



Intel HW roadmap and Architecture Specifics

OneAPI workshop with EuroCC/Castiell, February 16-17, 2022

Dr. Jean-Laurent Philippe, Intel EMEA HPC Director

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more at [www.Intel.com/PerformanceIndex](https://www.intel.com/PerformanceIndex)

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.

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

Includes the effect of Intel Thermal Velocity Boost, a feature that opportunistically and automatically increases clock frequency above single-core and multi-core Intel Turbo Boost Technology frequencies based on how much the processor is operating below its maximum temperature and whether turbo power budget is available. The frequency gain and duration is dependent on the workload, capabilities of the processor and the processor cooling solution.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

All product plans and roadmaps are subject to change without notice.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Agenda

- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update

Agenda

- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update



Front End

Out of Order Engine

Scalar Engine

Vector Engine

Memory Subsystem

Intel's **New** Efficient x86 Core Microarchitecture

Designed for throughput, enabling scalable multi-threaded performance for modern multi-tasking

Optimized for power and density efficient throughput with:

Deep Front-End

with on-demand length decode

Wide Back-End

with many execution ports

Optimized Design

for latest transistor technologies

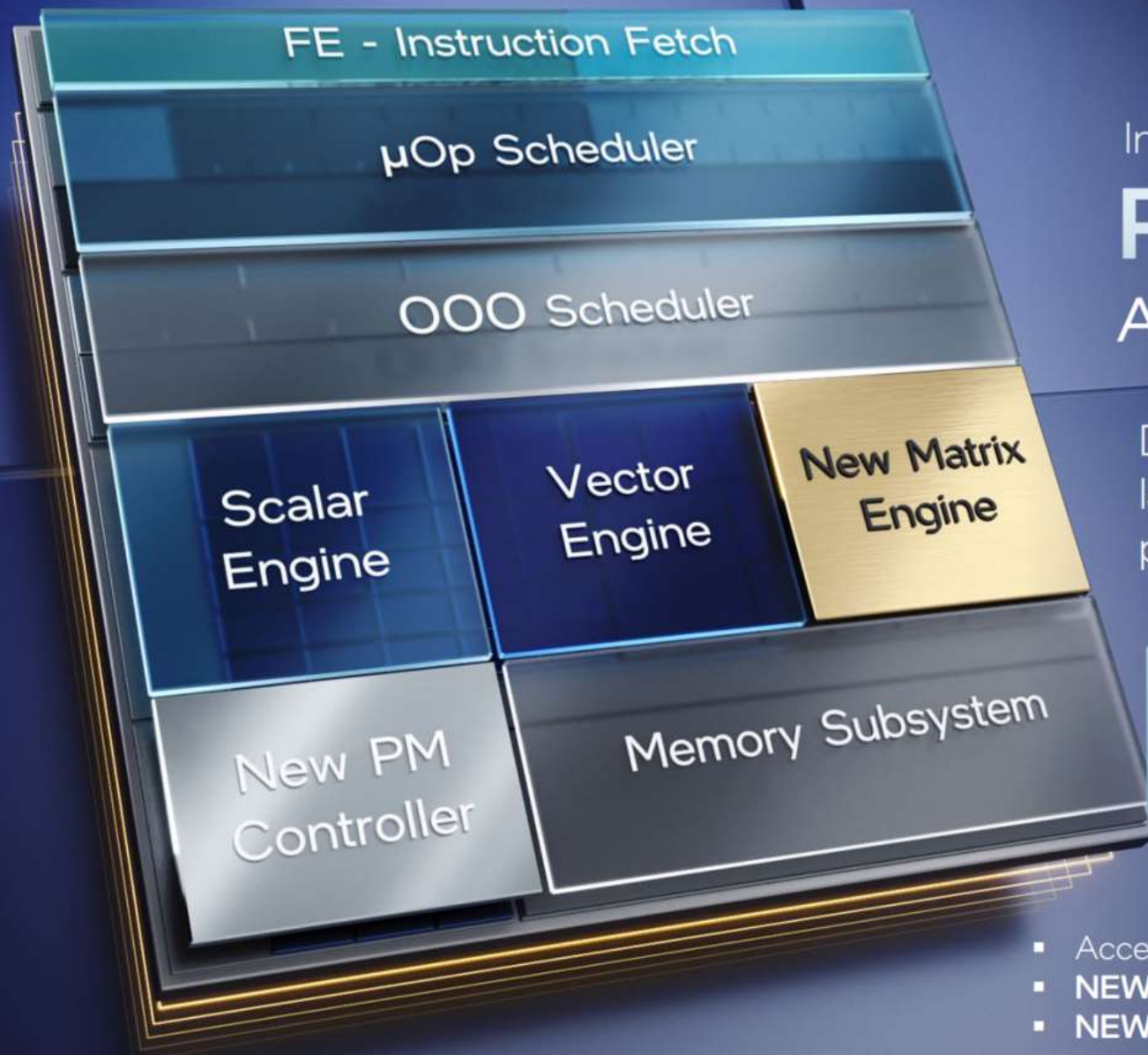


Intel's **New** Efficient x86 Core Microarchitecture

Designed for throughput, enabling scalable multi-threaded performance for modern multi-tasking

