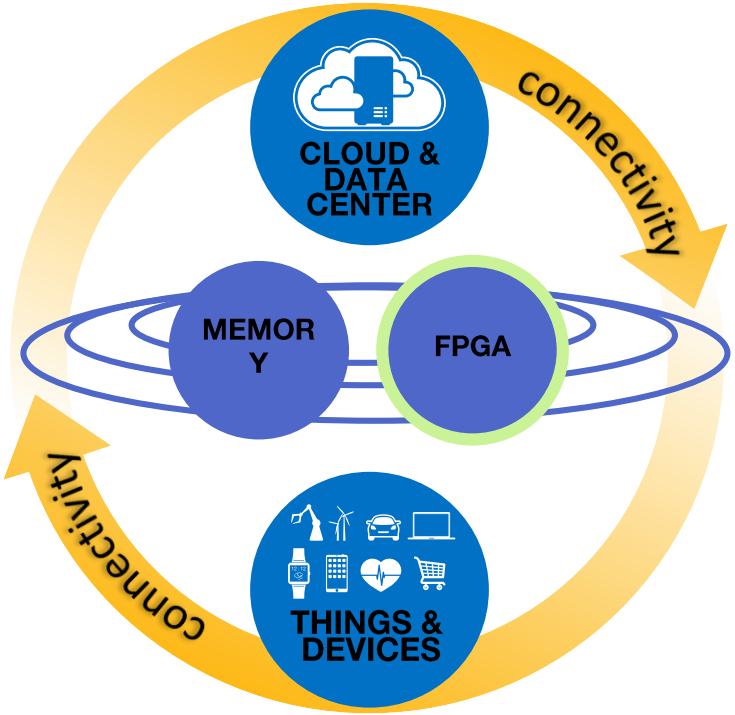
Enabling Software Developers to Harness FPGA Compute Accelerators

Bernhard Friebe
Senior Director, Product Planning
Programmable Solutions Group, Intel Corporation

