DELL AND INTEL AI ALIGNMENT

Gretchen Stewart,
Chief Data Scientist, Public Sector

Al Ford, Dell Technologies
Director AI Alliances

Key Take ways

- #S2MV – Joint AI strategy
- Democratizing AI – citizen data scientists
- Code Optimization and Code reuse - OpenSource
- Resources and Next steps

Federal AI Initiatives

Department of Energy – AI for Science - https://anl.app.box.com/s/bpp2xokglo8z 0

“On the server side, storage systems, such as those that support Lustre and bu t designed for and often perform poorly for these read-heavy, random access wo realization has led to pursuit of alternative memory technologies including NAND flashand 3D Xpoint (e.g.,Intel Optane), because they offer superior energy efficiency and density.”

NIST – US leadership in AI - https://www.nist.gov/system/files/documents/2019/08/1 ngagement_plan_9aug2019.pdf

Department of Defense – AI strategy - https://media.de Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRA

- USAF AI Annex to DoD st /www.af.mil/Portals/1/documents/5/USAF-AI-Annex-to-DoD-AI-Strategy.pdf
- Transformation of DoD ng up the Joint AI Center - https://www.ai.mil/

NOAA – https://nrc.noaa.go x?fileticket=0I2p2-Gu3rA%3D& id=0

“As data exploitation capabilities continue to increase exponentially with ystems and architecture, unmanned systems and commercial data sources, AI met e transformative advancements in the quality and timeliness of NOAA science, d services.”

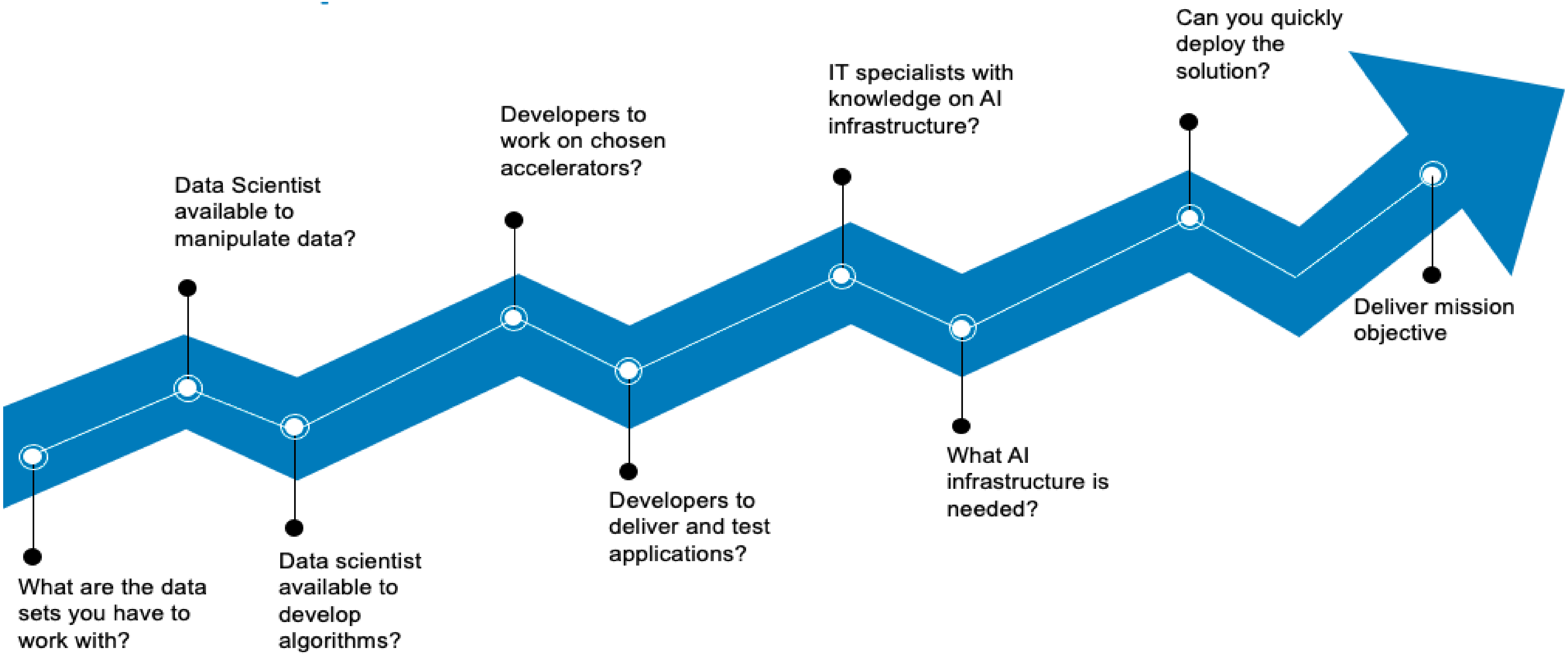
AI in Government Act 2020

WH Executive Order 2019

National AI Research Institutes - NSF

Centers of Excellence - GSA

Speed to Mission Value #S2MV



What are the data sets you have to work with?

Data Scientist available to manipulate data?

Data scientist available to develop algorithms?

Developers to work on chosen accelerators?

Developers to deliver and test applications?

IT specialists with knowledge on AI infrastructure?

What AI infrastructure is needed?

Can you quickly deploy the solution?

Deliver mission objective

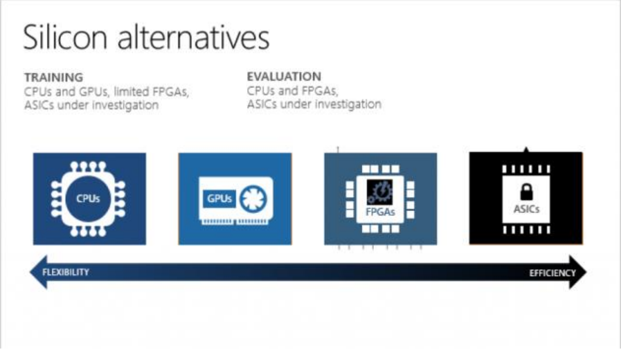
DELLEMC intel.

AI Strategy Powered by Dell and Intel

Lead with Software

Leverage Software portfolio
Optimize agency software
Evangelize to developers

Deliver the best AI Platforms



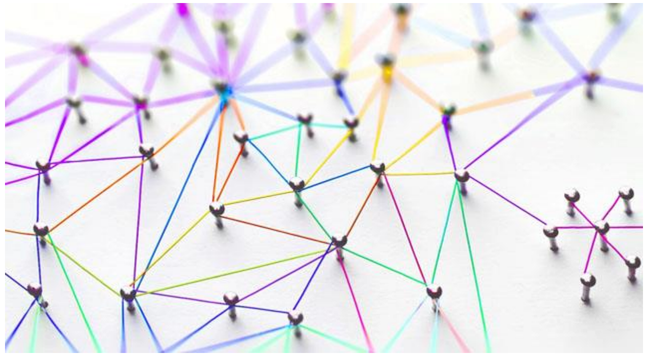
Silicon alternatives

TRAINING
CPUs and GPUs, limited FPGAs,
ASICs under investigation

EVALUATION
CPUs and FPGAs,
ASICs under investigation

Extend the CPU
Align compute and memory
Drive “Fit for Purpose” compute

Ignite the Ecosystem



Train the Team
Leverage >1000 Intel use cases
Pioneer AIoT deployments

DELLEMC intel.

