The Future of Intel Performance Analyzers

Dr. Sri Doddapaneni

Senior Director, Developer Software
Intel Corporation

Guest Speaker:

Ilias Katsardis

HPC Solution Lead
Google

SPDK, PMDK, Intel® Performance Analyzers | **Virtual Forum**

Agenda

01   oneAPI Overview
One Programming Model for Multiple Architectures and Vendors

02   oneAPI Analyzers
Intel Performance Analyzers getting oneAPI ready

03   Guest Speaker: Ilias Katsardis, Google
HPC at Google Cloud

04   Support for Cloud Environment
Evolve tools to enable new use-cases
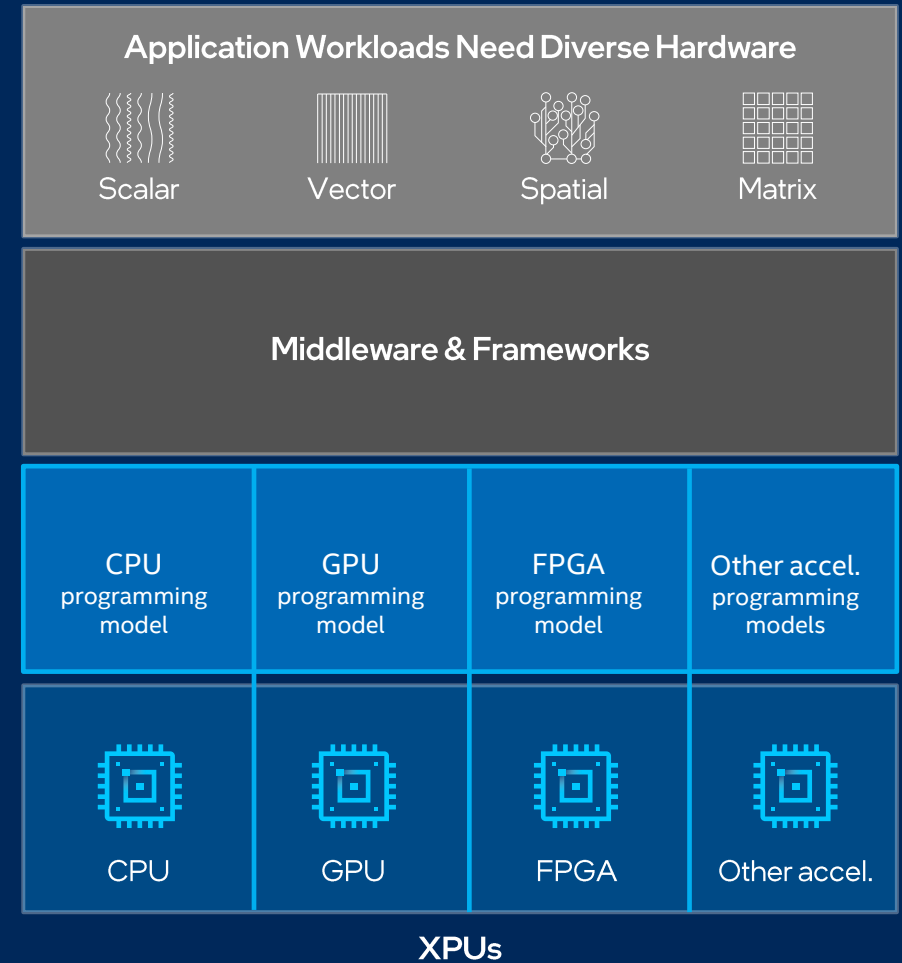
05   Conclusion
Thank You!

intel.

Programming Challenges
for Multiple Architectures

Growth in specialized workloads

Variety of data-centric hardware required

Separate programming models and toolchains for each architecture are required today

Software development complexity limits freedom of architectural choice

Application Workloads Need Diverse Hardware

| Scalar | Vector | Spatial | Matrix |

Middleware & Frameworks

| CPU programming model | GPU programming model | FPGA programming model | Other accel. programming models |

| CPU | GPU | FPGA | Other accel. |

XPUs

intel.