May 23, 2018

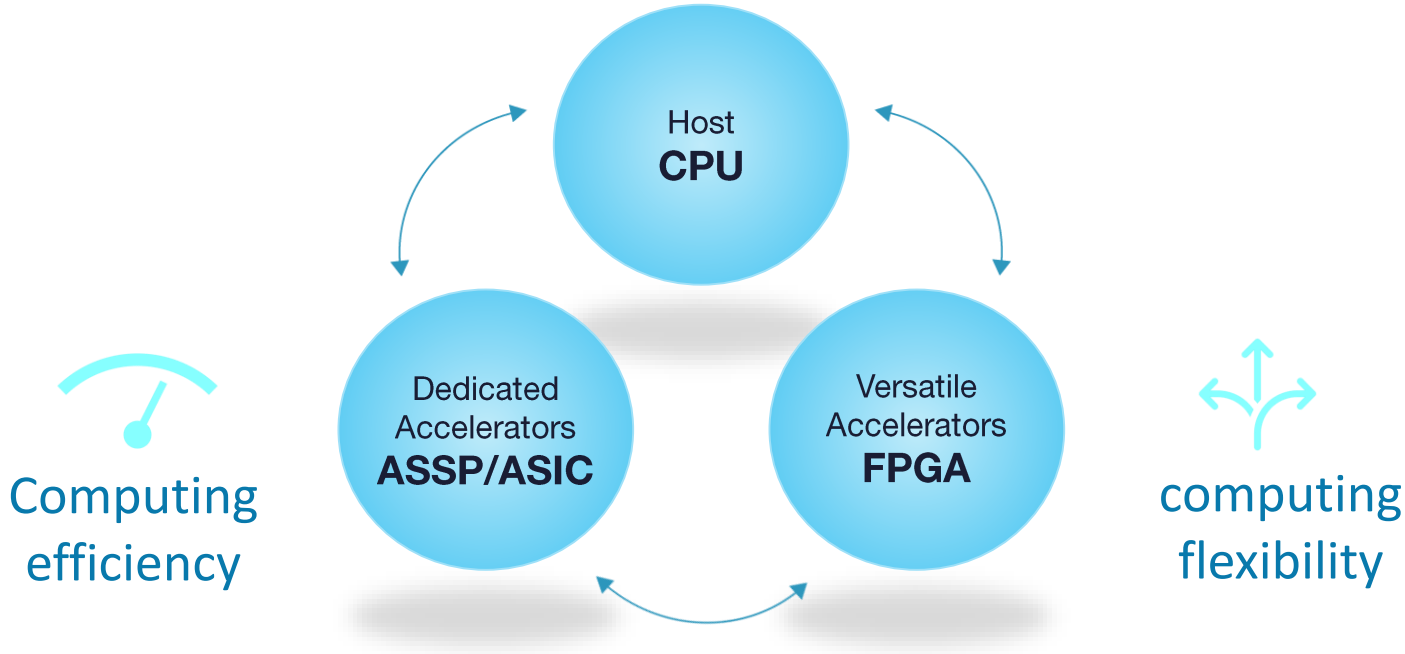
A Smart and Connected World



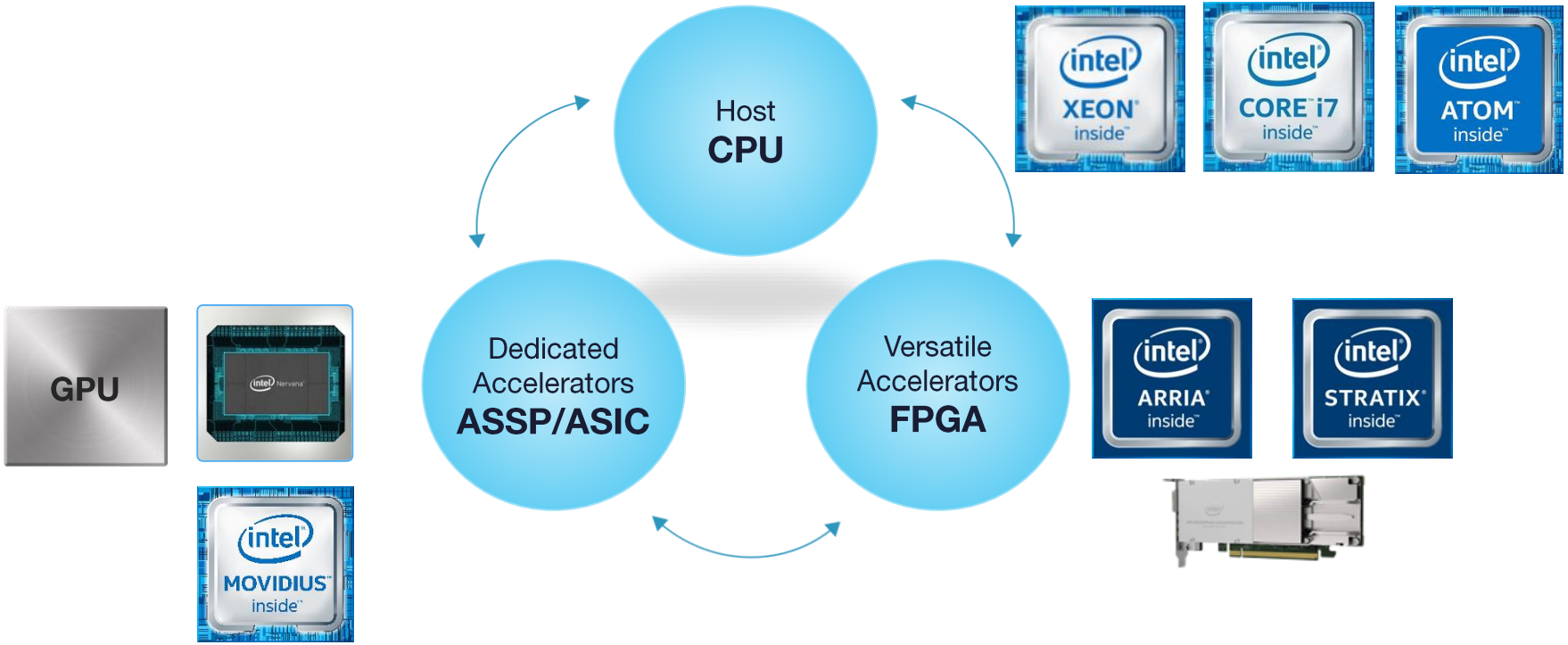
50 BILLION
