HLRS-Intel oneAPI HPC Workshop

October 26th, 2022

Edmund Preiss

oneAPI

intel

Objectives

- Get to know / Learn / Understand :
  - The oneAPI programming model
  - Building applications with DPC++/SYCL
  - Fundamentals of OpenMP offloading
  - How to use Intel’s oneAPI libraries (oneMLK, ...) and APIs
  - Intel’s heterogenous profiling and performance analysis tools
  - A basic understanding on (dynamical) debugging of applications using the oneAPI programming model
  - Intel's Compatibility tool that helps to migrate CUDA to SYCL code.

AGENDA

Day 1 : Oct 26th, 2022

			TOPIC	Presenter
09:00	09:10	00:10	Welcome and Introduction to Day 1	Dr Tobias Haas (HLRS), Edmund Preiss (Intel)
09:10	09:30	00:20	oneAPI – Introduction to a new Development Environment - Concept and oneAPI Standardization initiative - Intel's Tools Implementation – Intel oneAPI Toolkits and libs - Transition from Intel Parallel Studio XE to Intel oneAPI toolkits	Edmund Preiss (Intel)
09:30	09:50	00:20	Introduction to the DevCloud - Purpose: Demoing, testing and porting applications - Hardware and Software offerings - How to onboard & how to get an DevCloud account	Klaus-Dieter Oertel (Intel)
09:50	*10:00*	*00:10*	*Break*	
10:00	11:00	01:00	Direct programming with oneAPI Compilers (Part 1) – with Demos - Intro to heterogenous programming model with SYCL 2020 - SYCL features and examples o "Hello World" Example o Device Selection o Execution Model	Igor Vorobtsov (Intel)
11:00	*11:15*	*00:15*	*Break*	
11:15	12:30	01:15	Direct programming with oneAPI Compilers (Part 2) – with Demos o Compilation and Execution Flow o Memory Model; Buffers, Unified Shared Memory (USM) o Performance optimizations with SYCL features	Igor Vorobtsov (Intel)

AGENDA

Day 2 : Oct 27th, 2022

			Topic	Presenter
09:00	10:00	01:00	Intel OpenMP for Offloading – with Demos - Parallelizing heterogenous applications with OpenMP 5.1 - Mixing of OpenMP and SYCL	Alina Shadrina (Intel)
10:00	10:35	00:35	Intel oneAPI libraries (oneMKL) for HPC - with demos - Performance optimized libraries for numerical simulations and other purposes	Gennady Fedorov (Intel)
10:35	*10:50*	*00:15*	Break	
10:50	11:20	00:30	Intel Debugging Tools for heterogenous programming (CPU, GPU) - with demos	Alina Shadrina (Intel)
11:20	12:00	00:40	Open Source Compatibility tool for porting purposes(SYCLomatic) - with demo - Migration Cuda based GPU Applications to SYCL	Alina Shadrina (Intel)
12:00	12:30	00:30	Dynamic Debugging with Intel Inspector - with demos - Identifying Memory and Threading Errors (Data Races and Deadlocks)	Heinrich Bockhorst (Intel)
12:30	12:35	00:05	Questions and Answers - Wrap up of the day	Edmund Preiss (Intel)

AGENDA

Day 3 : Oct 28th, 2022

			TOPIC	Presenter
09:00	10:10	01:10	Application profiling for heterogenous hardware - Demos - Profiling Tools Interfaces for GPU - Open Source lightweight Tools - Profile heterogenous SYCL/OpenMP Workloads with Intel VTune Profiler - Share experiences/key findings with Gromacs related porting and optimization efforts	Heinrich Bockhorst (Intel)
10:10	*10:20*	*00:10*	Break	
10:20	11:30	01:10	Application profiling for heterogenous hardware - Demos - Estimate performance potential gains with Offload Advisor (CPU -> HW Accelerator) - Analyse heterogenous SYCL/OpenMP Workloads with Intel Advisor and Roofline analysis	Klaus-Dieter Oertel (Intel)
11:30	*11:40*	*00:10*	Break	
11:40	12:40	01:00	A 3rd Party oneAPI Case Study: GROMACS - A Molecular Dynamics Engine - Heterogenous Design consideration, alternatives and comparisons - Real Scheduling - SYCL - oneAPI and other Implementations - SYCL in GROAMCS 2022	Andrey Alekseenko (KTH, Sweden)
12:40	*12:45*	*00:05*	Break	
12:45	13:10	00:25	Programming for Distributed HPC Systems using Intel MPI	Rafael Lago (Intel)
13:10	13:15	00:05	- Questions and Answers - Wrap up	Edmund Preiss (Intel)

Call to Action

- Data Centre Admins
  - Prepare and Update your data center with performance optimized Intel oneAPI Toolkits to serve your users and developers
- Developers
  - Use your knowledge about Intel oneAPI Toolkits for application(s) development
  - Move CUDA code to SYCL
  - Develop applications with new LLVM based Intel C++ (ICX) and Fortran (IFX) Compilers
  - Practice with exercises available on Intel DevCloud

oneAPI -
A new Development Environment

October 26th, 2022

Edmund Preiss

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex. Results may vary.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details.
No product or component can be absolutely secure.

Slide 50 - **Texas Advanced Computing Center (TACC) Frontera references**
Article: *HPCWire: Visualization & Filesystem Use Cases Show Value of Large Memory Fat Notes on Frontera*.
www.intel.com/content/dam/support/us/en/documents/memory-and-storage/data-center-persistent-mem/Intel-Optane-DC-Persistent-Memory-Quick-Start-Guide.pdf
software.intel.com/content/www/us/en/develop/articles/introduction-to-programming-with-persistent-memory-from-intel.html
wreda.github.io/papers/assise-osdi20.pdf
KFBIO
KFBIO m. tuberculosis screening detectron2 model throughput performance on 2nd Intel® Xeon® Gold 6252 processor: NEW: Test 1 (single instance with PyTorch 1.6: Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel® Xeon® Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated Test 2 (24 instances with PyTorch 1.6: Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel Xeon Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated BASELINE: (single instance with PyTorch 1.4): Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel Xeon Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated.