The AI Buyer

1 Agency CIO / CDO

Key Concern(s): Agency mission effectiveness

Key Messaging:
- Time to results
- Alignment with Enterprise strategy

2 IT System Architect

Key Concern(s): TCO, security & manageability

Key Messaging:
- Infrastructure integration
- Create sharable platforms
- Time to results

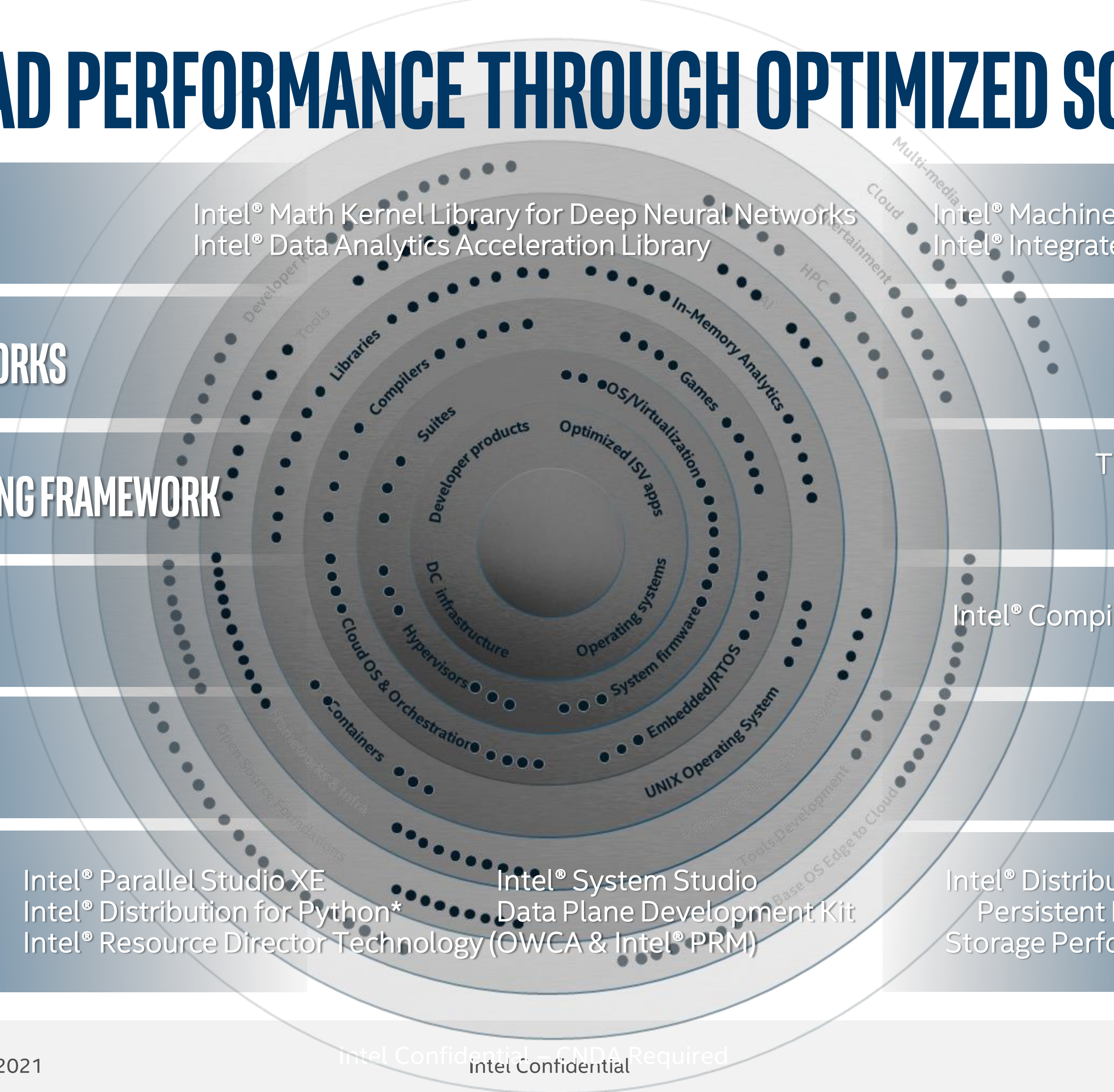
3 AI Developer, Programmer Data Scientist

Key Concern(s): Time-to-solution

Key Messaging:
- Scalability of Models
- Reduction in training times
- Development and Deployment Framework Support

DELLEMC intel

WORKLOAD PERFORMANCE THROUGH OPTIMIZED SOFTWARE



OPTIMIZED LIBRARIES	Intel® Math Kernel Library for Deep Neural Networks Intel® Data Analytics Acceleration Library	Intel® Machine Learning Scalable Library Intel® Integrated Performance Primitives
OPTIMIZED AI FRAMEWORKS		TensorFlow*, PyTorch* Caffe*, PaddlePaddle* MXNET*, …
PARALLEL PROGRAMMING FRAMEWORK		Threading Building Blocks Intel® MPI Library OpenMP*
COMPILERS		Intel® Compilers (LLVM, GCC, Fortran) nGraph
PROFILING TOOLS		Intel® VTune™ Amplifier Intel® Advisor Intel® Inspector
TOOLKITS / TOOL SUITES	Intel® Parallel Studio XE Intel® Distribution for Python* Intel® Resource Director Technology (OWCA & Intel® PRM) Intel® System Studio Data Plane Development Kit	Intel® Distribution of OpenVINO toolkit Persistent Memory Development Kit Storage Performance Development Kit

INTEL® PRM: INTEL® PLATFORM RESOURCE MANAGER

Intel Distribution for Python

software.intel.com/intel-distribution-for-python

FOR DEVELOPERS USING THE MOST POPULAR AND FASTEST-GROWING PROGRAMMING LANGUAGE FOR AI

EASY, OUT-OF-THE-BOX ACCESS TO HIGH-PERFORMANCE PYTHON

- Prebuilt, optimized for numerical computing, data analytics, HPC
- Drop-in replacement for your existing Python (no code changes required)

DRIVE PERFORMANCE WITH MULTIPLE OPTIMIZATION TECHNIQUES

- Accelerated NumPy/SciPy/Scikit-Learn with Intel Math Kernel Library (Intel MKL)
- Data analytics with pyDAAL, enhanced thread scheduling with TBB, Jupyter Notebook interface, Numba, Cython
- Scale easily with optimized MPI4Py and Jupyter notebooks

FASTER ACCESS TO LATEST OPTIMIZATIONS FOR INTEL ARCHITECTURE

- Distribution and individual optimized packages available through conda and Anaconda Cloud
- Optimizations upstreamed back to main Python trunk

ADVANCING PYTHON PERFORMANCE CLOSER TO NATIVE SPEEDS

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice. Optimization Notice

Optimized Deep Learning Frameworks and Toolkits

Gen on gen performance gains for ResNet-50 with Intel DL Boost

2S Intel Xeon Platinum 8280 Processor vs 2S Intel Xeon Platinum 8180 Processor

Intel Xeon Scalable Processor	2nd Gen Intel Xeon Scalable Processor	mxnet	PyTorch	TensorFlow	Caffe	OpenVINO
FP32	INT8 w/ Intel DL Boost	3.0x	3.7x	3.9x	4.0x	3.9x
INT8	INT8 w/ Intel DL Boost	1.8x	2.1x	1.8x	2.3x	1.9x

See Configuration Details 5 in backup.