Optimized for power and density efficient throughput with:

- Deep Front-End**
with on-demand length decode
- Wide Back-End**
with many execution ports
- Optimized Design**
for latest transistor technologies



Intel's New

Performance x86 Core Architecture

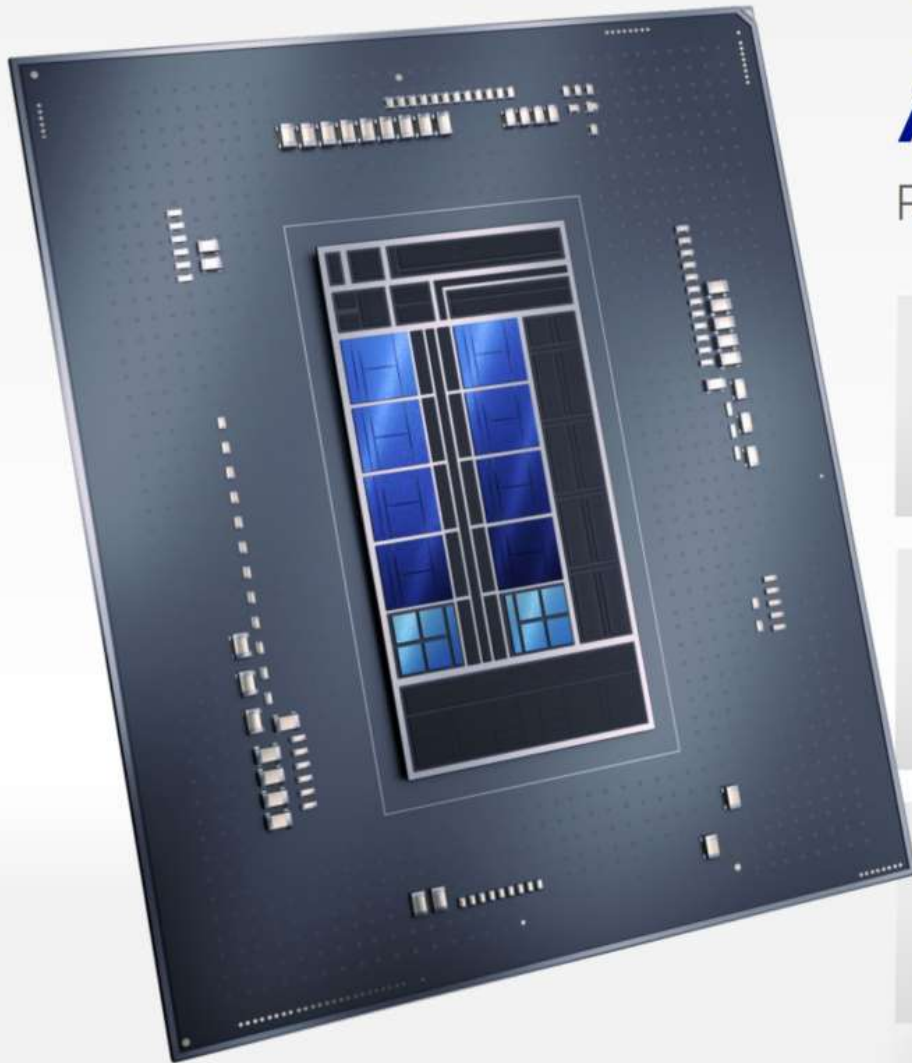
Designed for speed, pushing the limits of low latency and single threaded application performance via:

Wider

Deeper

Smarter

- Acceleration of workloads with large code footprint & large data sets
- **NEW** AI acceleration technology via coprocessor for matrix multiplication
- **NEW** smart PM controller for fine grain power budget management