CONNECTED DEVICES

2,300 EXABYTES
OF IP TRAFFIC/YEAR

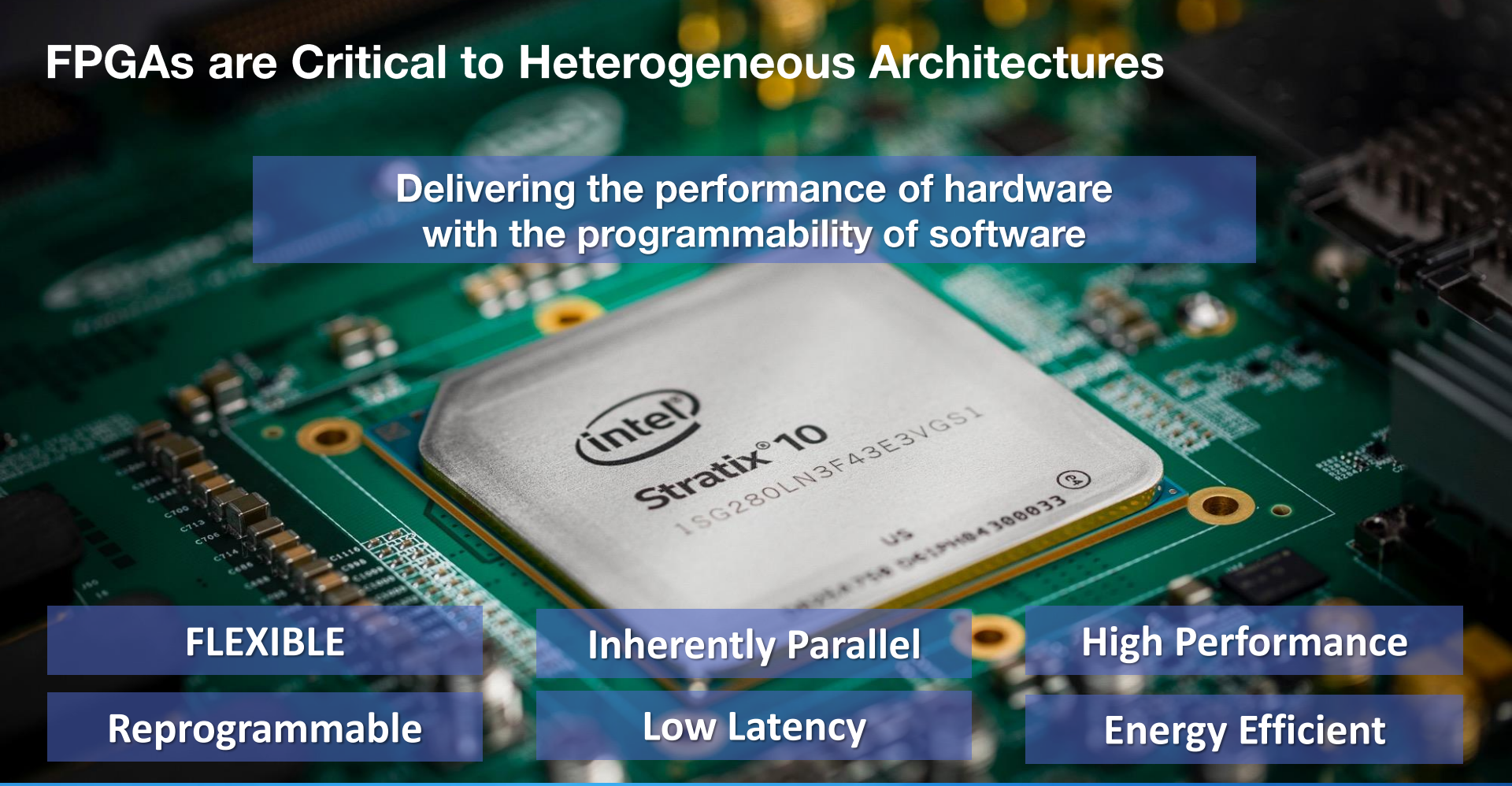
Heterogeneous Architectures are Emerging