Performance results are based on testing as of dates shown in configuration and may not reflect all publicly available security updates. No product can be absolutely secure. See configuration disclosure for details. 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. 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: http://www.intel.com/performance

Intel’s oneAPI Ecosystem

Built on Intel’s Rich Heritage of CPU Tools Expanded to XPUs

oneAPI

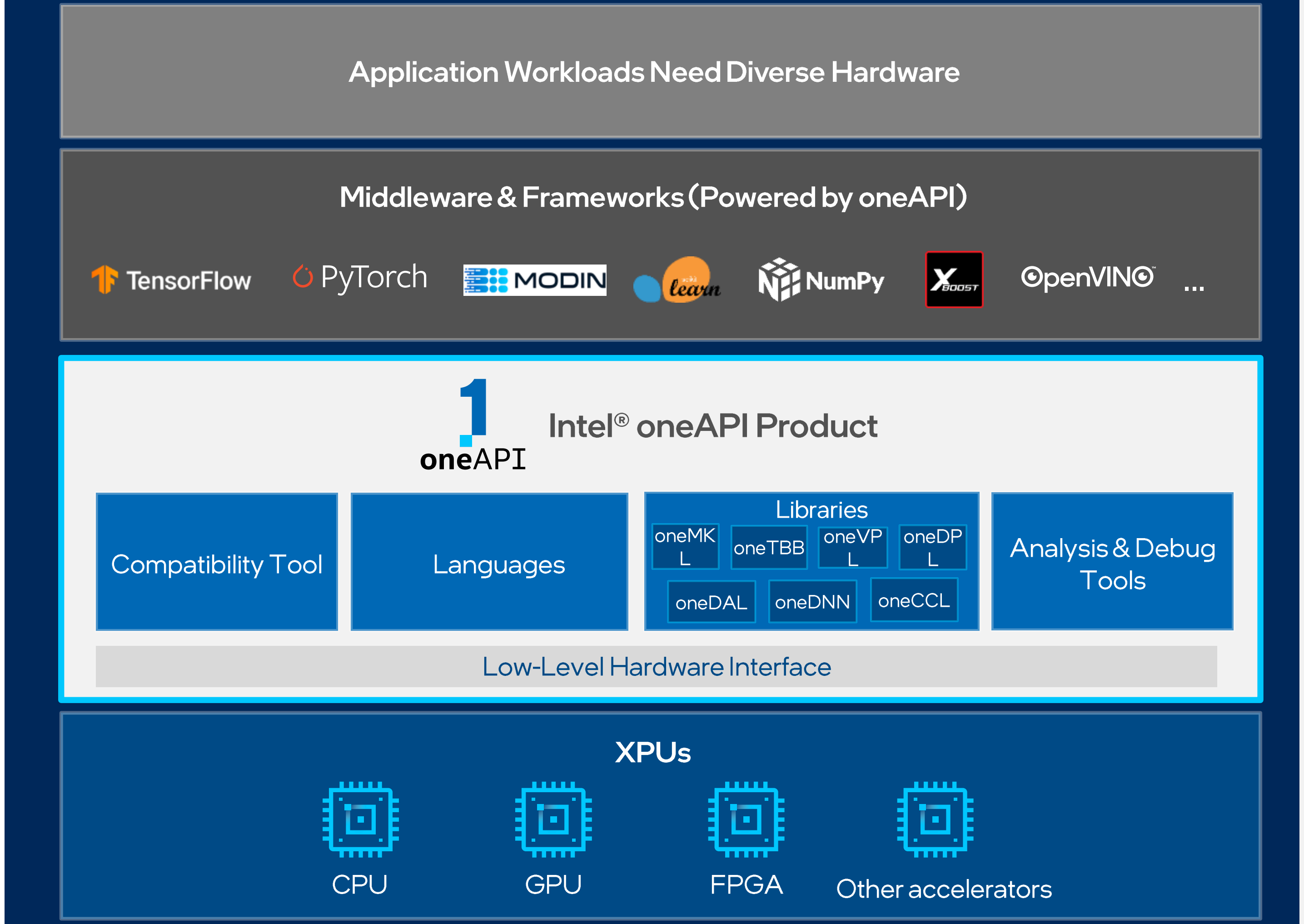
A cross-architecture language based on C++ and SYCL standards

Powerful libraries designed for acceleration of domain-specific functions

A complete set of advanced compilers, libraries, and porting, analysis and debugger tools

Powered by oneAPI

Frameworks and middleware that are built using one or more of the oneAPI industry specification elements, the DPC++ language, and libraries listed on oneapi.com.



Available Now

Visit software.intel.com/oneapi for more details
Some capabilities may differ per architecture and custom-tuning will still be required. Other accelerators to be supported in the future.

oneAPI Ecosystem Support

AI SINGAPORE Alibaba Cloud allegro.ai ANSYS Argonne NATIONAL LABORATORY @rchanan AsiaInfo 亚信科技 ATLAS EXPERIMENT AtoS BENTLEY

BittWare a molex company Brightskies technologies CANONICAL CERTI FACE CGG CHAOSGROUP CINECA CDAC CERN openlab codeplay

大势智慧 DASPATIAL DELL Technologies FOUNDRY. EURECOM Sophia Antipolis LIGO ARCHIVE GIGASPACES GeoEast Google Cloud Guerilla

HCL Hewlett Packard Enterprise URZ HEIDELBERG UNIVERSITY COMPUTING CENTRE IISER PUNE ILLINOIS Theoretical and Computational Biophysics Group inesc id lisboa TÉCNICO LISBOA KFBIO technology for health. LAIKA Laboratório Nacional de Computação Científica

Lenovo LOBACHEVSKY UNIVERSITY Microsoft Azure MEGH COMPUTING MEGWARE PHILIPS rENIAC SAP SAMSUNG MEDISON sas SENAI CIMATEC SISTEMA FIEB Federação das Indústrias do Estado da Bahia

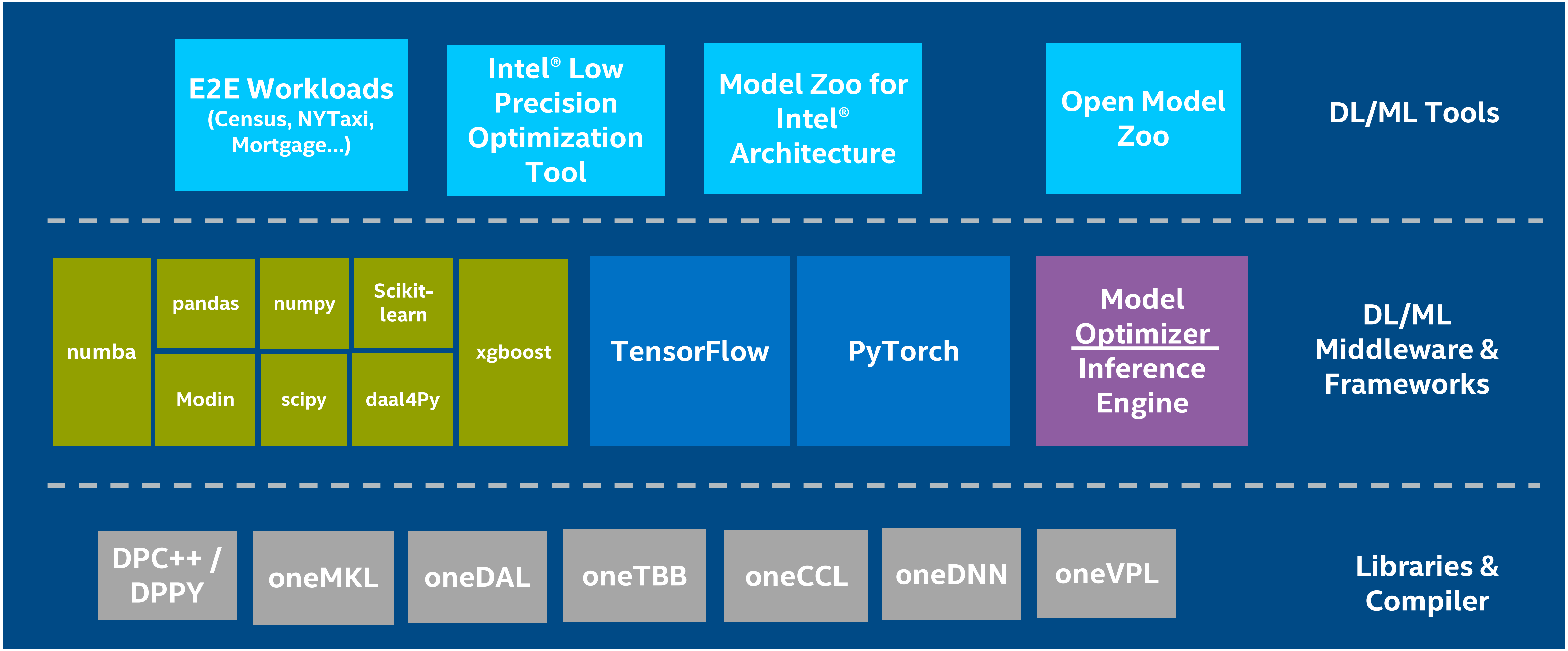
SUSE We adapt. You succeed. TANGENTSTUDIOS Taboola Tech Mahindra TensorFlow Tencent 腾讯 UNITED IMAGING 联影 ZUSE INSTITUTE BERLIN

