2020 embedded VISION summit®

Streamline, Simplify, and Solve for the Edge of the Future

Bill Pearson
VP IOTG, GM Developer Enabling
September 2020

intel®

Legal Disclaimer & Optimization Notice

[1]Testing by Intel as of October 12th, 2018

Deep Learning Workload Configuration. Comparing Intel® Movidius™ Neural Compute Stick based on Intel® Movidius™ Myriad™ 2 VPU vs. Intel® Neural Compute Stick 2 based on the Intel® Movidius™ Myriad™ X VPU with Asynchronous Plug-in enabled for (2xNCE engines). As measured by images per second across GoogleNetV1*. Base System Configuration: Intel® Core™ I7-8700K 95W TDP (6C12T at 3.7GHz base freq and 4.7GHz max turbo freq), Graphics: Intel® UHD Graphics 630 Total Memory 65830088 kB Storage: INTEL SSDSC2BB24 (240GB), Ubuntu* 16.04.5 Linux*-4.15.0-36-generic-x86_64-with-Ubuntu*-16.04-xenial, deeplearning_deploymenttoolkit_2018.0.14348.0, API version 1.2, Build 14348, myriadPlugin, FP16, Batch Size = 1

Software and workloads used in performance tests may have been optimized for performance only on Intel® microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

Performance results are based on testing as of October 12th, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

Intel technologies features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at Intel.com.

The benchmark results reported in this deck may need to be revised as additional testing is conducted. The results depend on the specific platform configurations and workloads utilized in the testing, and may not be applicable to any particular user's components, computer system or workloads. The results are not necessarily representative of other benchmarks and other benchmark results may show greater or lesser impact from mitigations.

INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS". NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

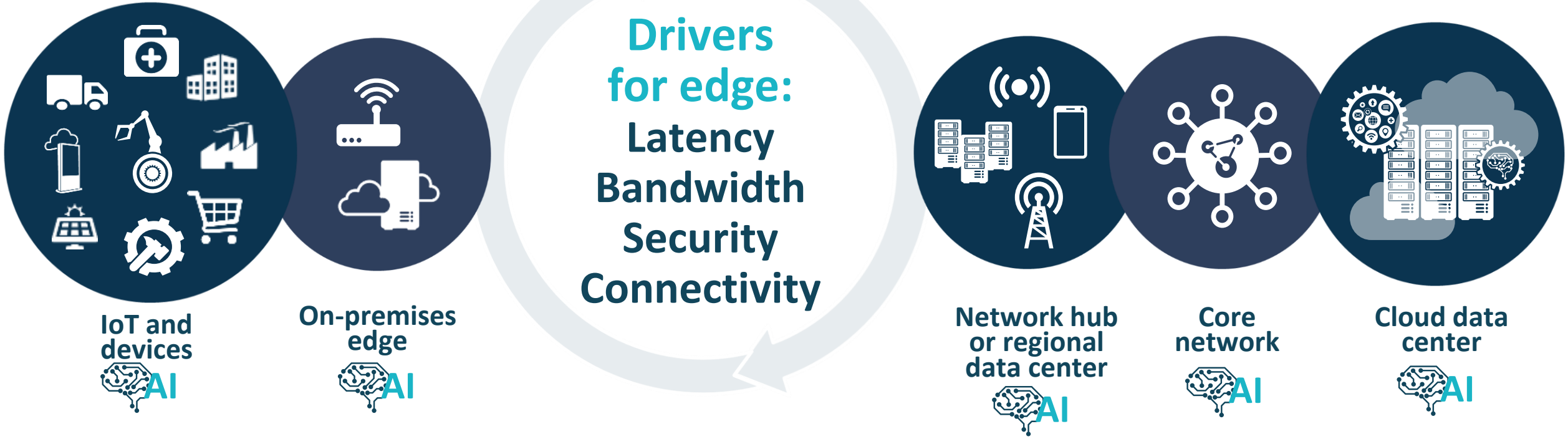
OpenVINO and the OpenVINO logo are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

* Other names and brands may be claimed as the property of others.