oneAPI

One Programming Model for Multiple Architectures and Vendors

Freedom to Make Your Best Choice

Choose the best accelerated technology the software doesn't decide for you

Realize all the Hardware Value

Performance across CPU, GPUs, FPGAs, and other accelerators
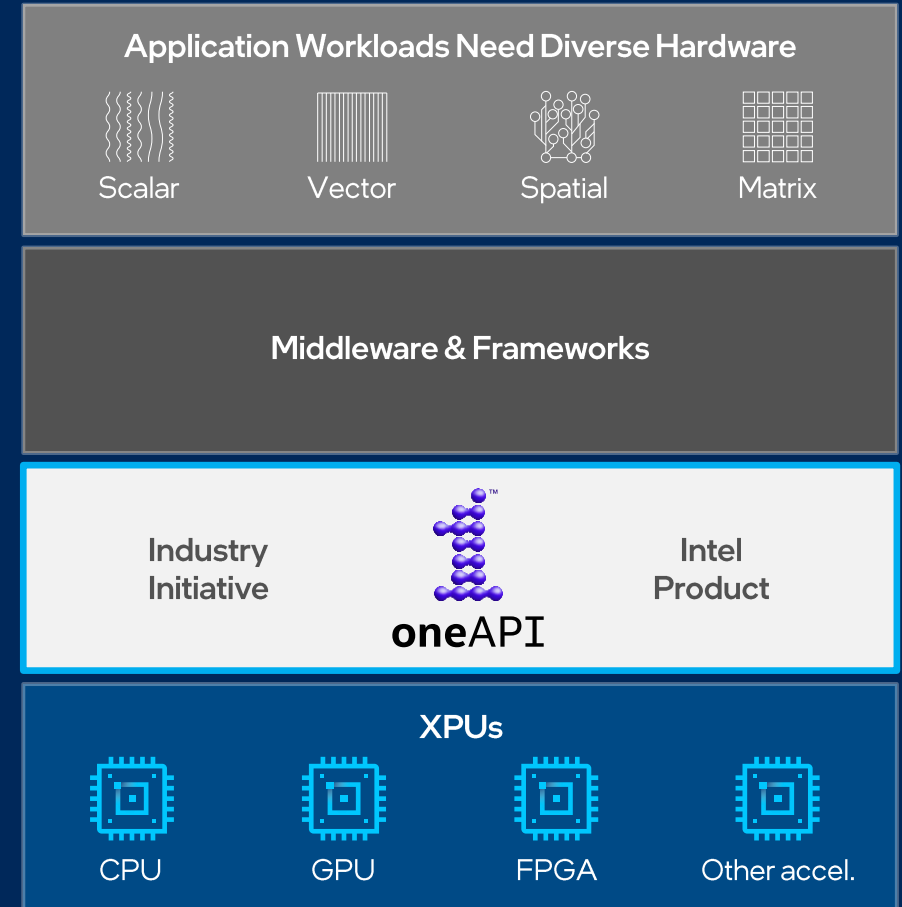
Develop & Deploy Software with Peace of Mind

Open industry standards provide a safe, clear path to the future

Compatible with existing languages and programming models including C++, Python, SYCL, OpenMP, Fortran, and MPI

oneAPI Industry Initiative
- Open to promote community and industry collaboration
- Enables code reuse across architectures and vendors

Application Workloads Need Diverse Hardware

Scalar	Vector	Spatial	Matrix

Middleware & Frameworks

Industry Initiative	oneAPI™	Intel Product

XPUs

CPU	GPU	FPGA	Other accel.

intel.

oneAPI Performance Analyzers

- Intel® VTune™ profiler supports offload hotspot analysis and in-depth analysis of kernel execution on GPU

- Intel®Advisor provides guidance on profitability of offloading kernels to target GPU based on CPU profile, and then analyzes for optimality of kernels using GPU roofline analysis

- *Coming soon:*

  - *VTune support for multiple GPUs, scalability analysis both intra-node and cluster-wide*