Ben-Gurion University of the Negev Indian Institute of Technology Delhi INDIAN INSTITUTE OF SCIENCE INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR INDIAN INSTITUTE OF TECHNOLOGY ROORKEE Stockholm University KTH VETENSKAP OCH KONST UNIVERSITY OF CAMBRIDGE UNIVERSITY OF OREGON UNIVERSIDAD DE MÁLAGA

These organizations support the oneAPI initiative 'concept' for a single, unified programming model for cross-architecture development. It does not indicate any agreement to purchase or use of Intel's products.
*Other names and brands may be claimed as the property of others.

AI Software Stack for Intel® XPUs

Intel offers a Robust Software Stack to Maximize Performance of Diverse Workloads



E2E Workloads (Census, NYTaxi, Mortgage…)

Intel® Low Precision Optimization Tool

Model Zoo for Intel® Architecture

Open Model Zoo

DL/ML Tools

numba

pandas

numpy

Scikit-learn

Modin

scipy

daal4Py

xgboost

TensorFlow

PyTorch

Model Optimizer Inference Engine

DL/ML Middleware & Frameworks

DPC++ / DPPY

oneMKL

oneDAL

oneTBB

oneCCL

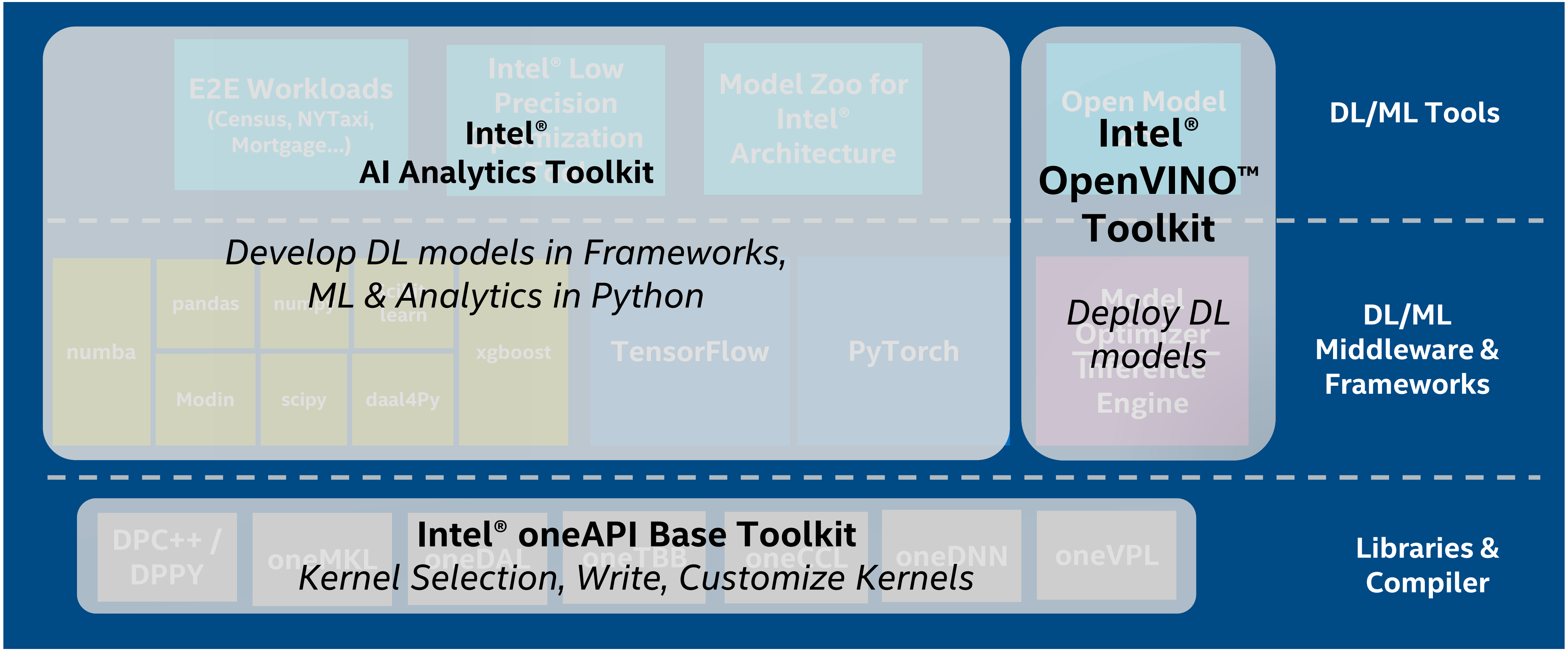
oneDNN

oneVPL

Libraries & Compiler

AI Software Stack for Intel® XPUs

Intel offers a robust software stack to maximize performance of diverse workloads



E2E Workloads (Census, NYTaxi, Mortgage...)

Intel® Low Precision Optimization Tool

Model Zoo for Intel® Architecture

Intel® AI Analytics Toolkit

Open Model Zoo

Intel® OpenVINO™ Toolkit

DL/ML Tools

Develop DL models in Frameworks, ML & Analytics in Python

numba

pandas

numpy

scikit-learn

xgboost

Modin

scipy

daal4Py

TensorFlow

PyTorch

Deploy DL models

Model Optimizer

Inference Engine

DL/ML Middleware & Frameworks

Intel® oneAPI Base Toolkit

Kernel Selection, Write, Customize Kernels

DPC++ / DPPY

oneMKL

oneDAL

oneTBB

oneCCL

oneDNN

oneVPL

Libraries & Compiler

Full Set of AI ML and DL Software Solutions Delivered with Intel's oneAPI Ecosystem

Intel® Distribution of OpenVINO™ toolkit

Powered by oneAPI

A toolkit for faster, more accurate real-world results using high-performance, AI and computer vision inference deployed into production across Intel® architecture from edge to cloud

Who needs this product?