Heterogeneous Platforms Enabled by Intel



FPGAs are Critical to Heterogeneous Architectures



Winning in Cloud: Intel FPGAs for Workload Acceleration

Microsoft bing

95%

GAIN IN THROUGHPUT

29%

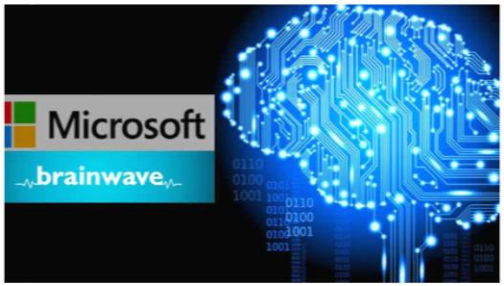
DECREASE IN LATENCY

8X

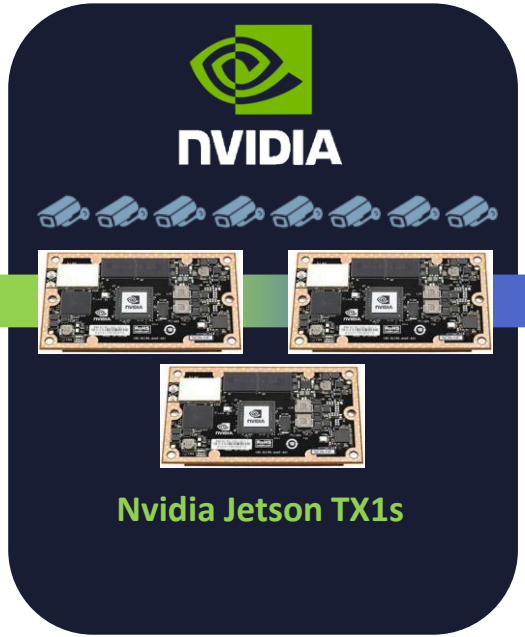
INCREASE IN SPEED WITH 15% LESS POWER