Introducing Alder Lake

Reinventing Multi Core Architecture

Single, Scalable SoC Architecture

All Client Segments – 9W to 125W – built on Intel 7 process

All-New Core Design

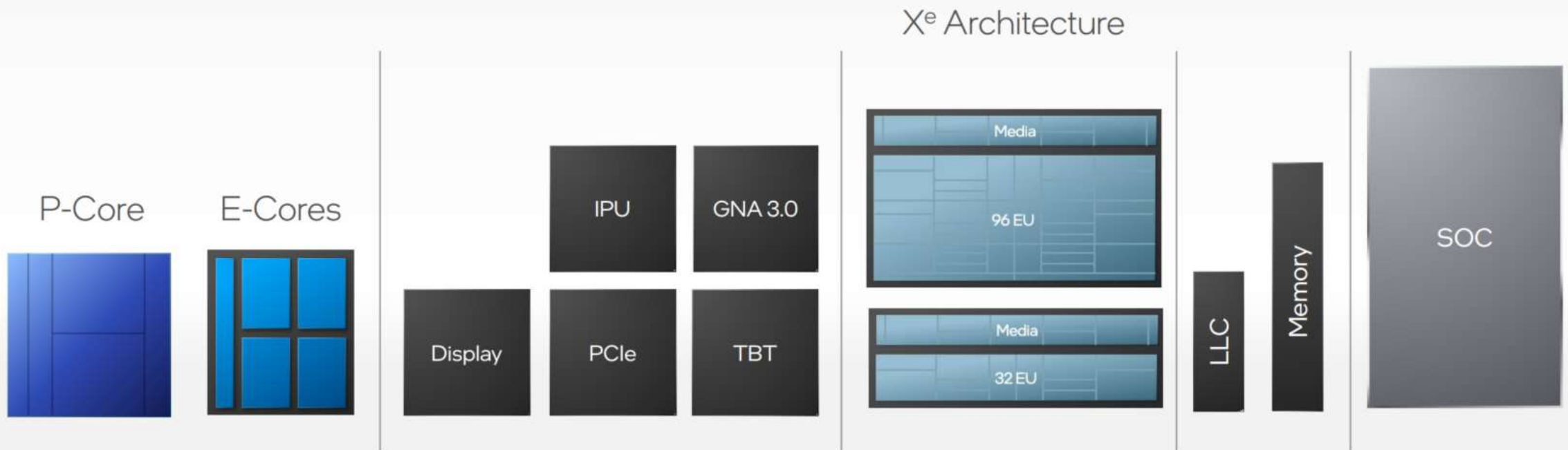
Performance Hybrid with Intel Thread Director

Industry-Leading Memory & I/O

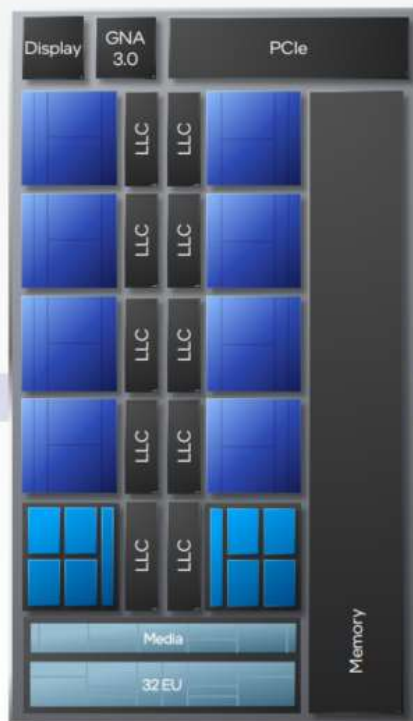
DDR5, PCIe Gen5, Thunderbolt™ 4, Wi-Fi 6E