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.
Notice revision #20110804

Edge AI is changing the industry



50% of data will be created outside the data center or cloud[1]

43% share of AI tasks will occur on edge devices (vs. cloud) in 2023[2]

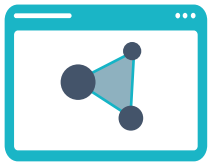
15x growth in devices with edge AI capabilities by 2023[2]

1. Gartner: December 2018 2. ABI Research

Developing for edge AI

Diverse types of edge AI applications

Wide range of unique requirements



Numerous choices in development process

Many complex options in edge AI requirements

Edge AI inference software workflow

Optimize

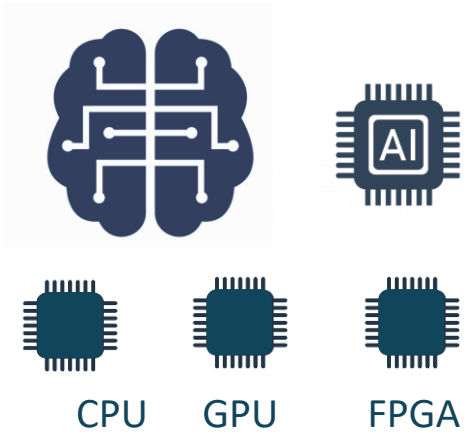
Create your application

Test

Prototype and benchmark

Deploy

Landing your solution



Scale

Rapidly deploy edge solutions

Complexity in diverse use cases, architectures, performance, tools

Edge AI has a wide range of requirements










ALGORITHMS

POWER

HARDWARE

INFERENCE

CPU

DEV TOOLS

USE CASES

SOFTWARE

SENSORS

FRAMEWORKS

DEPLOYMENT

SCALE

NETWORK

BENCHMARKS

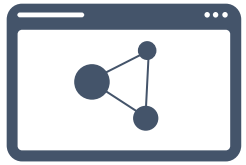
GPU

Many great frameworks, unique use cases, no one size fits all

Intel® Distribution of OpenVINO™ toolkit saves time for developers

2020 embedded VISION summit

Write once and deploy across many platforms



Trained model
-or-
Pretrained model

→

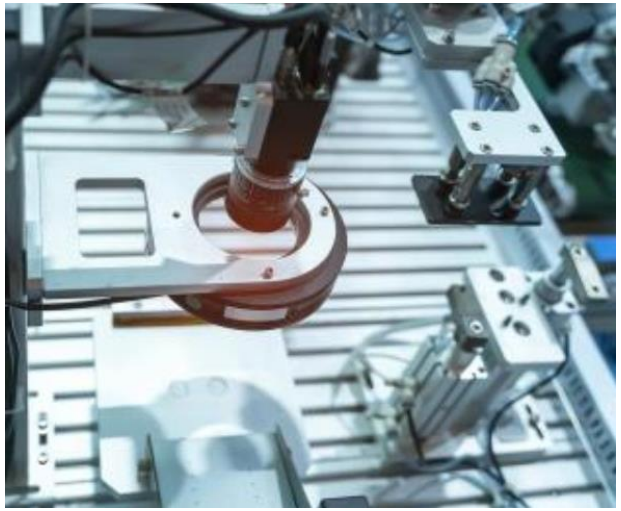
OpenVINO™

Streamlined and optimized deep learning inference

Heterogeneous, cross-platform flexibility

Free download ▸ software.intel.com/openvino-toolkit
Open source version ▸ 01.org/openvinotoolkit

Edge AI is changing our world



Rosmart

AI-enabled defect detection

Gwell Medical

Hospital logistics robot

Intel® Distribution of OpenVINO™ toolkit is making hospital logistics a good job for robots

Harrison LV

Chief Technology Officer

Gwell Medical

How to measure *real* performance

This image is provided for demonstrative purposes only. It does not reflect any actual benchmarks.

Gain access to diverse hardware

Be aware of benchmarks

Understand performance for your solution

Your performance is more than TOPS—measure the real system capability

Intel® DevCloud for the Edge: Accelerate production with cloud-based AI development