  - *Guidance on kernel code transformations in Advisor*

  - *Memory error checking of kernels offloaded to GPU in Intel® Inspector*
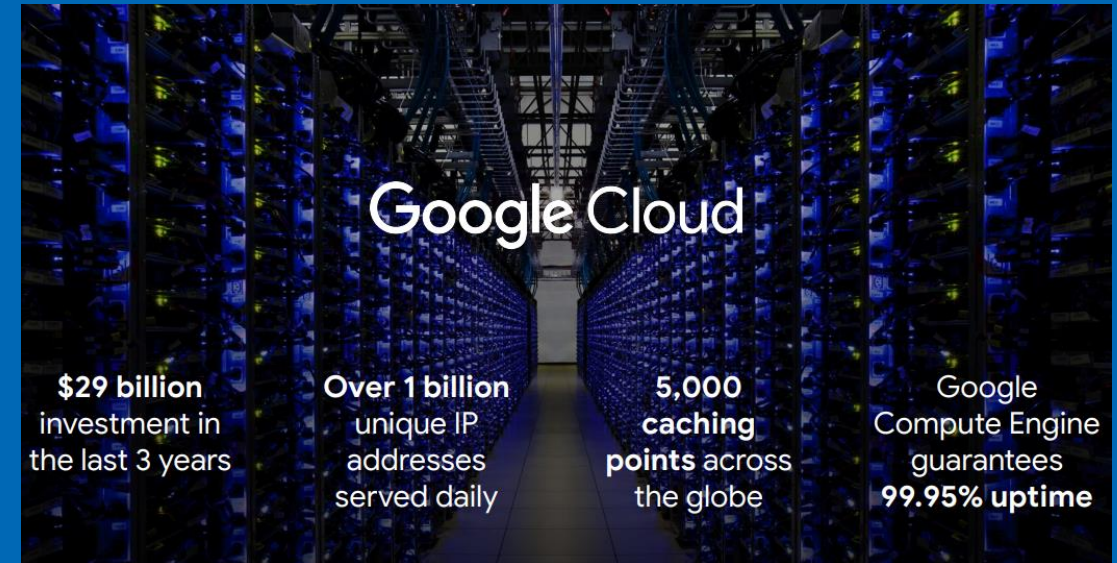
Accelerating Google Cloud for HPC
Democratize HPC & make it Universally Accessible & Useful

Cross-architecture, Cross-vendor oneAPI HPC Cloud Solution

- C2 provides great performance for HPC workloads requiring high, consistent performance with underlying hardware visibility: **40% higher performance/core**

- NUMA-aware for performance, high performance per thread, Isolation for latency sensitive workloads

- Memory-optimized VMs / compute-optimized workloads

- Performance-sensitive for CPU workloads, or licensed applications that may benefit from more powerful cores

- HPC: Simulations (finite element analysis, oil & gas, CFD, Monte Carlo, product simulation, weather, physics, chemistry), Financial Services (analysis, simulation), Genomic Analysis, Media Transcoding, Electronic Design Automation

- Runs on multiple generations of Intel® Xeon® processors, optimized by Intel® oneAPI **Base** & **HPC** Toolkits.

$29 billion investment in the last 3 years

Over 1 billion unique IP addresses served daily

5,000 caching points across the globe
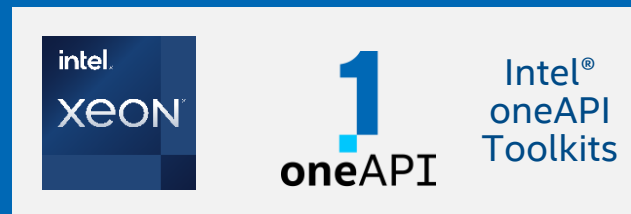
Google Compute Engine guarantees **99.95% uptime**

Learn more – Podcast: *Next Gen of Deep Learning with Google TensorFlow & Intel*
Image courtesy Google Cloud

Cross-vendor Multi-architecture Deployments

Used Multiple tools in Intel® oneAPI Toolkits

intel. XEON

1 oneAPI

Intel® oneAPI Toolkits

*"…We recommend using **Intel MPI** for best performance, and tools such as **VTune Profiler** and **Advisor** to help better understand performance optimizations and how to best migrate your workloads to the cloud."*
Ilias Katsardis, HPC Solution Lead, Google

intel.

New challenges from Cloud Environment

Intel performance analyzers are arguably the best at showing opportunities for improving performance of your software workloads on HW platforms

However, new challenges needed to be addressed:

- IaaS

- Microservices

- Serverless

- Others

Support for Cloud Apps

- IaaS: virtualization, software defined provisioning and configuration
  - Virtualization limits access to HW-based monitoring, but all is not lost
  - Automation for data collection and storage, local UI for analysis in Cloud

- Serverless / Microservices: containers, orchestration, load balancing
  - Deliver collectors as containers for integration into microservices
  - On-demand bring up of analyzer instances to support Web UI

- New Analysis Capabilities: paradigm shift from static HW to HW choice
  - Optimal selection of Cloud Compute Engines for my workload (perf or $)
  - Tune for target compute engine for perf or $