Tangent Studios
Configurations for Render Times with Intel® Embree, testing conducted by Tangent Animation Labs. Render farm: 8x Intel® Core™ processors +hyperthread*2 + 128gig. In-office workstations: Intel® Xeon® processors HP blade c7000 chassis, with HP460 gen8 blades - 2x Intel Xeon E5-2650 V2, Eight Core 2.6GHz-128GB. Software: Blender 2.78 with custom build using Intel® Embree. For more information on Tangent's work with Embree, watch this video:
www.youtube.com/watch?time_continue=251&v=_2Ia4h8q3xs&feature=emb_logo
Recreation of the performance numbers can be recreated using Agent327, Blender and Embree.

Chaos Group - Up to 90% Memory Reduction for Displacement
Testing conducted by Chaos Group with Intel® Embree 2020. Software Corona Renderer 5 with Intel Embree. Up to 90% memory reduction calculated using Corona Renderer 5 with regular displacement grids per triangle of 154 bytes versus Corona Renderer 5 with Intel Embree, which has a displacement capability grid of 12 bytes per grid triangle. (12/154 = 7.8% usage or >90% memory reduction.) Recreation of the performance numbers can be accomplished using Corona Renderer 5 and Embree. For more information, visit the Corona Renderer Blog: blog.corona-renderer.com/corona-renderer-5-for-3ds-max-released/

The Addams Family 2 - Gained a 10% to 20%—and sometimes 25%—efficiency in rendering, saving thousands of hours in rendering production time.
Testing Date: Results are based on data conducted by Cinesite 2020-21. 10% to up to 25% rendering efficiency/thousands of hours saved in rendering production time/15 hrs per frame per shot to 12-13 hrs.
Cinesite Configuration: 18-core Intel® Xeon® Scalable processors (W-2295) used in render farm, 2nd gen Intel Xeon processor-based workstations (W-2135 and -2195) used. Rendering tools: Gaffer, Arnold, along with optimizations by Intel® Open Image Denoise.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

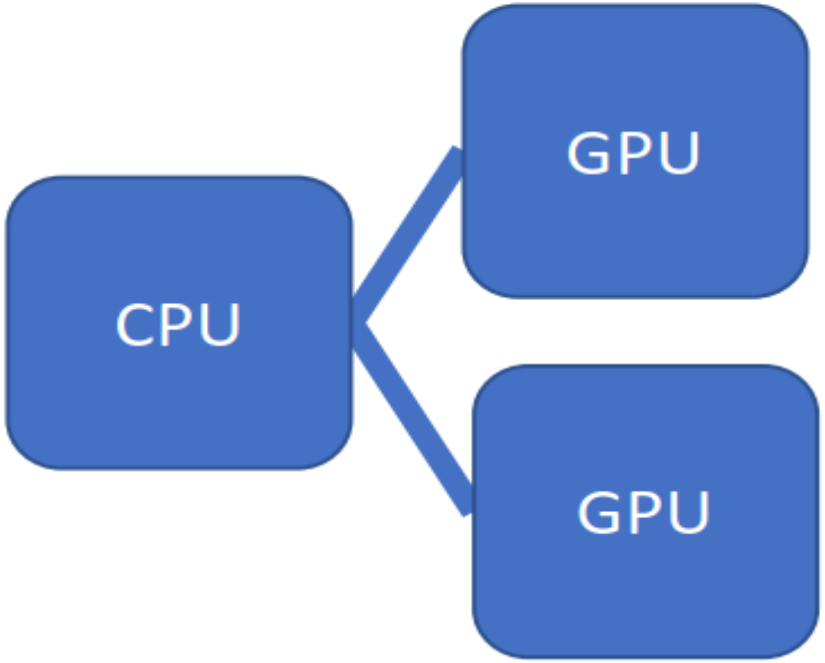
Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Agenda

- A Glimps on Future Evolution of the HPC Computer Architecture
- oneAPI Concept and the need for Standardization for heterogenous Programming
  - SYCL and Data Parallel C++
- The Intel® oneAPI Toolkits and Software Development Components
  - Key oneAPI Tool Components
- Examples of oneAPI Enabling & Workload Migration Activities
- Miscellenous / oneAPI Resources and useful links

How does a machine look like in a heterogenous world?

Sharing Parallism between CPU and additional Accelerators



Intel's diverse Computer Architecture

Diverse accelerators needed to meet today's performance requirements:
48% of developers target heterogeneous systems[1]