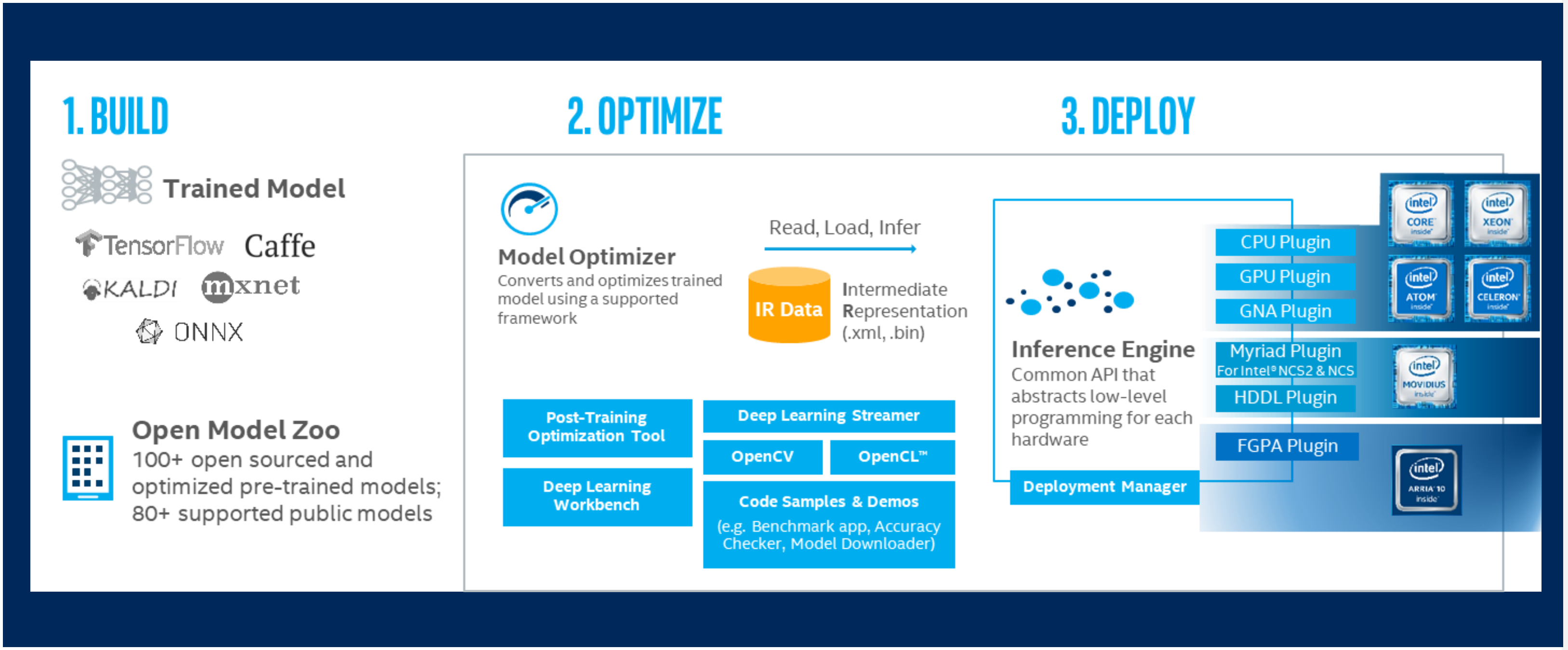
AI application developers, OEMs, ISVs, System Integrators, Vision and Media developers

Top Features/Benefits

High-performance, deep learning inference deployment

Streamlined development; ease of use

Write once, deploy anywhere



Proven, industry-leading accelerated technology

https://software.intel.com/en-us/openvino-toolkit

Getting Started with Intel® AI Analytics Toolkit

Overview

- Visit Intel® AI Analytics Toolkit (AI Kit) for more details and up-to-date product information
- Release Notes

Installation

- Download the AI Kit from Intel, Anaconda or any of your favorite package managers
- Get started quickly with the AI Kit Docker Container
- Installation Guide
- Utilize the Getting Started Guide

Hands on

- Code Samples
- Build, test and remotely run workloads on the Intel® DevCloud for free. No software downloads. No configuration steps. No installations.

Learning

- Machine Learning & Analytics Blogs at Intel Medium
- Intel AI Blog site
- Webinars and Articles at Intel® Tech Decoded

Support

- Ask questions and share information with others through the Community Forum
- Discuss with experts at AI Frameworks Forum

Download Now

Which Toolkit to Use When ?

	Intel® AI Analytics Toolkit	OpenVINO™ Toolkit
Key Value Prop	• Provide performance and easy integration across end-to-end data science pipeline for efficient AI model development • Maximum compatibility with opensource FWKs and Libs with drop-in acceleration that require minimal to no code changes • Audience: Data Scientists; AI Researchers; DL/ML Developers	• Provide leading performance and efficiency for DL inference solutions to deploy across any Intel HW (cloud to edge). • Optimized package size for deployment based on memory requirements • Audience: AI Application Developers; Media and Vision Developers
Use Cases	• Data Ingestion, Data pre-processing, ETL operations • Model training and inference • Scaling to multi-core / multi-nodes / clusters	• Inference apps for vision, Speech, Text, NLP • Media streaming / encode, decode • Scale across HW architectures – edge, cloud, datacenter, device
HW Support	• CPUs - Datacenter and Server segments – Xeons, Workstations • GPU - ATS and PVC (in future)	• CPU - Xeons, Client CPUs and Atom processors • GPU - Gen Graphics; DG1 (current), ATS, PVC (in future) • VPU - NCS & Vision Accelerator Design Products, • FPGA • GNA
Low Precision Support	**Use Intel® Low Precision Optimization Tool when using AI Analytics Toolkit** • Supports BF16 for training and FP16, Int8 and BF16 for Inference • Seamlessly integrates with Intel optimized frameworks • Available in the AI toolkit and independently	**Use Post Training Optimization Tool when using OpenVINO** • Supports FP16, Int8 and BF16 for inference • Directly works with Intermediate Representation Format • Available in the Intel Distribution of OpenVINO toolkit • Provides Training extension via NNCF for PyTorch with FP16, Int8

Exception: If a model is not supported by OpenVINO™ toolkit for Inference deployment, build custom layers for OV or fall back to the AI Analytics Toolkit and use optimized DL frameworks for inference.

Use both !

Toolkits are complimentary to each other and recommendation is to use them both based on your current phase of AI Journey

- *I am **exploring and analyzing data;** I am **developing models***
- *I want **performance and compatibility** with frameworks and libraries I use*
- *I would like to have **drop-in acceleration** with little to no additional code changes*
- *I prefer **not to learn any new tools** or languages*

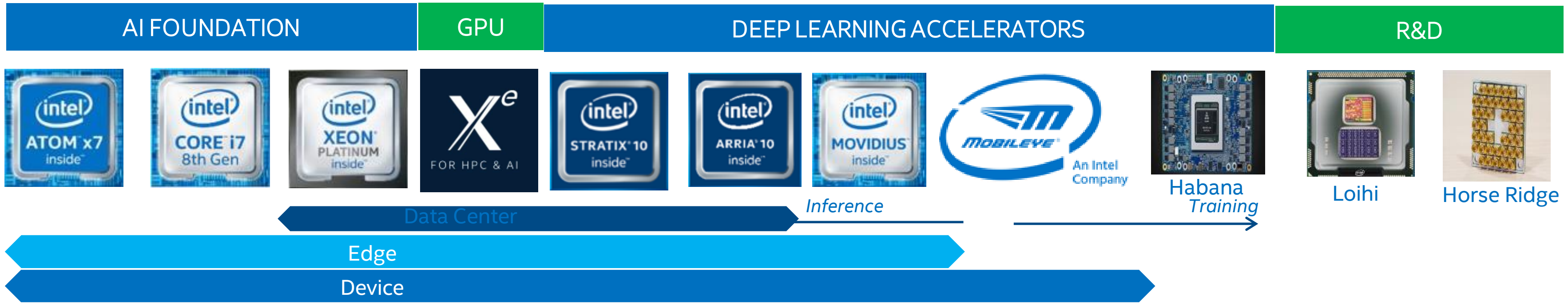
Data Scientist/ML Developer
Intel® AI Analytics Toolkit

- *I am **deploying models***
- *I want **leading performance and efficiency** across multiple target HW*
- *I’m concerned about **having lower memory footprint**, which is critical for deployment*
- *I am **comfortable with learning and adopting a new tool or API** to do so*