VTune Profiler Server Usage
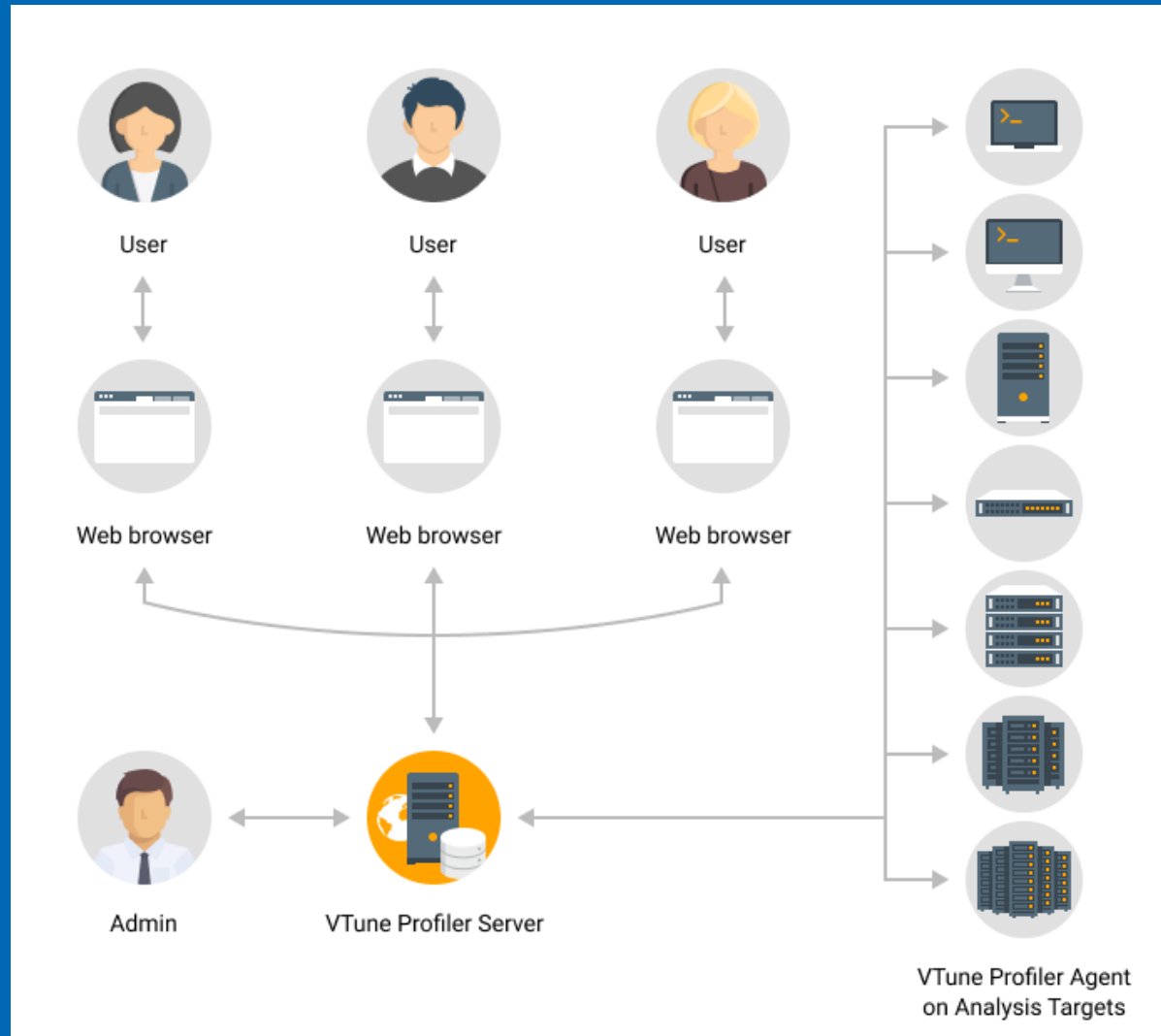
Presentations Today

1   Top-Down Topology-Aware I/O Performance Analysis with Intel® VTune™ Profiler

Ilia Kurakin, Software Engineer, Intel Corporation
9:00 AM - 9:45 AM

2   Is your Code GPU Offload Ready?

Cory Levels, Technical Consulting Engineer, Intel Corporation
09:45 AM – 10:30 AM

3   Debug your Threading and Memory Errors

Kevin O'leary, Lead Technical Consulting Engineer, Intel Corporation
10:30 AM – 11:00 AM

4   Design and Tune your Applications for GPU

Jennifer DiMatteo, Software Engineer, Technical Consulting Engineer, Intel Corporation
Cory Levels, Technical Consulting Engineer, Intel Corporation
11:00 AM – 12:00 PM

Conclusion

oneAPI programming model is an open specification that is gaining momentum in address challenges of multiple HW architectures

- oneAPI toolkits from Intel provide product quality implementation

Intel performance analyzers bring the industry leading capabilities to oneAPI programming model

Furthermore, we are expanding Intel performance analyzers to address new software architectures and Cloud environments

- Analyzer server architecture, Web UI, and automation support for monitoring

SPDK, PMDK, Intel® Performance Analyzers | **Virtual Forum**