 28, 2020

- Demonstrating the wide-ranging benefits of Glow NN compiler for vision- and voice-based machine learning applications - NXP is the first semiconductor vendor to deliver a 2-3x performance jump for MCUs over the standard version of Glow
- Originally developed by Facebook, the open source Glow compiler is now available in NXP's eIQ™ Machine Learning Software Development Environment delivering high performance inferencing for NXP's i.MX RT series of crossover MCUs
- NXP's implementation of Glow targets Arm ® Cortex®-M cores and the Cadence® Tensilica® HiFi 4 DSP, with platform-specific optimizations for its i.MX RT series of crossover MCUs

EINDHOVEN, The Netherlands, July 28, 2020 (GLOBE NEWSWIRE) -- NXP Semiconductors N.V. (NASDAQ: NXPI) today released its eIQ Machine Learning (ML) software support for Glow neural network (NN) compiler, delivering the industry's first NN compiler implementation for higher performance with low memory footprint on NXP's i.MX RT crossover MCUs. As developed by Facebook, Glow can integrate target-specific optimizations, and NXP leveraged this ability using NN operator libraries for Arm Cortex-M cores and the Cadence Tensilica HiFi 4 DSP, maximizing the inferencing performance of its i.MX RT685 and i.MX RT1050 and RT1060. Furthermore, this capability is merged into NXP's eIQ Machine Learning Software Development Environment, freely available within NXP's MCUXpresso SDK.

Glow Neural Network compiler for vision- and voice-based machine learning applications

Exploiting MCU Architectural Features using Glow
In May 2018, Facebook, the leading pioneer of PyTorch, introduced Glow (the Graph Lowering NN compiler) as an open source community project, with the goal of providing optimizations to accelerate neural network performance on a range of hardware platforms. As an NN compiler, Glow takes in an unoptimized neural network and generates highly optimized code. This differs from the typical neural network model processing whereby a just-in-time compilation is leveraged, which demands more performance and adds memory overhead. Directly running optimized code, like that possible with Glow, greatly reduces the processing and memory requirements. NXP has also taken an active role within the Glow open source community to help drive broad acceptance of new Glow features.

"The standard, out-of-the-box version of Glow from GitHub is device agnostic to give users the flexibility to compile neural network models for basic architectures of interest, including the Arm Cortex-A and Cortex-M cores, as well as RISC-V architectures," said Dwarak Rajagopal, Software Engineering Manager at Facebook. "By using purpose-built software libraries that exploit the compute elements of their MCUs and delivering a 2-3x performance increase, NXP has demonstrated the wide-ranging benefits of using the Glow NN compiler for machine learning applications, from high-end cloud-based machines to low-cost embedded platforms."

Optimized Machine Learning Frameworks for Competitive Advantage
The demand for ML applications is expected to increase significantly in the years ahead. TIRIAS Research forecasts that 98% of all edge devices will use some form of machine learning/artificial intelligence by 2025. Based on market projections, 18-25 billion devices are expected to include ML capabilities, even without dedicated ML accelerators, in that time frame. Consumer device manufacturers and embedded IoT developers will need optimized ML frameworks for low-power edge embedded applications using MCUs.

"NXP is driving the enablement of machine learning capabilities on edge devices, leveraging the robust capabilities of our highly integrated i.MX application processors and high performance i.MX RT crossover MCUs with our eIQ ML software framework," said Ron Martino, senior vice president and general manager, NXP Semiconductors. "The addition of Glow support for our i.MX RT series of crossover MCUs allows our customers to compile deep neural network models and give their applications a competitive advantage."

NXP's edge intelligence environment solution for ML is a comprehensive toolkit that provides the building blocks that developers need to efficiently implement ML in edge devices. With the merging of Glow into eIQ software, ML developers will now have a comprehensive, high-performance framework that is scalable across NXP's edge processing solutions that include the i.MX RT crossover MCUs and i.MX 8 application processors.