Alder Lake

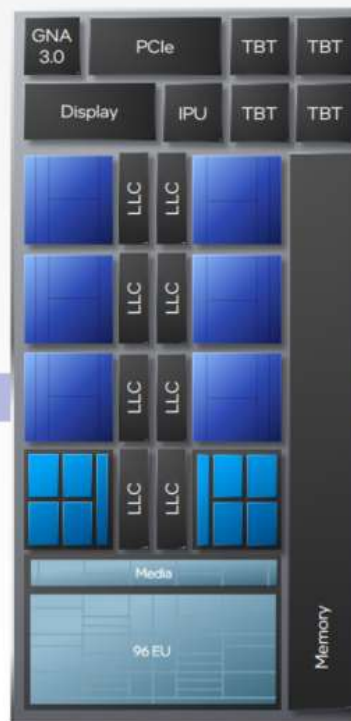
Building Blocks



Desktop



Mobile



Ultra Mobile



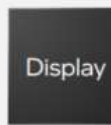
Building Blocks



P-Core



E-Cores



Display



PCIe



TBT



GNA 3.0



IPU



Media

32 EU



Media

96 EU



LLC



Memory



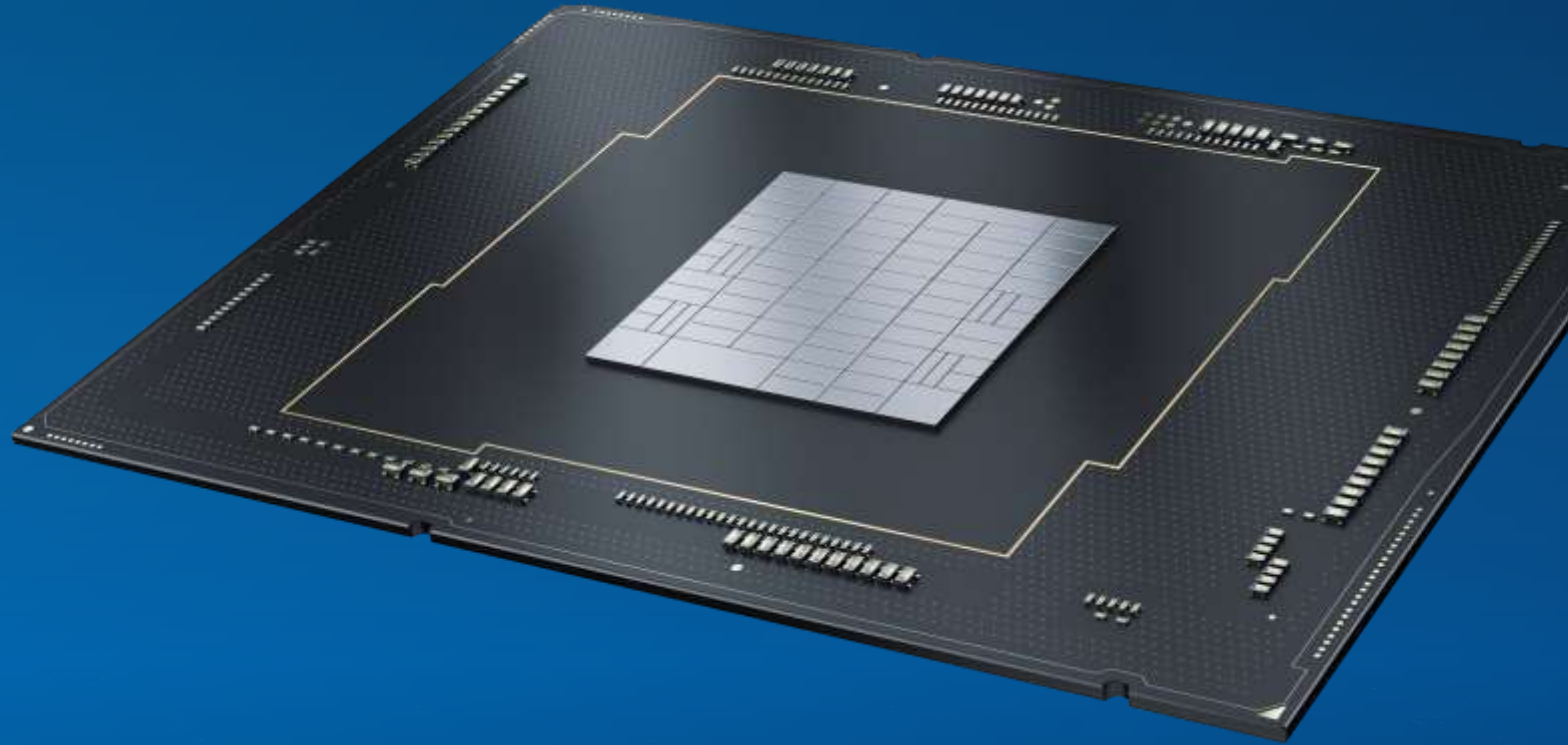
SOC

Agenda

- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update

3rd Gen Intel Xeon Scalable processor

Codenamed "Ice Lake"



Meeting the Demands of Real-World Applications

Updated core architecture, balanced design, fast I/O and increased core counts & memory bandwidth

Flexibility of core and frequency configurations on a single CPU to meet differing application performance requirements

Only mainstream CPU with built-in acceleration for both AI and HPC workloads

20+

years of ecosystem enablement and investment means most HPC applications are optimized for Intel® Xeon® Scalable processors



Easy to deploy on-prem or in Cloud, supported by an unmatched portfolio of libraries, tuners and compilers along with Intel validated solutions

