oneAPI:
A Single Programming Model to Deliver Cross-Architecture Performance

Kent Moffat, oneAPI Product Manager, Intel Corporation
oneAPI Concept and Industry Initiative
oneAPI Tools and Toolkits Overview
Questions and Answers
**Programming Challenges for Multiple Architectures**

Growth in specialized workloads

No common programming language or APIs

Inconsistent tool support across platforms

Each platform requires unique software investment

Diverse set of data-centric hardware required

---

<table>
<thead>
<tr>
<th>XPUs</th>
<th>Middleware / Frameworks</th>
<th>Language &amp; Libraries</th>
<th>Other Accel.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>DP1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>GPU</td>
<td>DP2</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>FPGA</td>
<td>DP3</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Other Accel.</td>
<td>DP4</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCING oneAPI

Unified programming model to simplify development across diverse architectures

Unified and simplified language and libraries for expressing parallelism

Uncompromised native high-level language performance

Based on industry standards and open specifications

Interoperable with existing HPC programming models
A standards based cross-architecture language, DPC++, based on C++ and SYCL

Powerful APIs designed for acceleration of key domain-specific functions

Low-level hardware interface to provide a hardware abstraction layer to vendors

Open standard to promote community and industry support

Enables code reuse across architectures and vendors

Visit oneapi.com for more details

Some capabilities may differ per architecture and custom-tuning will still be required.
DATA PARALLEL C++
STANDARDS-BASED, CROSS-ARCHITECTURE LANGUAGE

Language to deliver uncompromised parallel programming productivity and performance across CPUs and accelerators
  Allows code reuse across hardware targets, while permitting custom tuning for a specific accelerator
  Open, cross-industry alternative to single architecture proprietary language

Based on C++
  Delivers C++ productivity benefits, using common and familiar C and C++ constructs
  Incorporates SYCL* from the Khronos* Group to support data parallelism and heterogeneous programming

Community Project to drive language enhancements
  Extensions to simplify data parallel programming
  Open and cooperative development for continued evolution

Builds upon Intel’s years of experience in architecture and compilers

Optimization Notice
Copyright © 2019, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.

Custom-tuning for each architecture will still be required.
POWERSFUL APIs
LIBRARIES

Designed for acceleration of key domain-focused functions

Each can be custom-coded for any platform to deliver uncompromised performance

API-based Programming: Libraries

Math
Threading
DPC++ Library

Analytics/ML
DNN
ML Comm

Video Processing

Custom-tuning for each architecture will still be required.
Distributed through a core toolkit and a complementary set of add-on domain-specific toolkits

Includes DPC++ compatibility tool for code migration along with advanced performance analysis and debug tools

**Beta Available Now**

Some capabilities may differ per architecture and custom-tuning will still be required.
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.
TOOLKITS TAILORED TO YOUR NEEDS
Domain-specific sets of tools to get your job done quickly.

**Intel® oneAPI Base Toolkit**
A core set of high-performance tools for building Data Parallel C++ applications and oneAPI library based applications

*Learn More*

**Intel® oneAPI HPC Toolkit**
Everything HPC developers need to deliver fast C++, Fortran, & OpenMP® applications that scale

*Learn More*

**Intel® oneAPI IoT Toolkit**
Tools for building high-performing, efficient, reliable solutions that run at the network's edge

*Learn More*

**Intel® oneAPI DL Framework Developer Toolkit**
Tools for developers & researchers who build deep learning frameworks or customize existing ones so applications run faster

*Learn More*

**Intel® oneAPI Rendering Toolkit**
Powerful rendering libraries to create high-performance, high-fidelity visualization applications

*Learn More*

**Intel® System Bring-Up Toolkit**
Tools to debug & tune power & performance in pre- & post-silicon development

*Learn More*

**Intel® Distribution of OpenVINO™ Toolkit**
Tools to build high performance deep learning inference & computer vision applications (production-level tool)

*Learn More*

**Intel® AI Analytics Toolkit**
Tools to build applications that leverage machine learning & deep learning models

*Learn More*
DETAILS ABOUT INTEL® oneAPI TOOLKITS (BETA)

INTEL® oneAPI BASE TOOLKIT
Core set of frequently used tools and libraries for developing high-performance applications across diverse architectures—CPU, GPU, FPGA

Who Uses It?
A broad range of developers across industries
Add-on toolkit users since this is the base for all toolkits

Top Features/Benefits
Data Parallel C++ compiler, library, and analysis tools
DPC++ Compatibility tool helps migrate existing code written in CUDA*
Python distribution includes accelerated scikit-learn, NumPy, SciPy libraries
Optimized performance libraries for threading, math, data analytics, deep learning, and video/image/signal processing