2020 embedded VISION summit

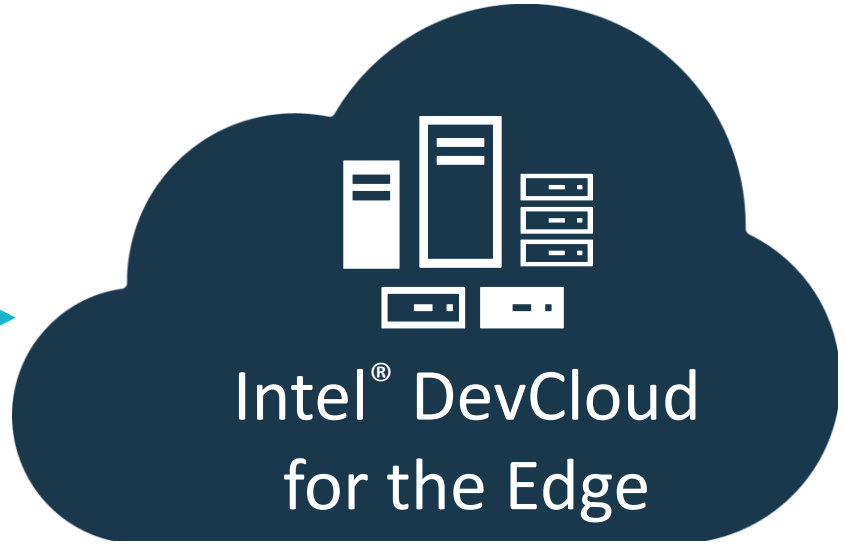
Accelerate prototype to production

OpenVINO™

Intel® Distribution of OpenVINO™ toolkit

Model optimizer

Inference engine



Latency

Bandwidth

Security

Connectivity

intel IRIS Pro GRAPHICS

Intel® GNA (IP)

intel XEON PLATINUM inside

intel CORE inside

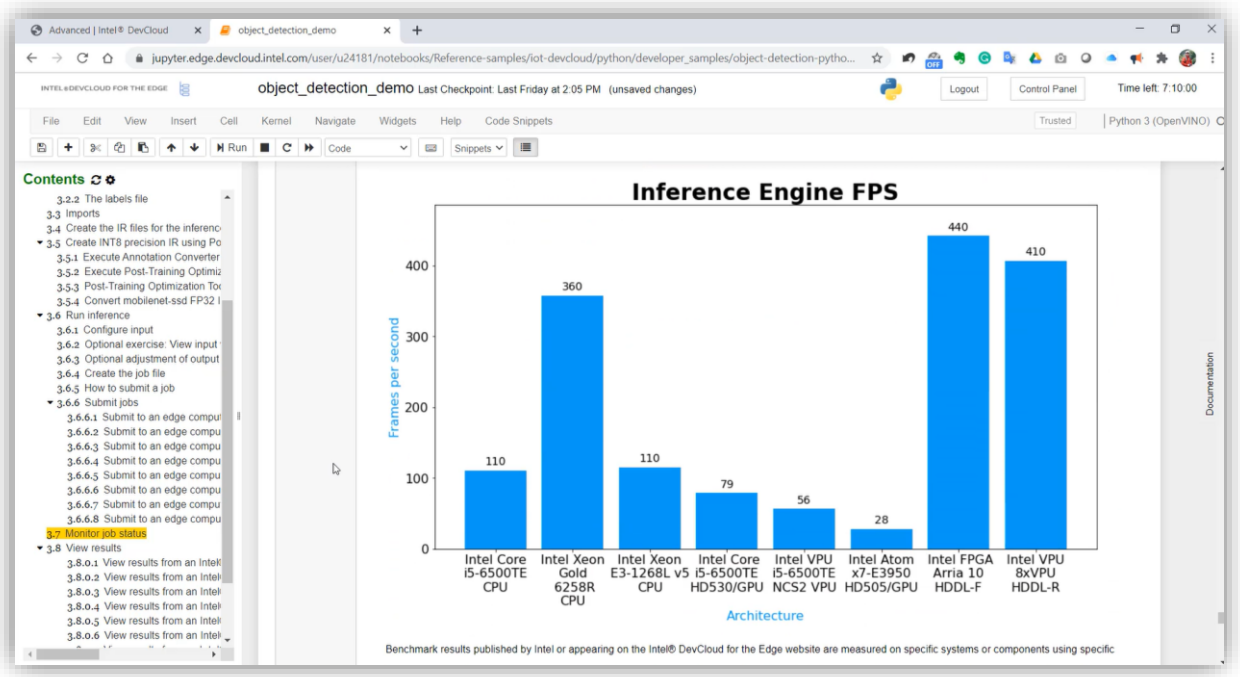
intel MOVIDIUS inside

intel ARRIA 10 inside

intel ATOM inside

devcloud.intel.com/edge

Measure hyper-customized performance comparisons among many choices



This image is provided for demonstration purposes only.