CPU

GPU

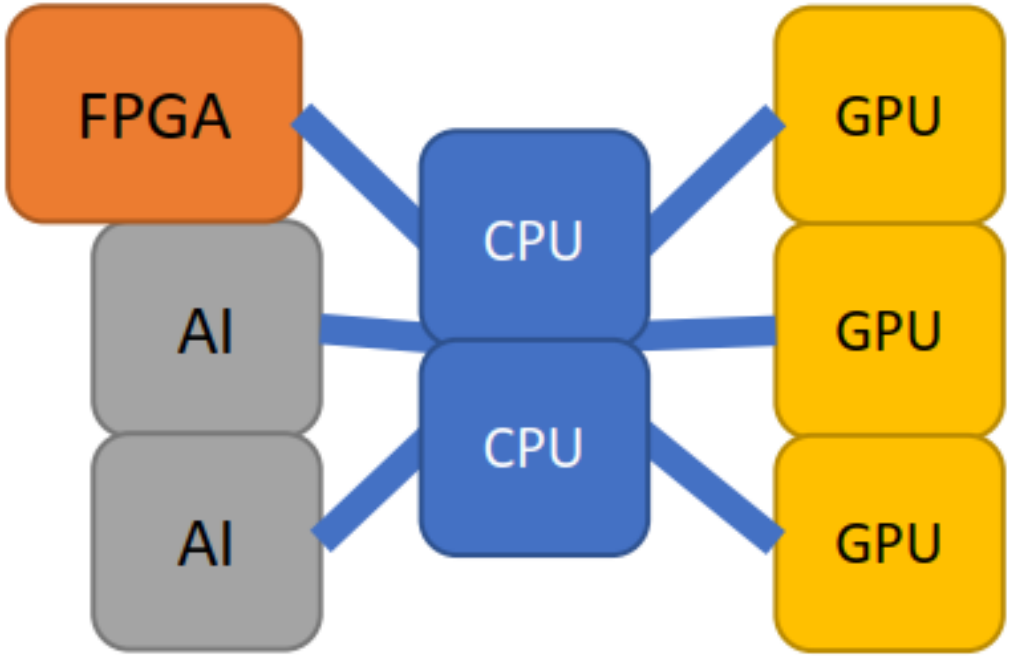
FPGA

Developer Challenges: Multiple Architectures, Vendors, and Programming Models

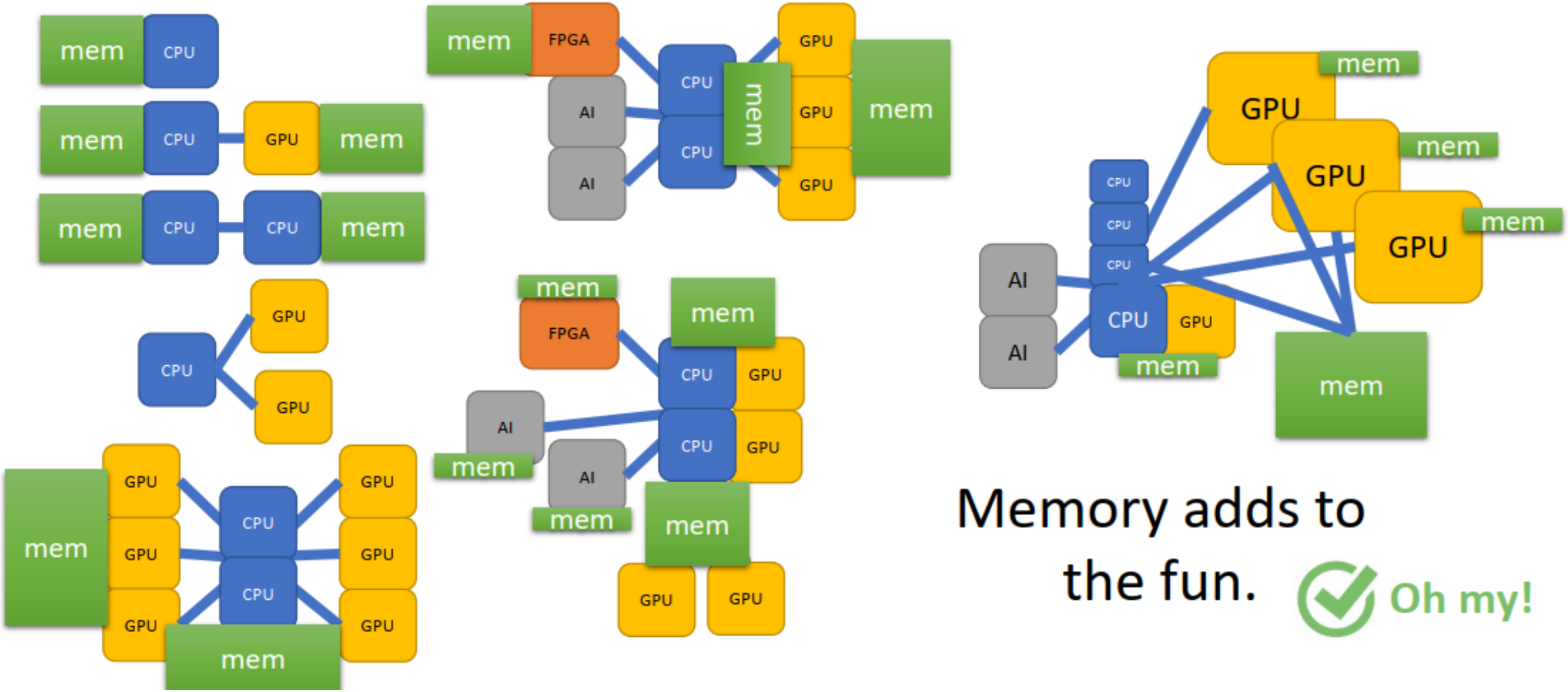
[1]Evans Data Global Development Survey 2020, Volume 2

How does a machine look like in a heterogenous world?

A mix of different Accelerators- All running in parallel



How does a machine look like in a heterogenous world ?



Can we really program XPUs (acceleration)?