Intel® oneAPI Base Toolkit

DIRECT PROGRAMMING
- Intel® oneAPI DPC++ Compiler
- Intel® DPC++ Compatibility Tool
- Intel® Distribution for Python*
- Intel® FPGA Add-on for oneAPI Base Toolkit

API-BASED PROGRAMMING
- Intel® oneAPI DPC++ Library
- Intel® oneAPI Math Kernel Library
- Intel® oneAPI Data Analytics Library
- Intel® oneAPI Threading Building Blocks
- Intel® oneAPI Video Processing Library
- Intel® oneAPI Collective Comms. Library
- Intel® oneAPI Deep Neural Network Library
- Intel® Integrated Performance Primitives

ANALYSIS TOOLS
- Intel® VTune™ Profiler
- Intel® Advisor
- GDB*
Compiler to deliver uncompromised parallel programming productivity and performance across CPUs and accelerators

- Allows code reuse across hardware targets, while permitting custom tuning for a specific accelerator
- Open, cross-industry alternative to single architecture proprietary language

**DPC++ is based on C++ and SYCL**

- Delivers C++ productivity benefits, using common and familiar C and C++ constructs
- Incorporates SYCL* from the Khronos* Group to support data parallelism and heterogeneous programming

**Builds upon Intel’s decades of experience in architecture and high performance compilers**

*Other names and brands may be claimed as the property of others.*
INTEL® DPC++ COMPATIBILITY TOOL (BETA)
MINIMIZES CODE MIGRATION TIME

Assists developers migrating code written in CUDA* to DPC++ once, generating human readable code wherever possible

~80-90% of code migrates automatically

Inline comments are provided to help the developer complete their code
**INTEL® oneAPI DPC++ LIBRARY (BETA)**

**ACCELERATE DPC++ KERNELS ON CPU, GPU & FPGA**

**Optimized C++ Standard Algorithms**
Contains 75 parallelized C++17 algorithms and utilities for efficient application development and deployment on a variety of hardware

**Based on parallel libraries that C++ developers are already familiar with**
Incorporates popular libraries Parallel STL and Boost.Compute for easier developer adoption

**Integrated with Intel® DPC++ Compatibility Tool**
Complements all oneAPI DPC++ components to simplify migration of developers’ CUDA* code to DPC++ code
Boost media and video application performance with hardware-accelerated codecs & programmable graphics on Intel® CPUs and Intel GPUs

Simple API that works the same on CPU and GPU

Using the API, developers have full control over codec visual quality and performance
INTEL® ONEAPI DEEP NEURAL NETWORK LIBRARY(BETA)
DELIVER HIGH PERFORMANCE DEEP LEARNING

- Helps developers create high performance deep learning frameworks
- Abstracts out instruction set and other complexities of performance optimizations
- Same API for both Intel CPU’s and GPU’s, use the best technology for the job
- Supports Linux*, Windows*
- Open sourced for community contributions
INTEL® oneAPI COLLECTIVE COMMUNICATIONS LIBRARY (BETA)

OPTIMIZE COMMUNICATION PATTERNS

Provides optimized communication patterns for high performance on Intel® CPUs and GPUs to distribute model training across multiple nodes

Transparentsly supports many interconnects, such as Intel® Omni-Path Architecture, InfiniBand*, and Ethernet

Built on top of lower-level communication middleware – MPI and libfabrics

Enables efficient implementations of collectives used for deep learning training – all-gather, all-reduce, and reduce-scatter
Analyze Data Parallel C++ (DPC++)
  See the lines of DPC++ that consume the most time

Tune for CPU, GPU & FPGA
  Optimize for any supported hardware accelerator

Optimize Offload
  Tune OpenMP* offload performance

Wide Range of Performance Profiles
  CPU, GPU, FPGA, threading, memory, cache, storage...

Most Popular Languages
  DPC++, C, C++, Fortran, Python*, Go*, Java*, or a mix

There will still be a need to tune for each architecture.
INTEL® ADVISOR (BETA)

DESIGN ASSISTANT — DESIGN FOR MODERN HARDWARE

Offload Advisor
Estimate performance of offloading to an accelerator

Roofline Analysis
Optimize CPU/GPU code for memory and compute

Vectorization Advisor
Add and optimize vectorization

Threading Advisor
Add effective threading to unthreaded applications

Flow Graph Analyzer
Create and analyze efficient flow graphs

Optimization Notice
Copyright © 2019, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.