App Developer
Intel® Distribution of OpenVINO™ toolkit

If you prefer working on primitives and optimize kernels and algorithms directly using oneAPI libraries (oneDNN, oneCCL & oneDAL), then use Intel® Base Toolkit

Fit for Purpose AI Compute


AI FOUNDATION

GPU

DEEP LEARNING ACCELERATORS

R&D

Habana
Training

Loihi

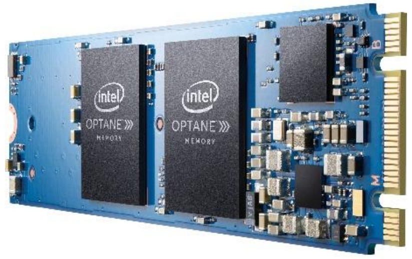
Horse Ridge

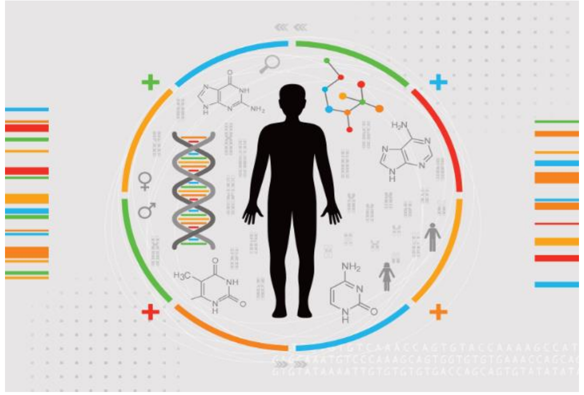
Data Center

Inference

Edge

Device





Align Compute and Memory

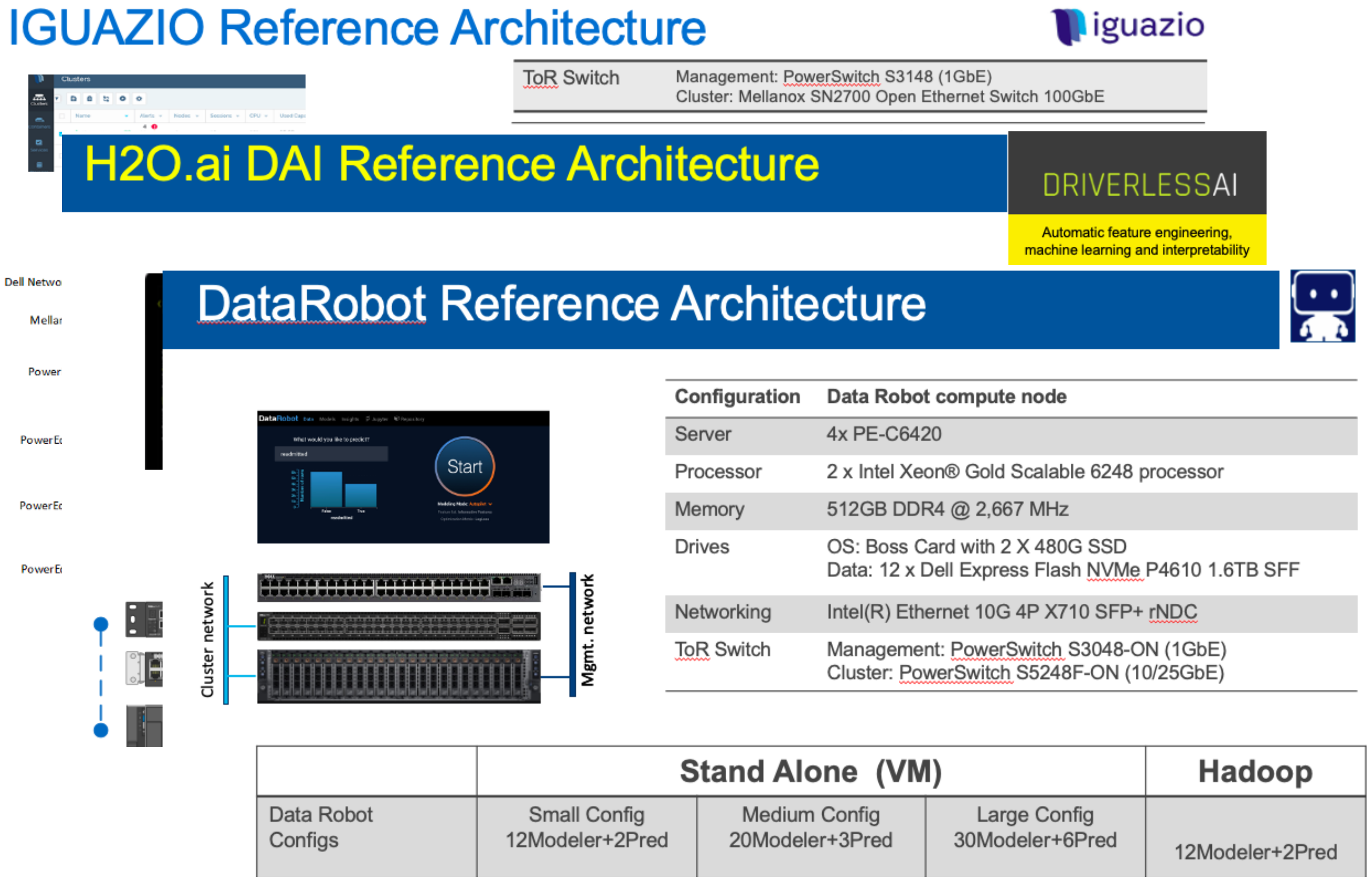
Hard Problem: Traditional AI architectures don’t address the challenge of accessing large and small data sets and the data latency from intensive ETL (Extract, Transform, Load) processes

+

- Supercharges AI Application Performance
- Increases Speed 2 Mission Value

Key Trend: Application developers want to run AI applications in persistent memory to eliminate bottlenecks and accelerate performance

Software Driven Market Ready Solutions



iguazio
Clusters
Name
Alerts
Nodes
Sessions
CPU
Used Capa
Dell Netwo
Mellar
Power
PowerEc
PowerEc
PowerEc
DataRobot
What would you like to predict?
readmitted
Start
Cluster network
Mgmt. network

IGUAZIO Reference Architecture

ToR Switch	Management: PowerSwitch S3148 (1GbE) Cluster: Mellanox SN2700 Open Ethernet Switch 100GbE

H2O.ai DAI Reference Architecture

DRIVERLESSAI

Automatic feature engineering, machine learning and interpretability

DataRobot Reference Architecture

Configuration	Data Robot compute node
Server	4x PE-C6420
Processor	2 x Intel Xeon® Gold Scalable 6248 processor
Memory	512GB DDR4 @ 2,667 MHz
Drives	OS: Boss Card with 2 X 480G SSD Data: 12 x Dell Express Flash NVMe P4610 1.6TB SFF
Networking	Intel(R) Ethernet 10G 4P X710 SFP+ rNDC
ToR Switch	Management: PowerSwitch S3048-ON (1GbE) Cluster: PowerSwitch S5248F-ON (10/25GbE)

	Stand Alone (VM)			Hadoop
Data Robot Configs	Small Config 12Modeler+2Pred	Medium Config 20Modeler+3Pred	Large Config 30Modeler+6Pred	12Modeler+2Pred

Demonstrating Customer Obsession

- 1x software ~10x hardware pull thru
- Reference Arch accelerates AI Adoption
- Software optimized for AI performance