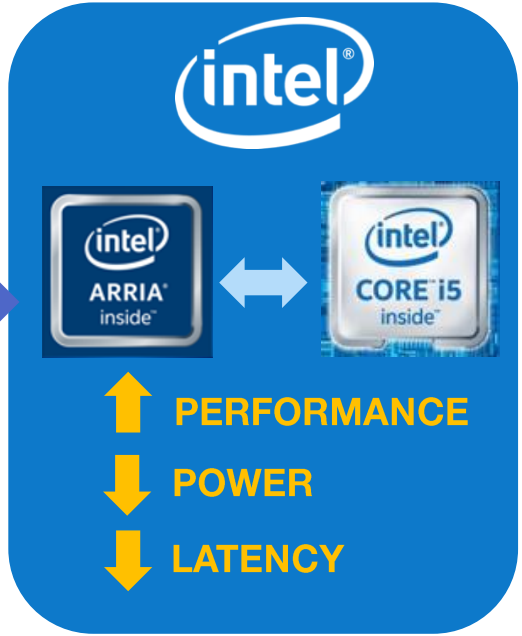
benchmarks done by Microsoft

Microsoft Azure



Success Story: Video Surveillance





Face Detection and Face Recognition

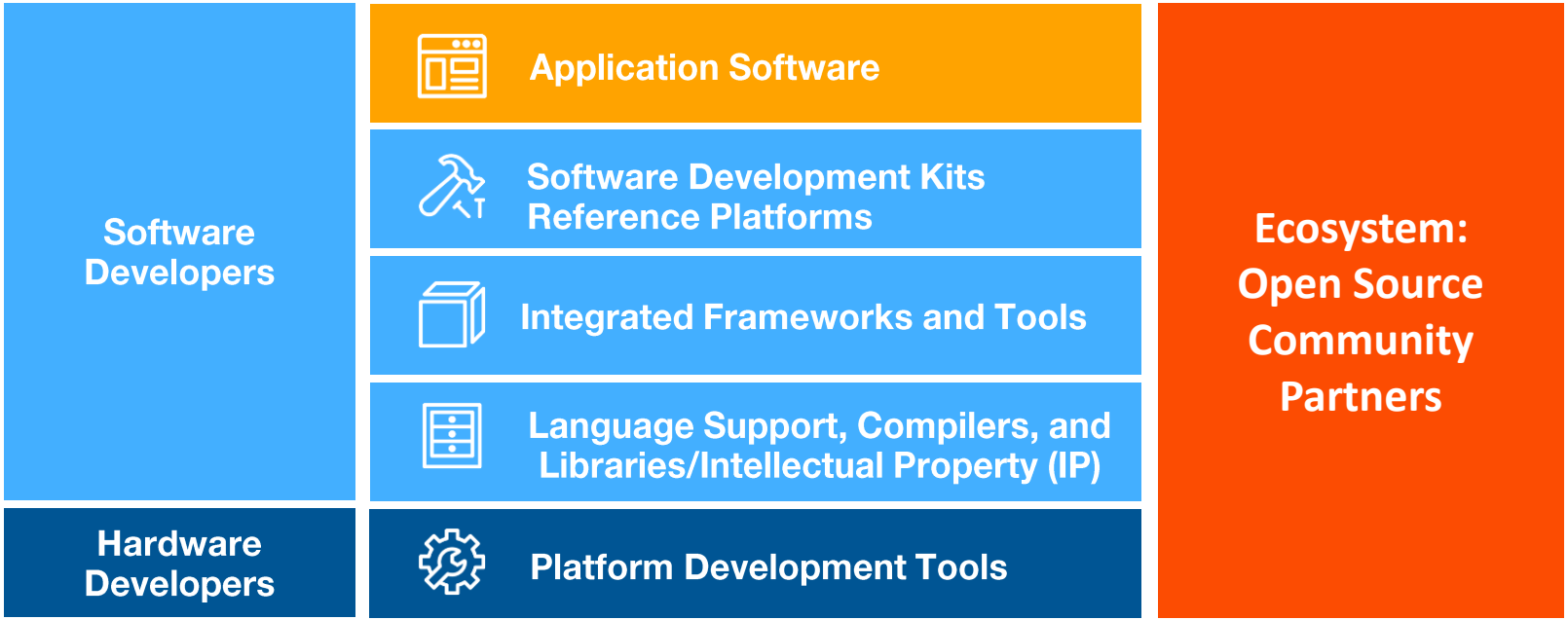
Face Attribute Classification

Software Developers are the New FPGA Developers

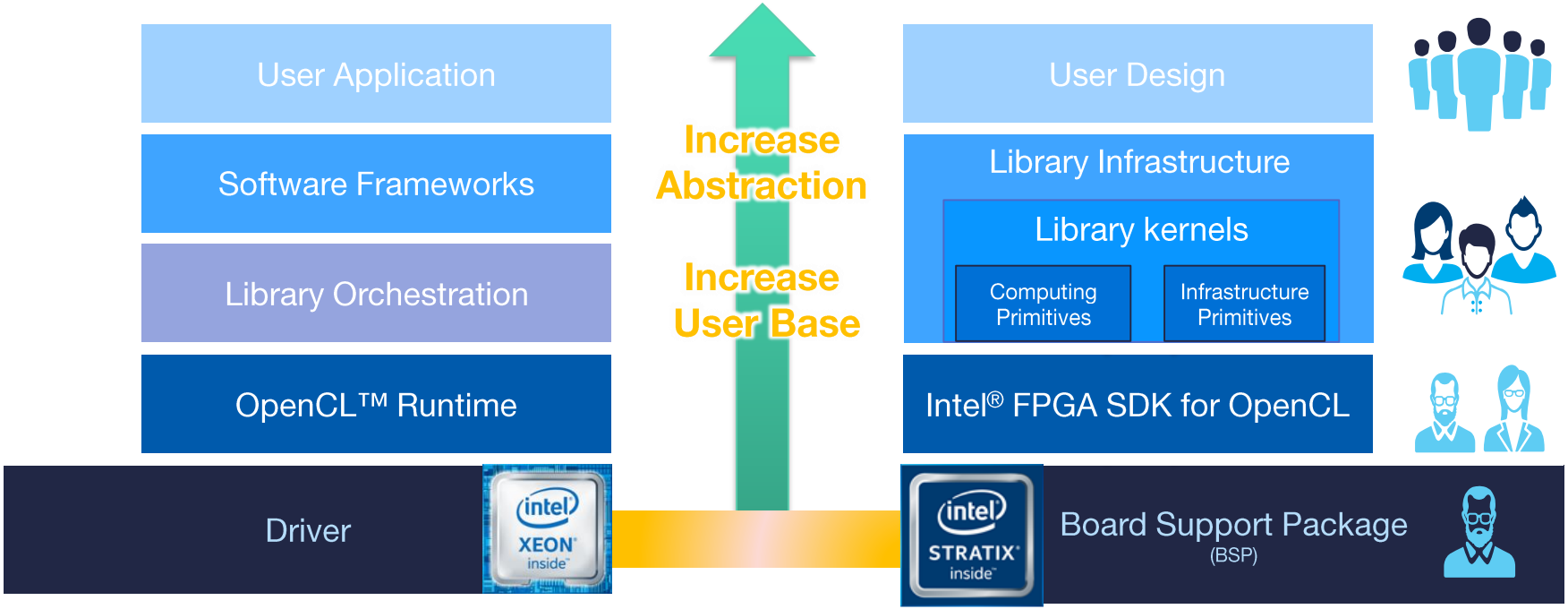
“I don’t speak FPGA!