Innovative large capacity system memory and storage architecture for fine grain data access and high performance

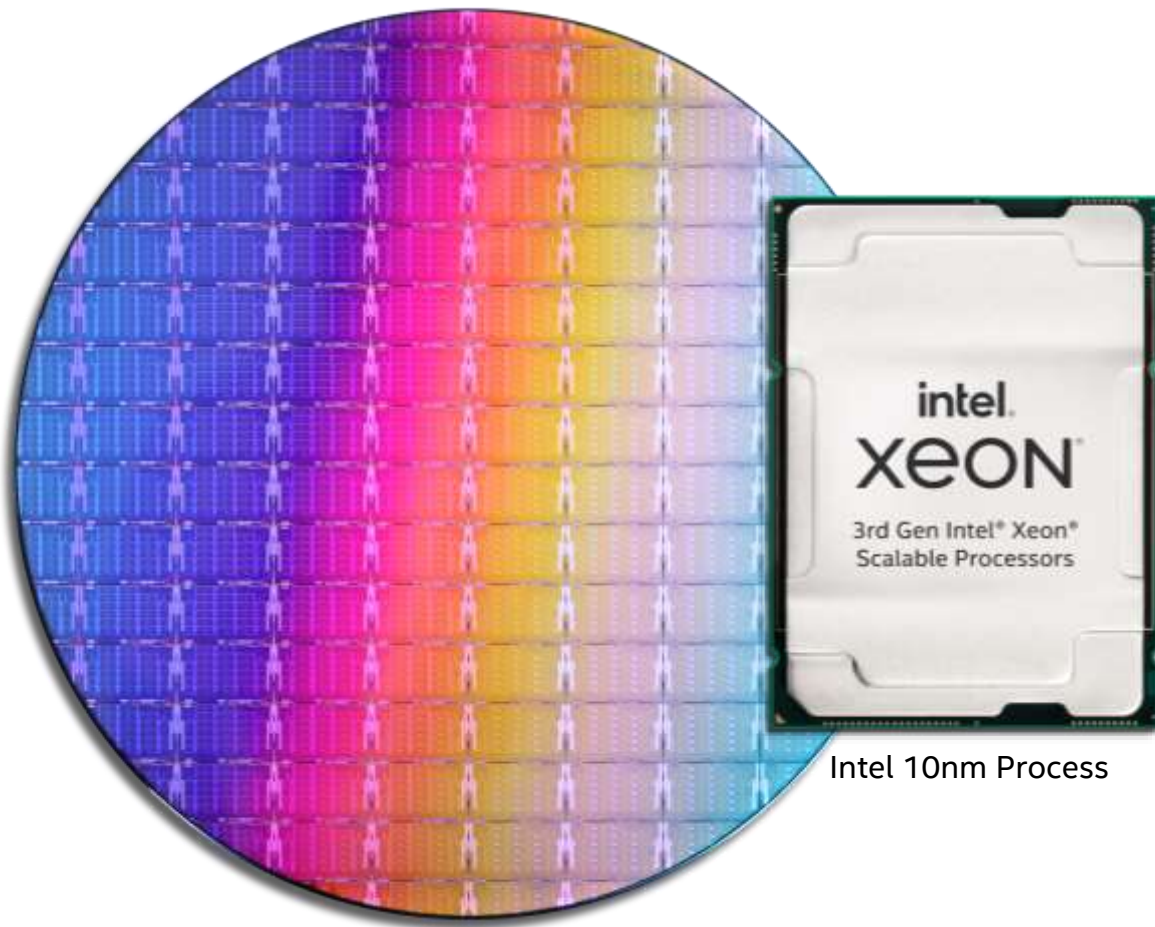
High core, high frequency SKU designed for liquid-cooled systems

Built-in set of new instructions and enhanced security capabilities that help address current and future privacy and security concerns