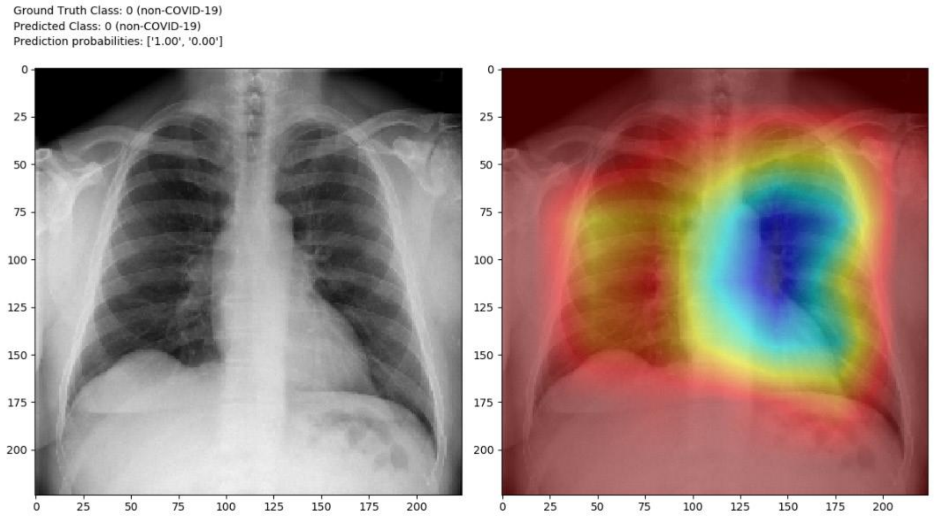
DELLEMC intel.

Case Study - Accrad

Better together – using both toolkits

CheXRad is a machine learning at the edge application that helps Radiologists and Physicians to identify COVID-19, viral pneumonia and other diseases on chest X-ray images and predict the need for ventilators

- CheXRad comes pre-configured with a COVID-19 and viral pneumonia classification neural network
- To architect, train and validate the neural network, Accrad used Intel Tensorflow from AI Analytics Toolkit and the infrastructure provided by Intel oneAPI DevCloud and developed the model.
- To further optimize its model for deployment, they used OpenVINO Toolkit and Intel DevCloud for Edge
- CheXRad could label pathologies in 140 chest x-rays in just 90 seconds—up to **160x faster** than radiologists, at comparable levels of accuracy, sensitivity and specificity.


Ground Truth Class: 0 (non-COVID-19)
Predicted Class: 0 (non-COVID-19)
Prediction probabilities: ['1.00', '0.00']
0
25
50
75
100
125
150
175
200
0
25
50
75
100
125
150
175
200
0
25
50
75
100
125
150
175
200
0
25
50
75
100
125
150
175
200

*"With the help of Intel, we were able to **train, optimize, and deploy** a machine learning model in **less time and at a lower operational cost** than available alternatives, enabling us to get to market fast with a powerful solution that's optimized for Intel® architecture." – Moloti Nakampe, R&D Director*

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Case Study - AbbVie

Better together – using both toolkits

Drop-in acceleration

- AbbVie is a research-based biopharmaceutical company using Xeons
- Abbelfish Machine Translation uses Intel® Optimization for TensorFlow of AI Analytics Toolkit
  - Custom model was used to provide more accurate translations than commercially available ones. Model includes 24 layers and over 500 million parameters, which took over four months to train;
  - Intel TF provided greater performance boost while the customer did not have to change their code / api’s from standard TF
- AbbVie Search uses Intel® Distribution of OpenVINO™ toolkit
  - OpenVINO provided great speed up to answer questions from a scientific article or clinical report when compared to standard TF.
  - Requires scaling across the company, so uses OpenVINO Model Server to serve inferences

Deployment

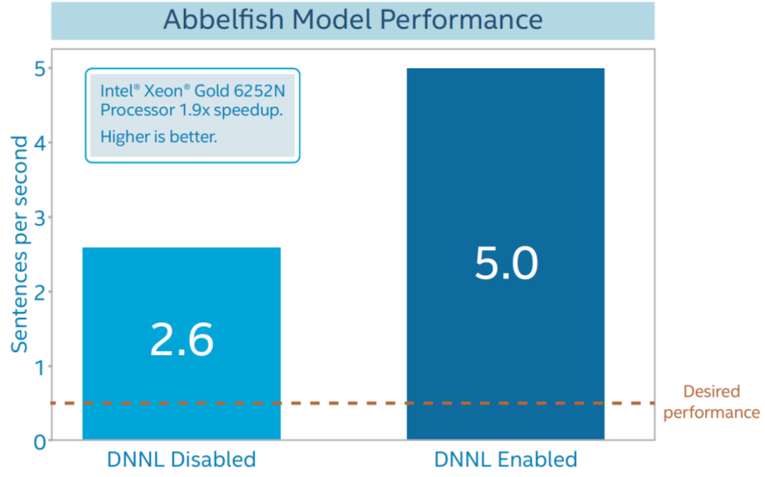


Figure 3. AbbVie's Abbelfish translated over five sentences per second using Intel Optimization for TensorFlow with oneAPI Deep Neural Network Library (oneDNN).[1]

AbbVie Search

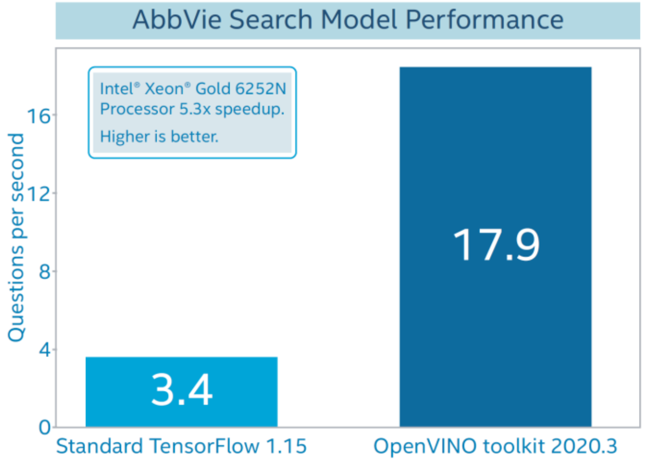


Figure 5. Comparison of AbbVie Search inference between unoptimized TensorFlow 1.15 (oneDNN disabled) and OpenVINO toolkit 2020.3.[1]

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel Neural Compute Stick 2

USB STICK FORM FACTOR
for neural network acceleration

REAL-TIME ON-DEVICE INFERENCE
no cloud connectivity required

NO ADDITIONAL PERIPHERALS
needed to start deploying solutions

OpenVINO

ACCELERATE DEVELOPMENT
with Intel Distribution of OpenVINO toolkit

INDUSTRY-LEADING PERFORMANCE[1]
with Intel Movidius Myriad™ X Vision Processing Unit (VPU)

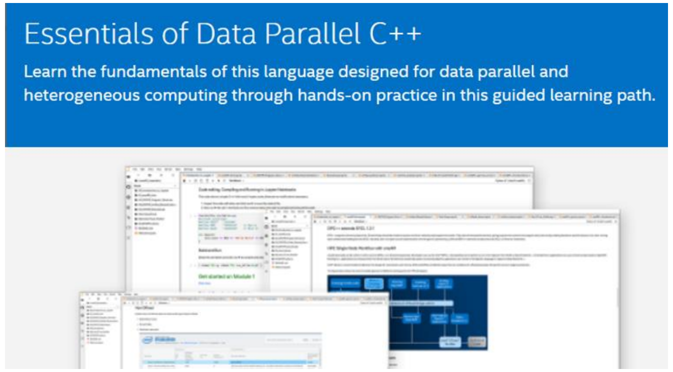
↑ 8X

HIGHER PERFORMANCE[1]
on deep neural networks compared to Intel Movidius Myriad 2 VPU

[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 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

Ecosystem Adoption & Support

Training



Essentials of Data Parallel C++

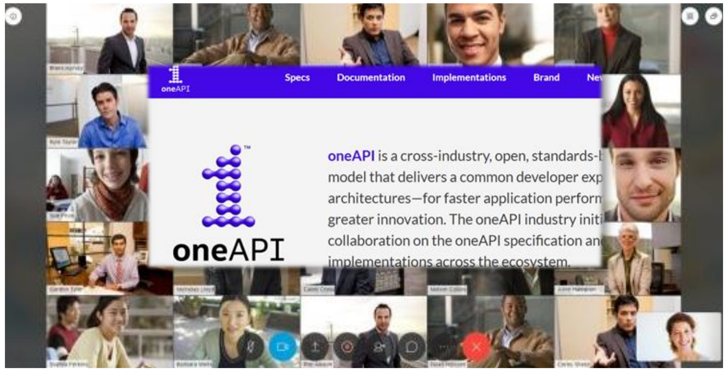
Learn the fundamentals of this language designed for data parallel and heterogeneous computing through hands-on practice in this guided learning path.