There will still be a need to tune for each architecture.
High-level language debug support

Multiple accelerator support: CPU, GPU, FPGA emulation

Auto-detect accelerator architecture during application runtime

Non-proprietary open-source solution based on GDB*
Ease of Use
Experienced FPGA users can take advantage of a streamlined programming model using DPC++

Real Time Processing
Process data faster with deterministic low latency and high throughput

Runtime Analysis Support
Collect profiling data at runtime to analyze CPU and FPGA interaction with Intel® VTune™ Profiler

Device Specific Optimizations
One-day class provides experienced FPGA developers training to begin optimizing oneAPI code for FPGA
DETAILS ABOUT INTEL® ONEAPI TOOLKITS\(^{(BETA)}\)

DOMAIN-SPECIFIC TOOLKITS FOR SPECIALIZED WORKLOADS
A toolkit that makes it easier to build, analyze, optimize & scale HPC applications for Intel® Xeon® Scalable, Intel® Core processors & Intel® Accelerators.

**Who Uses It?**
C/C++, Fortran, OpenMP & MPI application developers

**Top Features/Benefits**
- Optimized compilers & performance libraries for Intel® architectures
- Powerful analysis tools to identify optimization opportunities for threading, memory & offloading
- Standards driven to scale forward & preserve development investment

---

### Intel oneAPI Tools for HPC

#### DIRECT PROGRAMMING
- Intel® C++ Compiler with OpenMP®
- Intel® Fortran Compiler with OpenMP®
- Intel® oneAPI DPC++ Compiler
- Intel® Distribution for Python®
- Intel® FPGA Add-on for oneAPI Base Toolkit

#### API-BASED PROGRAMMING
- Intel® MPI Library
- Intel® oneAPI DPC++ Library
- Intel® oneAPI Math Kernel Library
- Intel® oneAPI Data Analytics Library
- Intel® oneAPI Threading Building Blocks
- Intel® oneAPI Video Processing Library
- Intel® oneAPI Collective Communications Library
- Intel® oneAPI Deep Neural Network Library

#### ANALYSIS TOOLS
- Intel® Inspector
- Intel® Trace Analyzer & Collector
- Intel® Cluster Checker
- Intel® VTune™ Profiler
- Intel® Advisor
- GDB®

---

Optimization Notice
Copyright © 2019, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.
Deliver Industry-leading C/C++ & Fortran code performance with OpenMP, unleash the power of the latest Intel® platforms

Develop optimized and vectorized code for Intel® architectures, including Intel® Xeon® processors

Leverage latest language and OpenMP* standards, and compatibility with leading compilers & IDEs

Beta support for Intel GEN9 GPU with OpenMP offload
This toolkit accelerates development of IoT applications for smart connected devices

Who Uses It?
Application developers creating highly reliable edge and IoT commercial devices

Top Features/Benefits
Leverage more cores and built-in technologies in platforms based on IA with the Intel® C++ Compiler
Connect sensors to devices and devices to the cloud with the IoT Connection Tools
Speed development and gain deep hardware and software insights with the Intel® System Debugger

Intel oneAPI Tools for IoT

**DIRECT PROGRAMMING**
- Intel® C++ Compiler
- Eclipse® IDE
- Linux® Kernel Build Tools
- Intel® oneAPI DPC++ Compiler
- Intel® DPC++ Compatibility Tool
- Intel® Distribution for Python*
- Intel® FPGA Add-on for oneAPI Base Toolkit

**API-BASED PROGRAMMING**
- IoT Connection Tools
- Intel® oneAPI DPC++ Library
- Intel® oneAPI Math Kernel Library
- Intel® oneAPI Data Analytics Library
- Intel® oneAPI Threading Building Blocks
- Intel® oneAPI Video Processing Library
- Intel® oneAPI Collective Communications Library
- Intel® oneAPI Deep Neural Network Library
- Intel® Integrated Performance Primitives

**ANALYSIS TOOLS**
- Intel® Inspector
- Intel® System Debugger
- Intel® VTune™ Profiler
- Intel® Advisor
- GDB*

Intel® oneAPI IoT Toolkit + Intel® oneAPI Base Toolkit
A set of 5 powerful, rendering libraries that deliver high-performance, high-fidelity, extensible, & efficient visualization applications and solutions on Intel® platforms.