Intel offers flexible performance to change with your HPC needs

3rd Gen Intel® Xeon® Scalable processors

for HPC in Financial Services



Intel 10nm Process

More & better-performing cores

20% IPC boost, Higher L1 & L2 cache

Higher memory bandwidth

8 DDR4 channels & 3200MT/s

Up to 6TB total memory / socket

with Intel® Optane™ PMem

Faster I/O

via PCIe Gen 4 with 64 lanes/socket

Built-in acceleration & flexibility

AVX-512, DL Boost, Speed Select

Key 3rd Gen Improvements

Designed to Move Faster, Store More, Process Everything

Feature	2nd Gen Intel® Xeon® Scalable Processor	3rd Gen Intel® Xeon® Scalable Processors	Notes
Cores per Socket	4-28	8-40	New Sunny Cove architecture
L1/L2/L3 cache per core	32KB/1MB/1.375MB	48KB/1.25MB/1.5MB	Larger caches to enable fast access to data
Memory Channels and DIMM Speed	6 Up to 2933	8 Up to 3200	Huge boost in memory bandwidth & support for Intel® Optane™ PMem 200
Processor Interconnect: UPI links, speed	2 or 3, 10.4 GT/s	2 or 3, 11.2 GT/s	Improved bandwidth between processors
PCIe lanes per socket	PCIe 3.0, 48 Lanes (x16, x8, x4)	PCIe 4.0, 64 lanes (x16, x8, x4)	2x bandwidth and more PCIe lanes to support new Gen 4 SSD, Ethernet and other adjacencies
Workload Acceleration Instructions	AVX-512 VNNI DDIO	AVX-512, VNNI, DDIO vAES, vPCLMULQDQ, VPMADD52, VBMI, PFR, Crypto, SHA extensions, TME, SGX	Enable new capabilities and speedup performance
Platform Adjacencies		Intel® Optane™ PMem 200 series, Intel® Optane™ P5800X SSD, Intel DC P5510 SSD, Intel E810-C ethernet adapter	

3rd Gen Intel® Xeon® Scalable Processors (Two-socket platform)



Performance made flexible

Up to
1.53x
higher geomean
HPC performance
vs prior gen¹

Up to
40
cores
per processor

Up to
3.7 GHz
Turbo frequency
single-core

Built-in
AI & HPC
Acceleration
Intel® Deep Learning Boost
& Intel® AVX-512

Up to
20%
increased instructions
per clock²

Next-Gen Platform

new Up to
6TB
System Memory
Capacity (Per Socket)
DRAM + PMem

new Up to
8CH
DDR4-3200
2 DPC
(Per Socket)

new Up to
2.6x
Memory Capacity Increase vs.
prior gen

new Up to
64
Lanes
PCI Express 4
(Per Socket)

Advanced Security Technologies

new 
Intel Software Guard
Extensions

new 
Intel Platform
Firmware Resilience

new 
Intel Total Memory
Encryption

new 
Intel Crypto
Acceleration

Scalable, Flexible, Customizable

new 
Intel Speed Select
Technology

new 
Intel® Optane™ Persistent
Memory 200 series

1
oneAPI
Optimized
Software

See [108] at www.intel.com/3gen-xeon-config. Results may vary.

Intel® AVX -512

Built-in acceleration and outstanding performance

3rd Gen Intel Xeon Scalable processors are the only data center CPU with 512-bit instruction processing. This wider vectorization speeds computation processes per clock, increasing frequency over the prior generation.

AVX 512 takes advantage of the processors' increased memory bandwidth, new core architecture, improved frequency management. Additionally, 2 *FMA throughput is now available across the Platinum, Gold and Silver skus