Online webinars & courses, developer guides, sample code

Academia

oneAPI Centers of Excellence: research, enabling code, curriculum, teaching

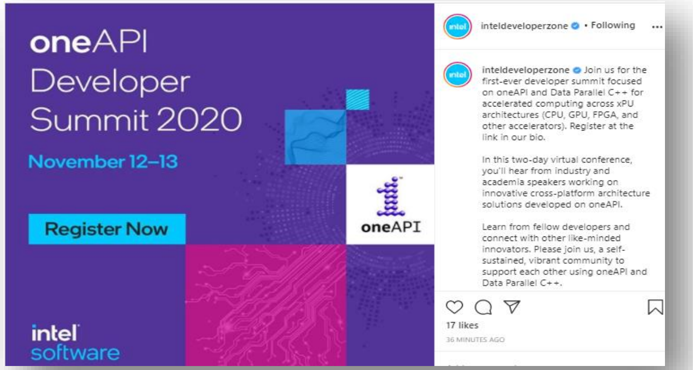
Community



oneAPI is a cross-industry, open, standards-
model that delivers a common developer exp
architectures—for faster application perform
greater innovation. The oneAPI industry initi
collaboration on the oneAPI specification an
implementations across the ecosystem.

oneAPI open specification, DevMesh innovators, community support forums

Summits & Workshops



oneAPI Developer Summit 2020

November 12–13

Register Now

oneAPI

intel software

inteldeveloperzone • Following

inteldeveloperzone Join us for the first-ever developer summit focused on oneAPI and Data Parallel C++ for accelerated computing across xPU architectures (CPU, GPU, FPGA, and other accelerators). Register at the link in our bio.

In this two-day virtual conference, you'll hear from industry and academia speakers working on innovative cross-platform architecture solutions developed on oneAPI.

Learn from fellow developers and connect with other like-minded innovators. Please join us, a self-sustained, vibrant community to support each other using oneAPI and Data Parallel C++.

17 likes

36 MINUTES AGO

Live & on-demand virtual workshops, community-led sessions

Industry Experts

intel

CERTIFIED INSTRUCTOR

AS A CERTIFIED INSTRUCTOR FOR

DATA PARALLEL C++ ESSENTIALS

oneAPI

JEFFREY MCVEIGH
VICE PRESIDENT & GENERAL MANAGER, DATACENTER XPU PRODUCTS & SOLUTIONS

MEGHANA RAO
ONEAPI CERTIFIED INSTRUCTOR PROGRAM MANAGER

Training by leading technical training companies worldwide

Intel® DevCloud

A DEVELOPMENT SANDBOX FOR DATA CENTER TO EDGE WORKLOADS

Develop, test, and run your workloads on a cluster of the latest Intel® hardware and software. With integrated Intel® optimized frameworks, tools, and libraries, you'll have everything you need for your projects.

Try Out a Diverse Collection of Intel® Hardware

Expand your skills and experiment with this state-of-the-art cluster that offers capabilities such as natural language processing and time-series analysis, as well as edge acceleration hardware.

State-of-the-art software and hardware

Intel® oneAPI Toolkits + latest Intel® Xeon®processors, GPUs (integrated & discrete), FPGAs

Key Takeaways & Call to Action

$3-4 SW generate $10 in HW

- ➢ Create 3 – 5 proof of concepts. Ready to partner and we can aid with customer analytics meetings.
- ➢ The toolkits and software are optimized, and we can help with development and deployment to achieve performance and efficiency across different stages of AI Journey
- ➢ Recommend the toolkits based on current phase of customer pipeline

Your customers can download the toolkits for free. Intel partnered with DELL we can deliver workshops, coding sessions and POC's/

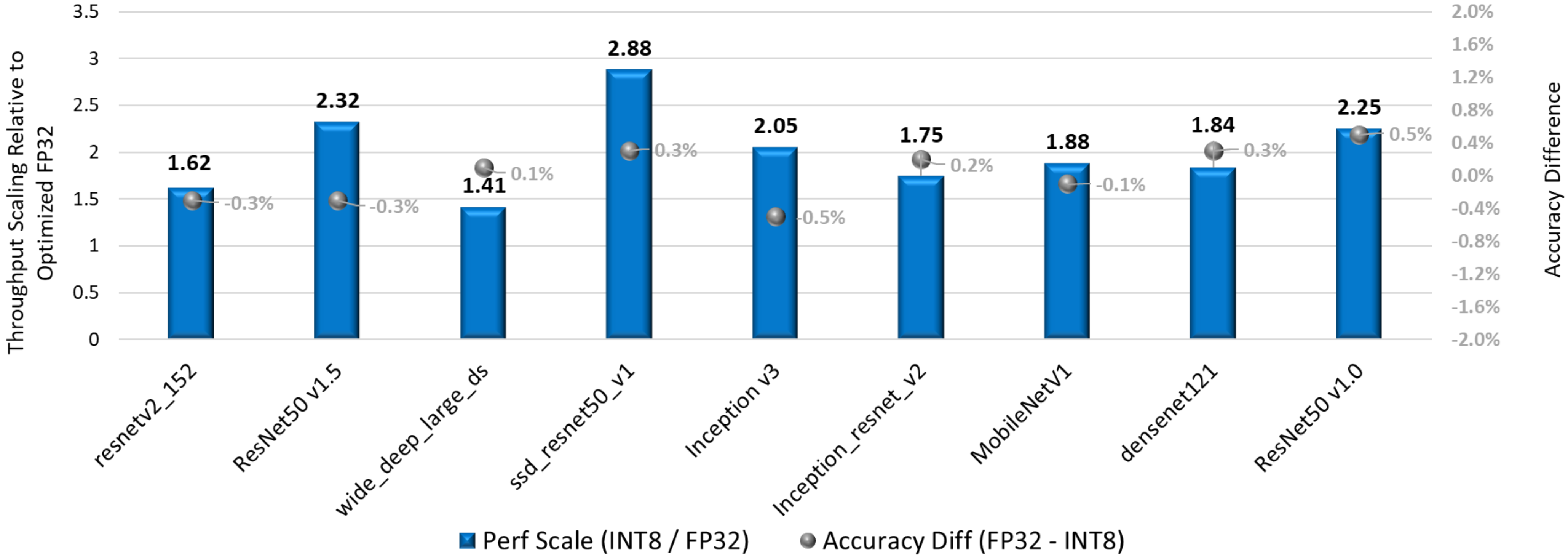
Intel® AI Analytics Toolkit

Intel® Distribution of OpenVINO ™ toolkit

Intel® oneAPI Base Toolkit

INT8 Quantized Inference Performance

Uses Intel® Optimization for Tensorflow and Intel® Low Precision Optimization Tool



INT8 Inference Throughput Scaling up to 3.8x and Accuracy Drop within 0.6%

For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.
See backup for configuration details.