1. Freedom = Choice of XPUs
2. Value = Maintain Performance across XPUs
3. Trustworthy = Maintain One Source Code for Future XPUs

oneAPI: One Name, Two Distinct Objectives

oneAPI

- Open industry specification
- Open-source repo and development
- Community driven
- Supports multi-vendor implementation

---

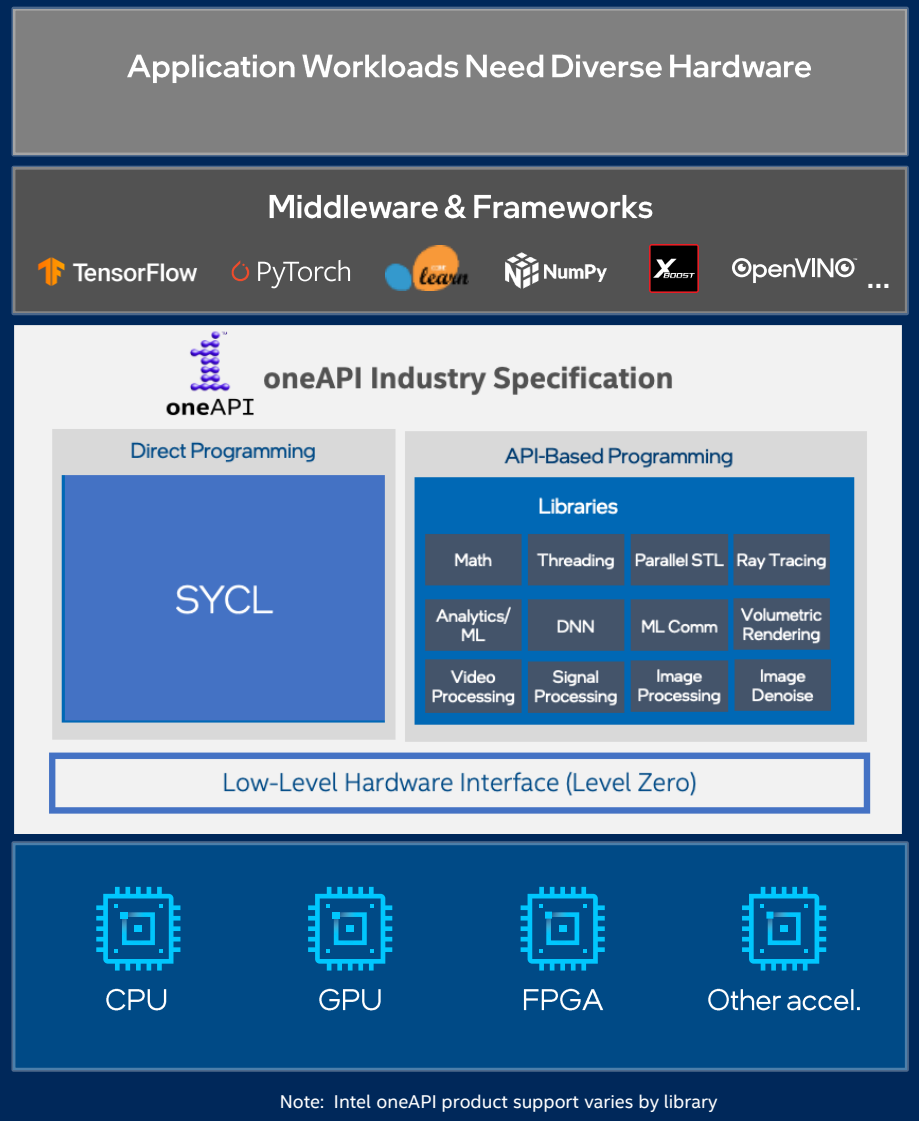
oneAPI

- Visit oneapi.io for more details
- Intel's implementation of oneAPI standards + additional languages and programming models
- Toolkits optimized for Intel CPUs, GPUs, and FPGAs
- Broadly available for download with paid priority support

Open, standard language Khronos SYCL

Standardized interfaces for common libraries

Standardized hardware interface



Note: Intel oneAPI product support varies by library

Data Parallel C++

Standards-based, Most Comprehensive, Cross-architecture Implementation of SYCL

DPC++ = ISO C++ and Khronos SYCL and community extensions

Freedom of Choice: Future-Ready Programming Model

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

DPC++ = ISO C++ and Khronos SYCL and community extensions

- Designed for data parallel programming productivity
- Provides full native high-level language performance on par with standard C++ and broad compatibility
- Adds SYCL from the Khronos Group for data parallelism and heterogeneous programming

Community Project Drives Language Enhancements

- Provides extensions to simplify data parallel programming
- Continues evolution through open and cooperative development