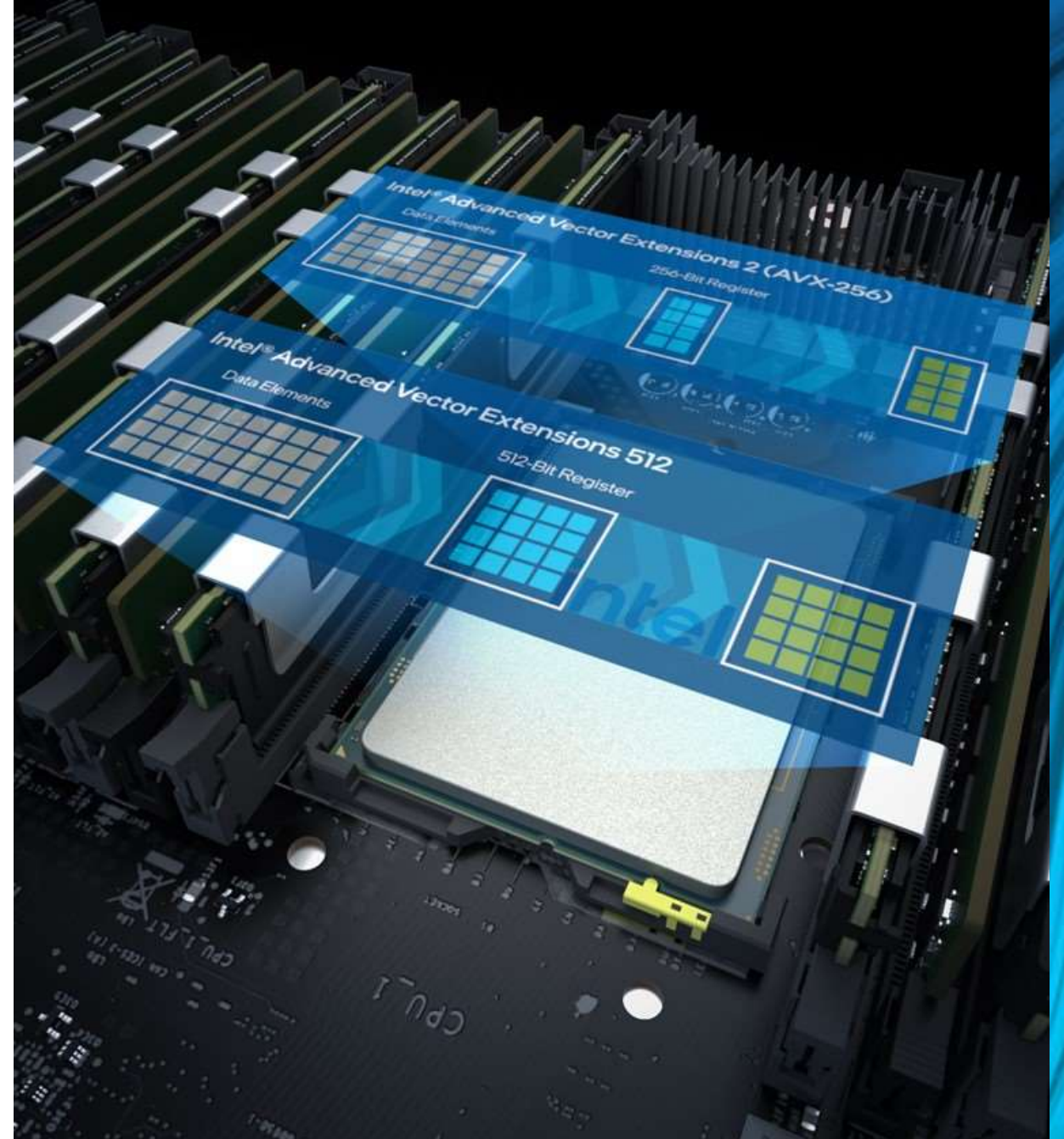
The latest version of Intel AVX-512 is easy to enable by using Intel tools, Intel compilers, Intel Math Kernel Library (Intel MKL), and Intel oneAPI Data Analytics Library. All of these have Intel AVX-512 built right in.

Vectorization can dramatically increase performance for neural network processing, and for other operations that can be parallelized in this way.

2x More
Registers
vs. AVX-256

Learn more at intel.com/avx512

¹ Intel® AVX 2.0 delivers 16 double precision and 32 single precision floating point operations per second per clock cycle within the 256-bit vectors, with up to two 256-bit fused-multiply add (FMA) units. For workloads and configurations visit www.intel.com/PerformanceIndex. Results may vary. Check full sku configurations for feature availability.

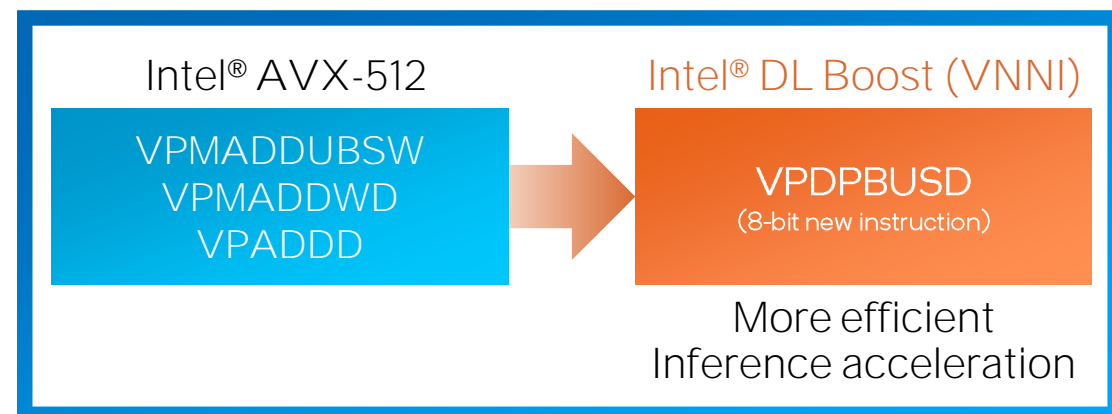


Intel® Deep Learning Boost

A Vector Neural Network Instruction (VNNI)

Extends Intel® AVX-512 to accelerate AI/DL Inference

Problems Solved



Low Precision Integer Operations

End Customer Value

Designed to accelerate AI/Deep Learning workloads (image classification, speech recognition, language translation, object detection and more)



up to **1.56x**
Higher AI
Inference
performance
with Intel DL Boost
vs prior gen¹⁵

Value Pillar



Performance

Software Optimization Examples



TensorFlow

Caffe



Intel®
MKL-DNN

Framework Libraries

Customer Segments



Cloud Service
Providers



Enterprise



Comms Service
Providers

For workloads and configurations see backup and visit www.Intel.com/PerformanceIndex. Results may vary.