What is the programming model,
and where are the compilers, libraries and tools I am used to?”

Empowering Developers

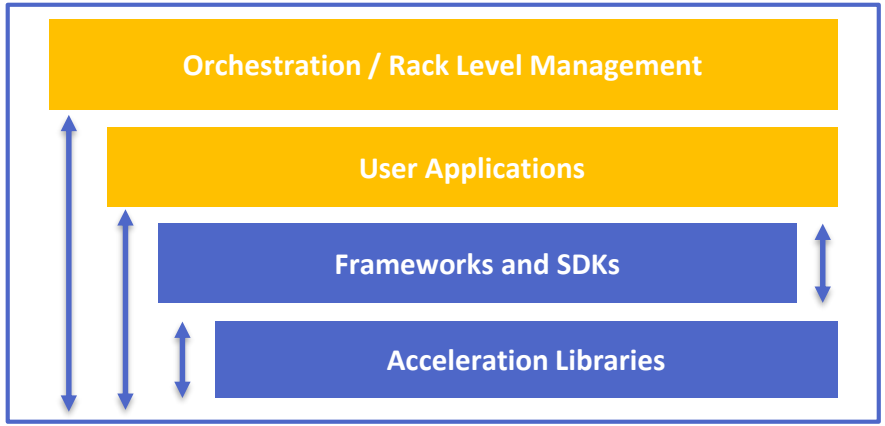


FPGA Acceleration Stack with Libraries



Acceleration Stack for Intel® Xeon® CPU with FPGAs

- Fast, hot-swapping of accelerator functions
- Accessible from virtual machines and containers
- Support for leading cloud orchestrators
- Enables the developer and ecosystem