Direct Programming:
SYCL/Data Parallel C++

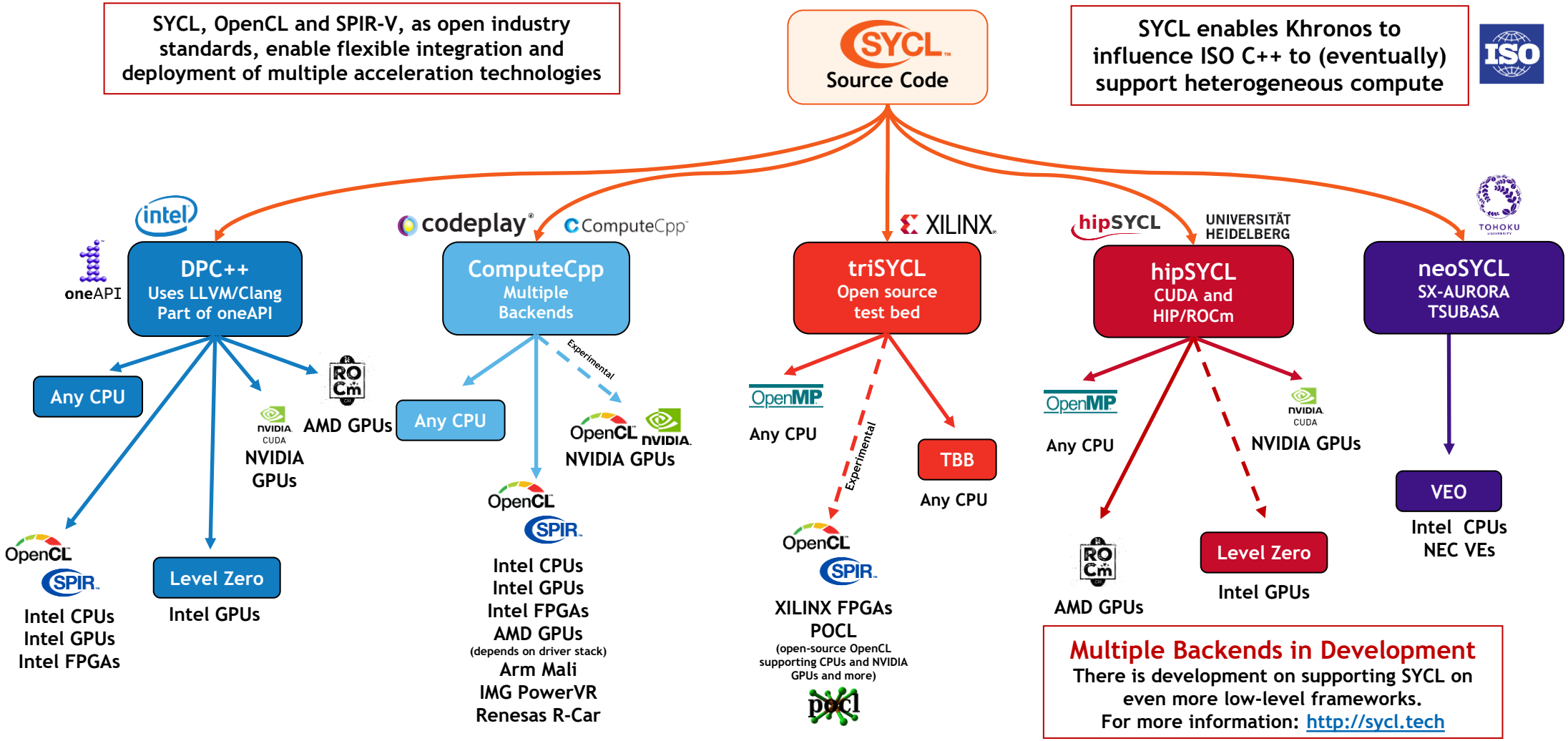
Community Extensions

Khronos SYCL

ISO C++

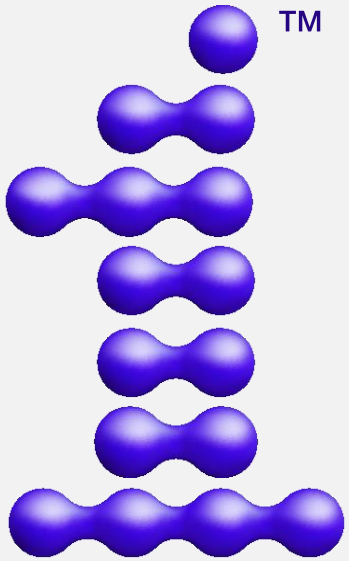
The open source and Intel DPC++/C++ compiler supports Intel CPUs, GPUs, and FPGAs.
Codeplay announced a DPC++ compiler that targets Nvidia GPUs.

SYCL ecosystem is growing



https://www.khronos.org/blog/sycl-2020-what-do-you-need-to-know

+ Celerity: SYCL on MPI+SYCL



oneAPI

oneAPI

Industry Specification

spec.oneapi.com/oneAPI/

- Notices and Disclaimers
- Contribution Guidelines
- Introduction
- Software Architecture
- Library Interoperability
- oneAPI Elements
- Data Parallel C++ (DPC++)
- oneAPI Data Parallel C++ Library (oneDPL)
- oneAPI Deep Neural Network Library (oneDNN)
- oneAPI Collective Communications Library (oneCCL)
- oneAPI Level Zero (Level Zero)
- oneAPI Data Analytics Library (oneDAL)
- oneAPI Threading Building Blocks (oneTBB)
- oneAPI Video Processing Library (oneVPL)
- oneAPI Math Kernel Library (oneMKL)
- Contributors

oneAPI Ecosystem Support

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.

oneAPI: Open Accelerator Ecosystem

Freedom of Choice in Hardware Drives Productivity

oneAPI

Codeplay contribution to DPC++ brings SYCL support for NVIDIA GPUs

oneAPI oneDNN on Arm for A64FX Fugaku

Extending DPC++ with Support for Huawei AI Chipset

NERSC, ALCF, CODEPLAY PARTNER ON SYCL FOR NEXT-GENERATION SUPERCOMPUTERS

on Nvidia

ARGONNE, ORNL AWARD CODEPLAY CONTRACT TO STRENGTHEN SYCL SUPPORT FOR AMD GPUS

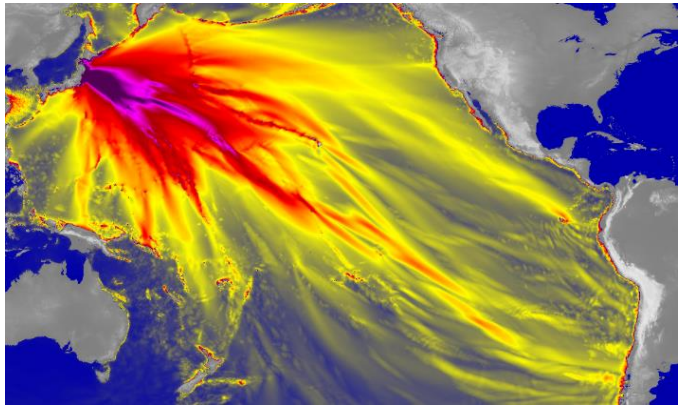
European exascale combines SiPearl's CPU RHEA with Intel's Xᵉ GPU PVC

"DPC++ and oneAPI helped us to develop much faster the accelerators for machine learning algorithms."

– Chris Kachris, co-founder, InAccel

"If you like modern, standard C++ and you want to target GPUs or other accelerators, you will love SYCL!"