Evaluate your AI application

Immediate feedback—frames per second, performance, and more with access to the latest Intel® hardware

Edge AI Automating shelf inspection

Vispera Shelfsight

Visual inventory management

Barriers in scaling your use case to the edge


VISION


RETAIL


INDUSTRIAL


ROBOTICS


NETWORK

Ad hoc packages with unknown interoperability

Unique vertical solution needs and compliance standards

Nonmodular architecture makes it difficult to integrate applications and share data

Take the guesswork out of creating your solution

Intel® Edge Software Hub

Quickly deploy use case–specific packages for edge

OpenVINO™

OpenNESS

Intel Edge Software Hub

Use case–specific
reference implementations

EDGE INSIGHTS FOR VISION

EDGE INSIGHTS FOR INDUSTRIAL

CONVERGED EDGE INSIGHTS

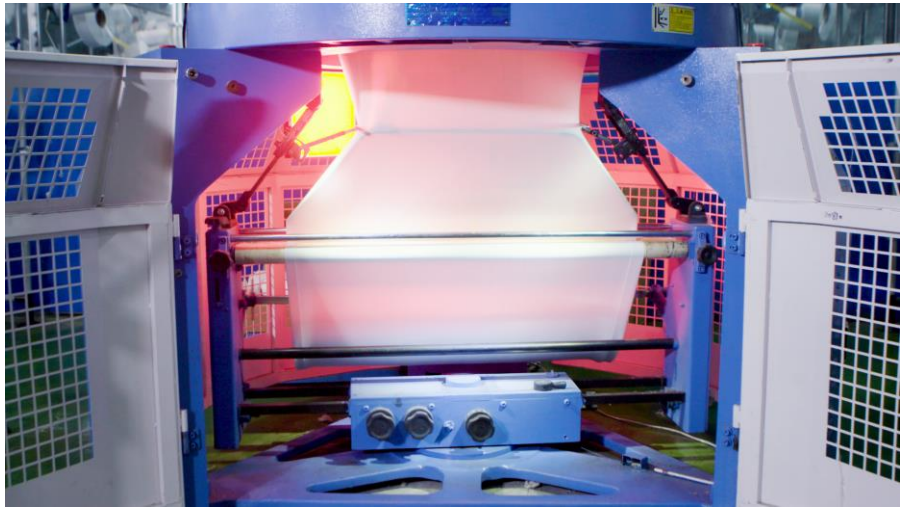
AND MORE

Complementary
ecosystem offerings

intel.com/edgesoftwarehub

Textile Factory Automation moves to the edge

Kinco 步科



Kinco

AI enables Large Scale Clothing
Production to Customizations

Simplify and streamline edge AI development cloud to edge—you can deploy anywhere

2020 embedded VISION summit

Optimize

Create your application

OpenVINO™

OpenNESS

oneAPI

Intel® Edge Software Hub

Test

Prototype and benchmark

Intel® DevCloud for the Edge

Deploy

Landing your solution

OpenVINO™

Scale

Rapidly deploy edge solutions

"Intel's DevCloud for the Edge is a 2020 Vision Product of the Year Award Winner!"

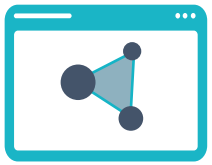
June 2020 https://www.edge-ai-vision.com/the-alliance/vision-product-of-the-year-awards/2020-vision-product-of-the-year-winners/.

edge ai + vision ALLIANCE
Formerly the Embedded Vision Alliance

Developing for edge AI

Edge AI is transforming the industry in solving important problems

No one size fits all—we provide the tools to find the right choice



Streamline development to create your optimized application

What will you solve next?

Transform the world with your DL superpowers!

Solve real-world problems using the Intel® Distribution of OpenVINO™ toolkit

Prize categories:

- Education and Economic Growth
- Environmental Sciences and Sustainability
- Industrial and Manufacturing
- Health and Well-Being
- Safety and Security in Public Spaces

Win up to $14,000 in prizes!