Customers will be better equipped to develop ML voice applications, object recognition and facial recognition, among other applications, on i.MX RT MCUs and i.MX application processors.

Accelerated Performance with NXP's Glow Neural Network Implementation
eIQ now includes inferencing support for both Glow and TensorFlow Lite, for which NXP routinely performs benchmarking activities to measure performance. MCU benchmarks include standard NN models, such as CIFAR-10. Using a CIFAR-10 model as an example, the benchmark data acquired by NXP shows how to leverage the performance advantage of the i.MX RT1060 device (with 600MHz Arm Cortex-M7), i.MX RT1170 device (with 1GHz Arm Cortex-M7), and i.MX RT685 device (with 600 MHz Cadence Tensilica HiFi 4 DSP).

NXP's enablement for Glow is tightly coupled with the Neural Network Library (NNLib) that Cadence provides for its Tensilica HiFi 4 DSP delivering 4.8GMACs of performance. In the same CIFAR-10 example, NXP implementation of Glow achieves a 25x performance advantage by using this DSP to accelerate the NN operations.

"The Tensilica HiFi 4 DSP was originally integrated in the i.MX RT600 crossover MCU to accelerate a broad range of audio and voice processing applications. However, as the number of ML inference applications targeting low-cost, low-power MCU-class applications has increased, the inherent DSP computational performance of the HiFi 4 DSP makes it an ideal target to accelerate these NN models," said Sanjive Agarwala, corporate VP, Tensilica IP at Cadence. "Through NXP's Glow implementation in eIQ ML software, customers of i.MX RT600 MCUs can leverage the DSP to address a number of ML applications including keyword spotting (KWS), voice recognition, noise reduction and anomaly detection."

"NXP's inclusion of the Arm CMSIS-NN software library in eIQ is designed to maximize the performance and minimize the memory footprint of neural networks on Arm Cortex-M cores," said Dennis Laudick, VP Marketing, Machine Learning at Arm. "Using a CIFAR-10 neural network model as an example, NXP is able to achieve a 1.8x performance advantage with CMSIS-NN. Other NN models should yield similar results, clearly demonstrating the benefits of this advanced compiler and our optimized NN operator library."

Availability
NXP's eIQ for Glow NN compiler is available now, delivered via MCUXpresso SDK for i.MX RT600 Crossover MCUs, as well as i.MX RT1050 and i.MX RT1060 crossover MCUs. eIQ for Glow NN compiler will be available for other NXP MCUs in the future.

About the i.MX RT Series of Crossover MCUs
The i.MX RT series is the industry's first crossover MCU portfolio, featuring a high-performance Arm Cortex-M core, real-time functionality and MCU usability at an affordable price. The series represents the convergence of low-power applications processors and high-performance microcontrollers. The i.MX RT series bridges the gap between the traditional MCUs and i.MX applications processor space, allowing MCU customers a path for significant performance and integration improvements, without sacrificing ease-of-use.

For more information, go to www.nxp.com/eiq and www.nxp.com/eiq/glow

About NXP Semiconductors
NXP Semiconductors N.V. enables secure connections for a smarter world, advancing solutions that make lives easier, better, and safer. As the world leader in secure connectivity solutions for embedded applications, NXP is driving innovation in the automotive, industrial & IoT, mobile, and communication infrastructure markets. Built on more than 60 years of combined experience and expertise, the company has approximately 29,000 employees in more than 30 countries and posted revenue of $8.88 billion in 2019. Find out more at www.nxp.com.

For more information, please contact:

America and Europe	**Greater China / Asia**
Jason Deal	Ming Yue
Tel: +44 7715228414	Tel: +86 21 2205 2690
Email: jason.deal@nxp.com	Email: ming.yue@nxp.com

NXP-IoT

A photo accompanying this announcement is available at https://www.globenewswire.com/NewsRoom/AttachmentNg/7a13fd76-8c48-43aa-91a8-e5cf6ceffa10

Source: NXP USA, Inc.