– Marcel Breyer



Visualization of *easyWave* tsunami simulation application - Courtesy Zuse Institute Berlin (ZIB)

oneAPI

Intel® oneAPI Toolkits and Components

intel®

Intel® oneAPI Tools

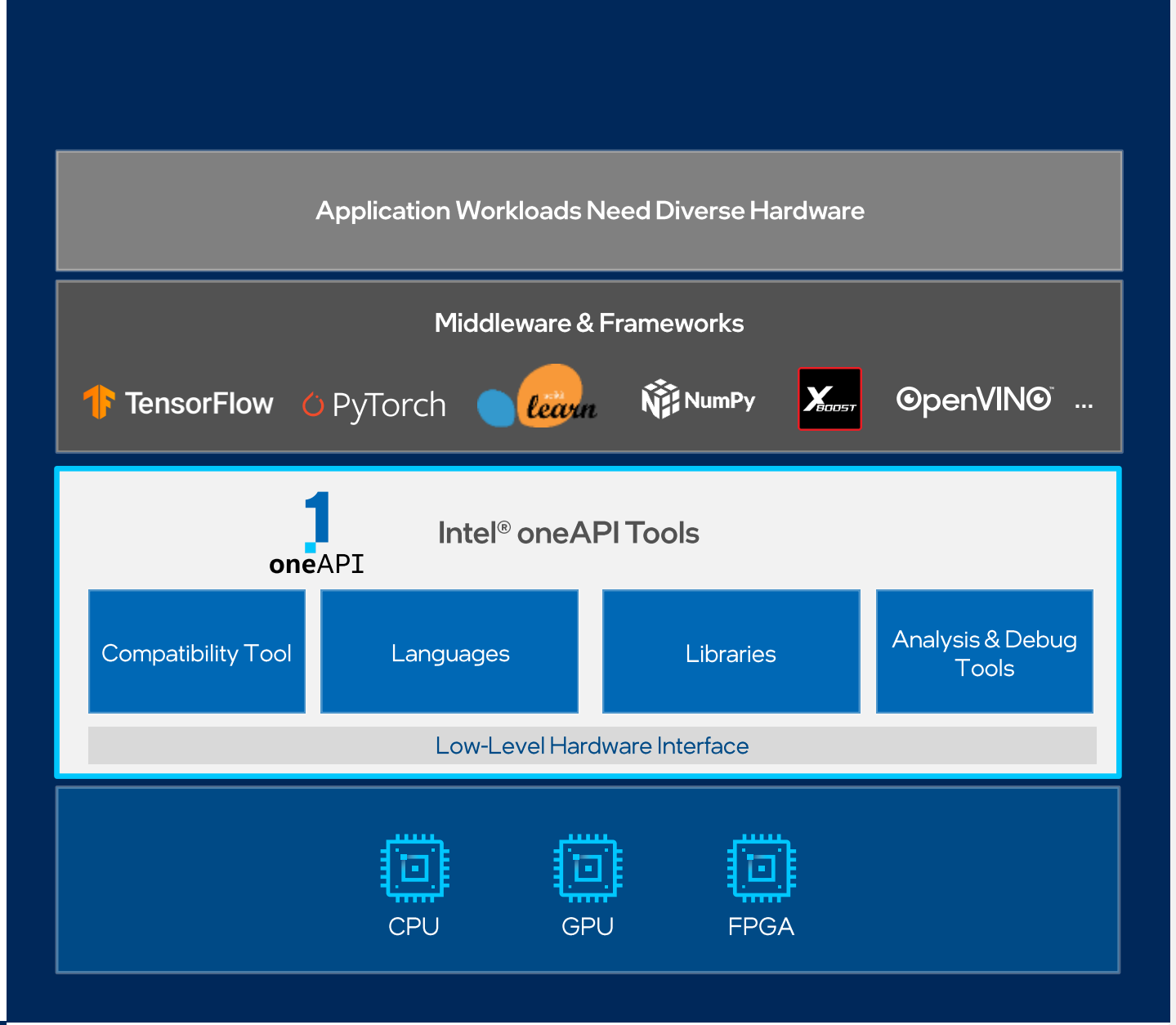
Built on Intel's Rich Heritage of CPU Tools Expanded to XPUs

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

- Accelerates compute by exploiting cutting-edge hardware features
- Interoperable with existing programming models and code bases (C++, SYCL, Fortran, Python, OpenMP, etc.), developers can be confident that existing applications work seamlessly with oneAPI
- Eases transitions to new systems and accelerators
- Using a single code base frees developers to invest more time on innovation

Available with paid Commercial Support

Latest version is 2022.2



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.

Analysis & Debug Tools

Get More from Diverse Hardware

oneAPI

Design

Intel® Advisor

- Efficiently offload code to GPUs
- Optimize your CPU/GPU code for memory and compute
- Enable more vector parallelism and improve efficiency
- Add effective threading to unthreaded applications

Debug

Intel® Distribution for GDB

- Multiple accelerator support with CPU, GPU and FPGA
- Enables deep, system-wide debug of Data Parallel C++ (DPC++), C, C++, and Fortran code

Tune

Intel® VTune™ Profiler

- Tune for GPU, CPU, and FPGA
- Optimize offload performance
- Supports DPC++, C, C++, Fortran, Python, Go, Java or a mix of languages