Submissions close
October 13, 2020

intel

Intel® Distribution of OpenVINO™ toolkit streamlines deployment for DL workloads

Elizabeth Campbell

General manager, Americas

ADLINK Technology

Workshops and demos to visit

In-depth technical workshops

- Friday, September 18, 9:00 a.m. to 1:30 p.m. PDT: **Using the Intel® Distribution of the OpenVINO™ Toolkit for Deploying Accelerated Deep Learning Applications**
- Monday, September 21, 9:00 a.m. to 1:30 p.m. PDT: **Intel's Edge AI for Retail**
- Wednesday, September 23, 9:00 a.m. to 1:30 p.m. PDT: **Intel's Edge AI for Industrial**

Technical presentations

- Alexander Kozlov, Deep Learning R&D engineer, Intel: **Recent Advances in Post-Training Quantization**
- Dr. Manas Pathak, Global AI lead for oil and gas, Intel: **Acceleration of Deep Learning for 3D Seismic**
- Tara K. Thimmanaik, solutions architect, Intel: **Smarter Manufacturing Achieved with Intel's Deep Learning-Based Machine Vision**
- Gary Brown, Director of AI Marketing, Intel: **Getting Efficient DL Inference Performance: Is it Really All About the TOPS?**
- Rama Karamsetty, Edge AI Marketing Manager, Intel: **Edge Inferencing—Scaling w/Vision Accelerator Cards**
- Vaidyanathan Krishnamoorthy, edge inference solutions architect, Intel: **Federated Edge Inferencing**

Dedicated demos and networking space

ADLINK Leading EDGE COMPUTING · algolux · FOXCONN · IOTech · SKYLAB · ThunderSoft · XINJE

intel

Thank you

Edge computing resources

Edge developer offerings

OpenVINO™ toolkit | OpenNESS | Open Visual Cloud

- Intel® DevCloud for the Edge
- Intel® Edge Software Hub
- Intel® developer kits
- Intel® Edge AI nanodegree program with Udacity
- OpenVINO™ course on Coursera
- AI OpenVINO™ Course on China University MOOC
- OpenVINO™ + Azure ONNX RT on Azure Marketplace
- Intel® DevCloud for the Edge on Azure Marketplace

Edge computing portfolios

- Intel® Edge AI Booklet | Intel® Edge Ecosystem Booklet
- Intel® Edge Computing Overview

Edge ecosystem and offerings

- Intel® IoT Market Ready Solutions
- Intel® IoT RFP Ready Kits
- Intel® IoT Solution Aggregators
- Intel® Select Solutions
- Intel® AI: In Production | Partner case studies
- Intel® IoT Solutions Alliance
- Intel® Network Builders for the Edge

Workshops and demos to visit

General session speaker

- Bill Pearson, VP IOTG, GM Developer Enabling, Tuesday, September 15, 10:00 a.m. to 10:30 a.m. PDT: **Streamline, Simplify and Solve for the Edge of the Future**

In-depth technical workshops

- Friday, September 18, 9:00 a.m. to 1:30 p.m. PDT: **Using the Intel® Distribution of the OpenVINO™ Toolkit for Deploying Accelerated Deep Learning Applications**
- Monday, September 21, 9:00 a.m. to 1:30 p.m. PDT: **Intel's Edge AI for Retail**
- Wednesday, September 23, 9:00 a.m. to 1:30 p.m. PDT: **Intel's Edge AI for Industrial**

Technical presentations

- Alexander Kozlov, Deep Learning R&D engineer, Intel: **Recent Advances in Post-Training Quantization**
- Dr. Manas Pathak, Global AI lead for oil and gas, Intel: **Acceleration of Deep Learning for 3D Seismic**
- Tara K. Thimmanaik, solutions architect, Intel: **Smarter Manufacturing Achieved with Intel's Deep Learning-Based Machine Vision**
- Gary Brown, Director of AI Marketing, Intel: **Getting Efficient DL Inference Performance: Is it Really All About the TOPS?**
- Rama Karamsetty, Edge AI Marketing Manager, Intel: **Edge Inferencing—Scaling w/Vision Accelerator Cards**
- Vaidyanathan Krishnamoorthy, edge inference solutions architect, Intel: **Federated Edge Inferencing**

Dedicated demos and networking space

ADLINK Leading EDGE COMPUTING · algolux · FOXCONN · IOTech · SKYLAB · ThunderSoft · XINJE

intel