Who Uses It?
Developers working on high-performance, high-fidelity visualization applications

Key Usages
Creation of studio animation/visual effects content & HPC scientific visualization

Top Features/Benefits
Enables parallelism & performance already in your CPU-based Intel® platforms
Gain performance-optimized ray tracing kernels for photorealistic rendering
Efficiently use threading & vectorization to create interactive, high-fidelity applications

Intel oneAPI Tools for Rendering & Ray Tracing

DIRECT PROGRAMMING
- Intel® oneAPI DPC++ Compiler
- Intel® DPC++ Compatibility Tool
- Intel® Distribution for Python
- Intel® FPGA Add-on for oneAPI Base Toolkit

API-BASED PROGRAMMING
- Intel® Embree
- Intel® Open Image Denoise
- Intel® OpenSWR
- Intel® Open Volume Kernel Library
- Intel® OSPRay
- Intel® oneAPI DPC++ Library
- Intel® oneAPI Data Analytics Library
- Intel® oneAPI Deep Neural Network Library
- Intel® oneAPI Video Processing Library
- Intel® oneAPI Threading Building Blocks
- Intel® oneAPI Integrated Performance Primitives
- Intel® oneAPI Collective Comms. Library

ANALYSIS TOOLS
- Intel® VTune™ Profiler
- Intel® Advisor
- GDB*

Intel® oneAPI Rendering Toolkit
Intel® oneAPI Base Toolkit
INTEL® oneAPI DL FRAMEWORK DEVELOPER TOOLKIT (BETA)

Designed for developers & researchers who want to create the next great deep learning framework or optimize existing ones.

Key Usages
Deep Learning Framework Development
Deep Learning Research

Top Features/Benefits
Create fast deep neural networks that can take advantage of Intel's CPU and accelerators
Scale your framework from one node to multiple nodes providing faster analysis for the framework's workload

Intel oneAPI Tools for Deep Learning Framework Developers

API-BASED PROGRAMMING

Intel® oneAPI Deep Neural Network Library
Intel® oneAPI Collective Communications Library

Optimization Notice
Copyright © 2019, Intel Corporation. All rights reserved.
*Other names and brands may be claimed as the property of others.
A toolkit that provides optimized software tools to accelerate End to End AI development

Who Uses It?
AI Researchers & application developers, data scientists

Key Usages
AI Research & applications across Finance, Retail, e-commerce, robotics, transportation & more

Top Features/Benefits
Achieve greater deep learning performance for training and inference phases with optimized frameworks.
Accelerate data science and analytics stages with Python* packages enhanced for Intel® architectures
A toolkit to accelerate development of high performance deep learning inference & computer vision into vision/AI applications used from edge to cloud. It enables deep learning on hardware accelerators & easy deployment across multiple types of Intel® platforms.

**Who needs this product?**
Computer vision, deep learning software developers
Data scientists
OEMs, ISVs, System Integrators

**Usages**
Security surveillance, robotics, retail, healthcare, AI, office automation, transportation, non-vision use cases (speech, NLP, Audio, text) & more.
This toolkit helps accelerate system bring-up & optimization for Intel® platforms.

Who Uses It?
Hardware & system software developers creating highly reliable & optimized Intel-based solutions

Key Usages
Used by IoT, client & server customers to bring-up a platform & optimize the power/thermal & performance characteristics

Top Features/Benefits
Strengthen system reliability with a powerful debug and trace tool providing deep hardware and software insight
Analyze and optimize how Intel silicon operates with respect to power
Take advantage of hardware capabilities for optimal system performance on the Intel® platforms
ECOSYSTEM ADOPTION & SUPPORT

Drive adoption of Data Parallel C++ language & oneAPI library APIs

**Developer Enabling Programs**

- oneAPI Zone
- Academic Programs & Curriculum
- Developer Cloud for latest hardware & software
- Development Kits, Reference Architectures
- ISV & Application Enabling

**Support**

- Developer Guides, Training & More
- Community Forums
- Intel Tools Support Forums

**Ecosystem Engagement**

- Existing Software Standards (OpenMP*, MPI)
- Strategic Industry Collaborations
- Open Source Community
- Third Party Solutions

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

Optimization Notice

Copyright © 2019, Intel Corporation. All rights reserved.
A development sandbox to develop, test and run your workloads across a range of Intel CPUs, GPUs, and FPGAs using Intel’s oneAPI beta software

**NO DOWNLOADS | NO HARDWARE ACQUISITION | NO INSTALLATION | NO SETUP & CONFIGURATION**

**GET UP & RUNNING IN SECONDS!**
NOTICES & DISCLAIMERS

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request. No product or component can be absolutely secure. Intel technologies’ features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

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.

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.

Copyright © 2019, Intel Corporation. All rights reserved. Intel, the Intel logo, Xeon, Core, VTune, and OpenVINO are trademarks of Intel Corporation or its subsidiaries in the U.S. and other countries.

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