Intel® oneAPI Toolkits

A complete set of proven developer tools expanded from CPU to XPU (accelerators)

oneAPI

Intel® oneAPI **Base** Toolkit

A core set of high-performance libraries and tools for building C++, SYCL and Python applications

intel. 1 oneAPI BASE TOOLKIT

Add-on **Domain-specific** Toolkits

intel. 1 oneAPI HPC TOOLKIT

Intel® oneAPI Tools for HPC

Deliver fast Fortran, OpenMP & MPI applications that scale

intel. 1 oneAPI IoT TOOLKIT

Intel® oneAPI Tools for IoT

Build efficient, reliable solutions that run at network's edge

intel. 1 oneAPI AI ANALYTICS TOOLKIT

Intel® oneAPI AI Analytics Toolkit

Accelerate machine learning & data science pipelines with optimized DL frameworks & high-performing Python libraries

intel. 1 oneAPI RENDERING TOOLKIT

Intel® oneAPI Rendering Toolkit

Create performant, high-fidelity visualization applications

Toolkit **powered by oneAPI**

OpenVINO™

Intel® Distribution of OpenVINO™ Toolkit

Deploy high performance inference & applications from edge to cloud

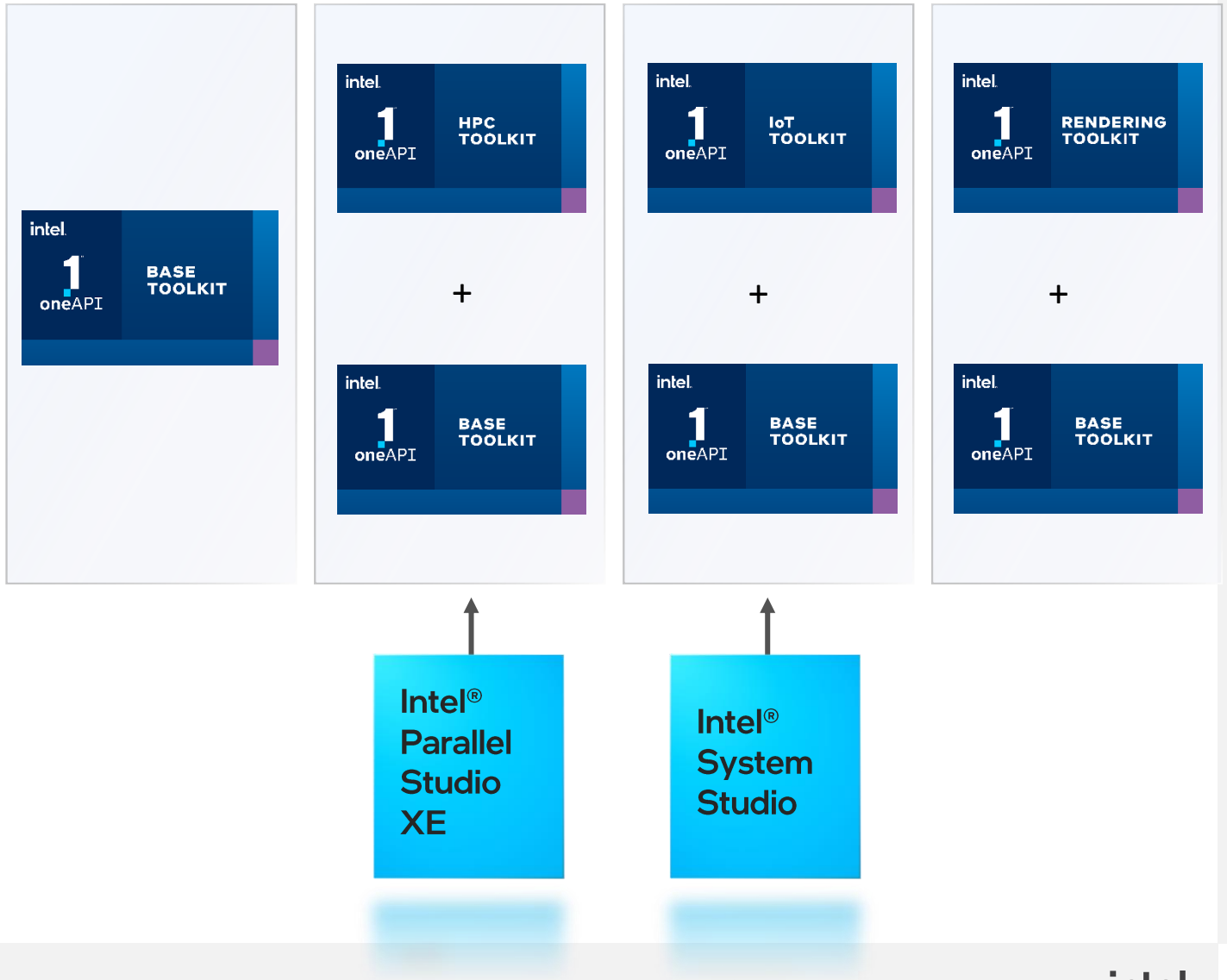
Latest version available 2022.1

Commercial Toolkits Deliver Priority Support (Paid Support Licenses)

Next Generation of Commercial Intel® Software Development Products

- Worldwide support from Intel technical consulting engineers
- Prior commercial tool suites, Intel® Parallel Studio XE and Intel® System Studio, transition to oneAPI products

oneAPI



Intel® oneAPI Toolkits Availability

Get Started Quickly

Code Samples, Quick-start Guides, Webinars, Training

software.intel.com/oneapi

oneAPI

Run the tools locally

- Downloads
- Repositories
- Containers

Or run the tools in intel DevCloud

- 1 Minute to Code
- No Hardware Acquisition
- No Download, Install or Configuration
- Samples & Tutorials
- Supports Jupyter Notebooks, Visual Studio Code

Get Up & Running In Seconds!