Intel® OpenVINO™ Toolkit: Enabling Camera to Cloud

What's inside OpenVINO™

Intel Deep Learning Deployment Toolkit

Cross-platform approach
to deep learning inference

Model Optimizer
Convert and optimize
trained modules

Inference Engine
Run optimized
inferences

OpenCV

Optimized
functions for
Intel processors

Create your own customer
kernels or use a library
of functions

Optimized Libraries and OpenVX*

Runtimes, emulator,
kernels, workload samples

Enhanced, graphical
development using the
Vision Algorithm Designer

Hardware Support

GPU FPGA CPU

GPU CPU

GPU IPU CPU

Getting Started with FPGAs for Deep Learning Inferencing

Targeting

Edge

Data Center

Machine Learning Developer
Software Developer
Data Scientists

Intel® OpenVINO™

Intel Deep Learning Deployment Toolkit

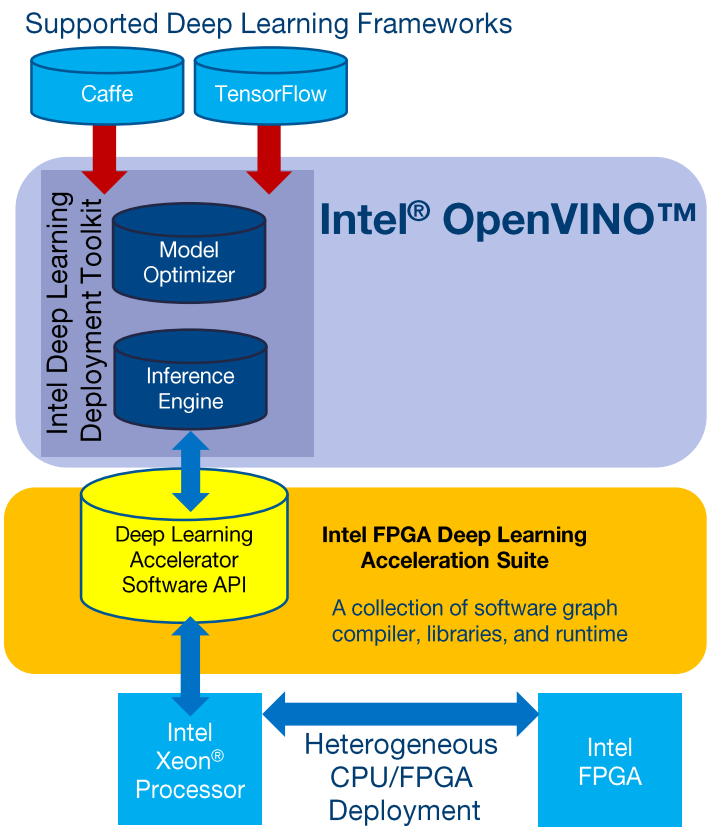
New

Intel FPGA Deep Learning Acceleration Suite

Intel Programmable Acceleration Card with Intel Arria® 10 GX FPGA

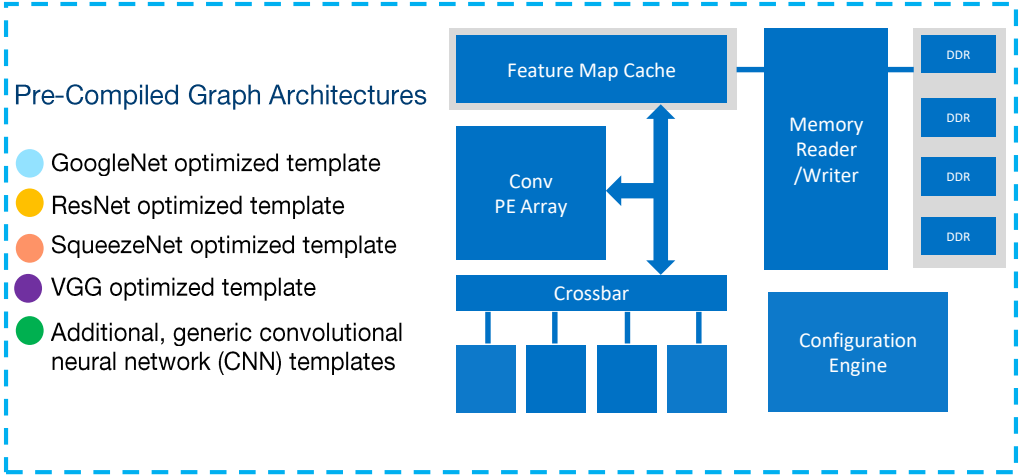
Intel FPGA Evaluation:
Intel Programmable Acceleration Card (Intel PAC) for Data Centers (production available via OEM)
IEI for Edge (coming soon)

Intel® FPGA Deep Learning Acceleration Suite



Current Supported Topologies
(more variants coming soon)

AlexNet	GoogleNet	Tiny Yolo	LeNet	SqueezeNet
VGG16	ResNet 18	...	ResNet 50	ResNet 101



Get Started Today

Heterogeneous architectures are emerging targeting flexible computing acceleration with FPGA

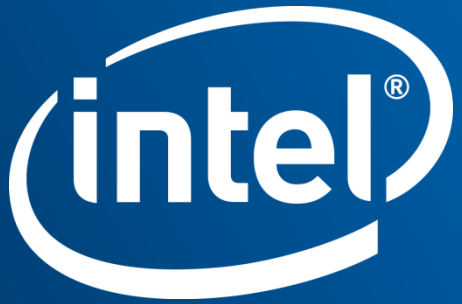
Intel® FPGAs offer Edge and Data Center workload acceleration

Intel FPGAs enabled by software-developer-friendly programming models

Meet with us at the Intel booth, Embedded Vision Summit

Find out more online at intel.com, Intel FPGA website, and Embedded Summit website

Plug and play with Intel OpenVINO™ pre-packaged with Intel FPGA Deep Learning Acceleration Suite for real-time deep learning inferencing on Intel FPGAs. Download


intel®