Intel® oneAPI Toolkits – Proven Performance

Top Takeaways & Proof Points

- **HPC Cross-architecture** – **Argonne National Labs** is running Exascale-class applications efficiently on current and future generations of Intel CPUs and GPUs
- **HPC Cross-architecture** – **Zuse Institute Berlin (ZIB)** ported the tsunami simulation *easyWave* application from CUDA to Data Parallel C++ delivering performance across multiple architectures from multiple vendors
- **HPC & AI** – **CERN** uses Intel® DL Boost and oneAPI to speed simulations with inference acceleration by nearly **2x** without accuracy loss*
- **Hyper-real Visualization & AI Using Advanced Ray Tracing** – **Bentley Motors Limited's** AI-based car configurator processes **1.7M+** images with up to **10B** possible configurations per model*
- **IoT** – **Samsung Medison** accelerates ultrasound image processing at the edge on multiple Intel® architectures for improved accuracy and fast diagnosis
- **Major CSPs & Framework endorse oneAPI** – **Microsoft Azure, Google Cloud, TensorFlow**
- **FPGA** – Using oneAPI, **Bittware** had its application running **in days** vs. what typically would take several weeks using Verilog or VHDL*
- And more… 250+ applications developed with oneAPI tools > view catalog

Video [3:45]

*Detailed slides per customer are noted in the oneAPI Customer Use Cases deck. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy. See Notices & Disclaimers for more details.

GROMACS – Using oneAPI

Intel oneAPI Tools: Empowering GROMACS Cross-Architecture Development

@IntelDevTools

„... "The part of oneAPI that is most important to me and my team is that, of course, it's an open standard. We are firm believers in open standards, particularly in the long run, because that means we can rely on it no matter what the vendors do. …" , Erik Lindahl; **Video 2** @sycl.tech

Click on the image to run the first video

oneAPI Resources

software.intel.com/oneapi

Get Started

- software.intel.com/oneapi
- Documentation + dev guides
- Code Samples
- Intel® DevCloud

Industry Initiative

- oneAPI.io
- oneAPI open Industry Specification
- Open-source Implementations

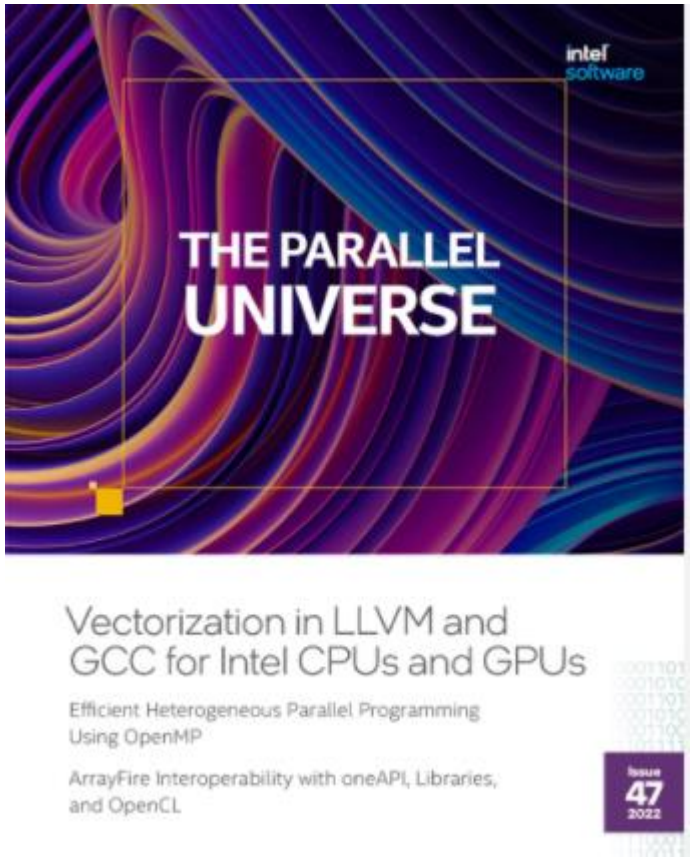
Learn

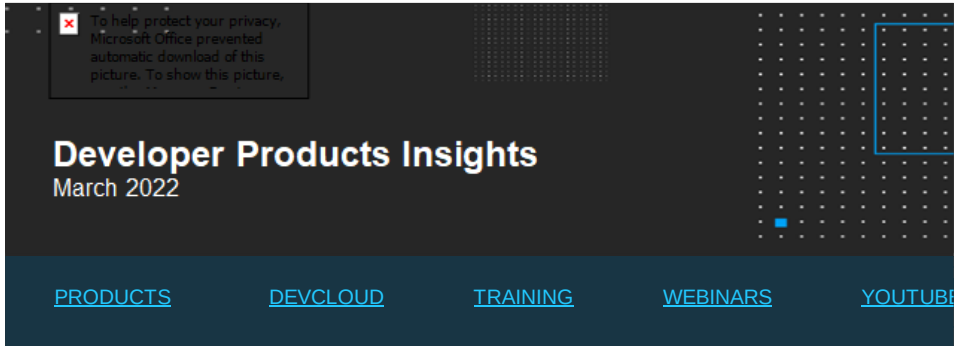
- Training: Webinars & courses
- Intel® DevMesh Innovator Projects
- Summits & Workshops: Live & on-demand virtual workshops, community-led sessions
- Training by certified oneAPI experts worldwide for HPC & AI

Ecosystem

- Community Forums
- Intel® DevMesh Innovator Projects
- Academic Programs: oneAPI Centers of Excellence: research, enabling code, curriculum, teaching

Other useful Content Resources





Forwarded from a friend? Subscribe now →

Featured Content

5 Outstanding Additions in SYCL

SYCL 2020 offers C++ programmers 5 new features to take advantage of accelerators and the potential of open, cross-platform development. Find out what they are and how you benefit.

Read it →

Click on images to activate links & to subscribe

Summary

.

- oneAPI cross-architecture, one source programming model provides freedom of XPU choice.
  Apply your skills to the next innovation, not to rewriting software for the next hardware platform.
- Intel® oneAPI Toolkit products take full advantage of accelerated compute by maximizing performance across Intel CPUs, GPUs, and FPGAs.
- Develop confidently with a proven set of cross-architecture libraries and advanced tools that interoperate with existing performance programming models.


intel