Intel® Speed Select Technology



Flexible Performance

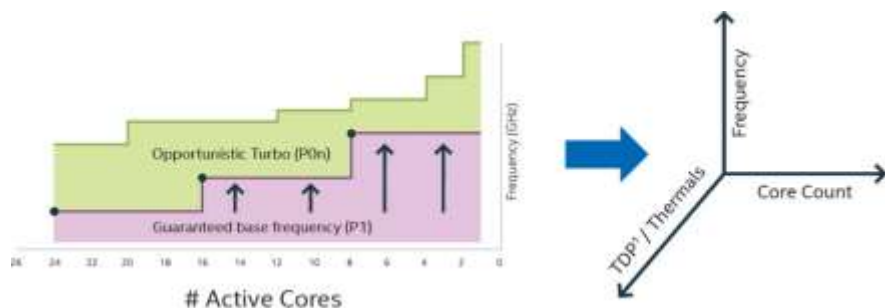
Family of features designed to give you more active and nuanced control over CPU performance

- 3 configurations in a single CPU
- Match your CPU to your workload needs
- Dynamically re-configure processors at runtime
- Enables more efficient systems
- Reduce complexities & costs

Intel® Speed Select Technology Features

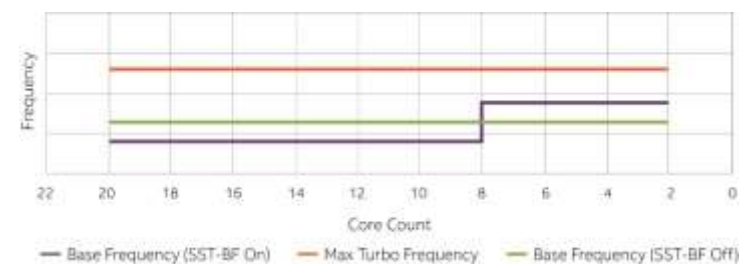
Intel® Speed Select Technology is an umbrella term for a collection of features that provide more granular control over CPU performance

Intel® Speed Select Technology–Performance Profile (Intel® SST-PP, formerly Speed Select)

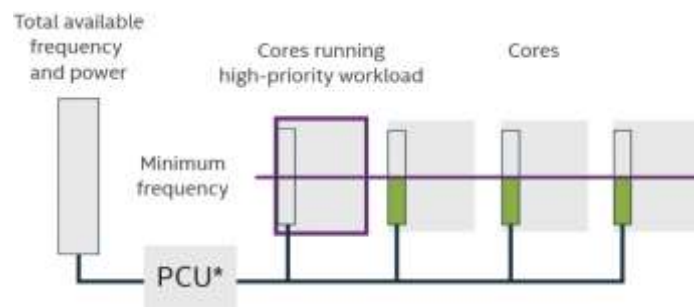


Available in
3rd Gen Intel® Xeon®
Scalable Processor
2S Only

Intel® Speed Select Technology–Base Frequency

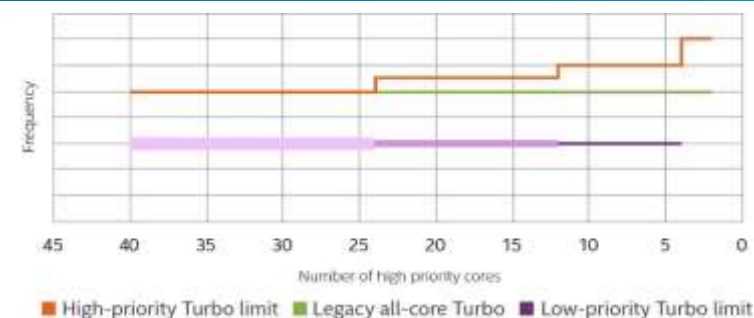


Intel® Speed Select Technology–Core Power



Available in
3rd Gen Intel® Xeon®
Scalable Processor

Intel® Speed Select Technology–Turbo Frequency

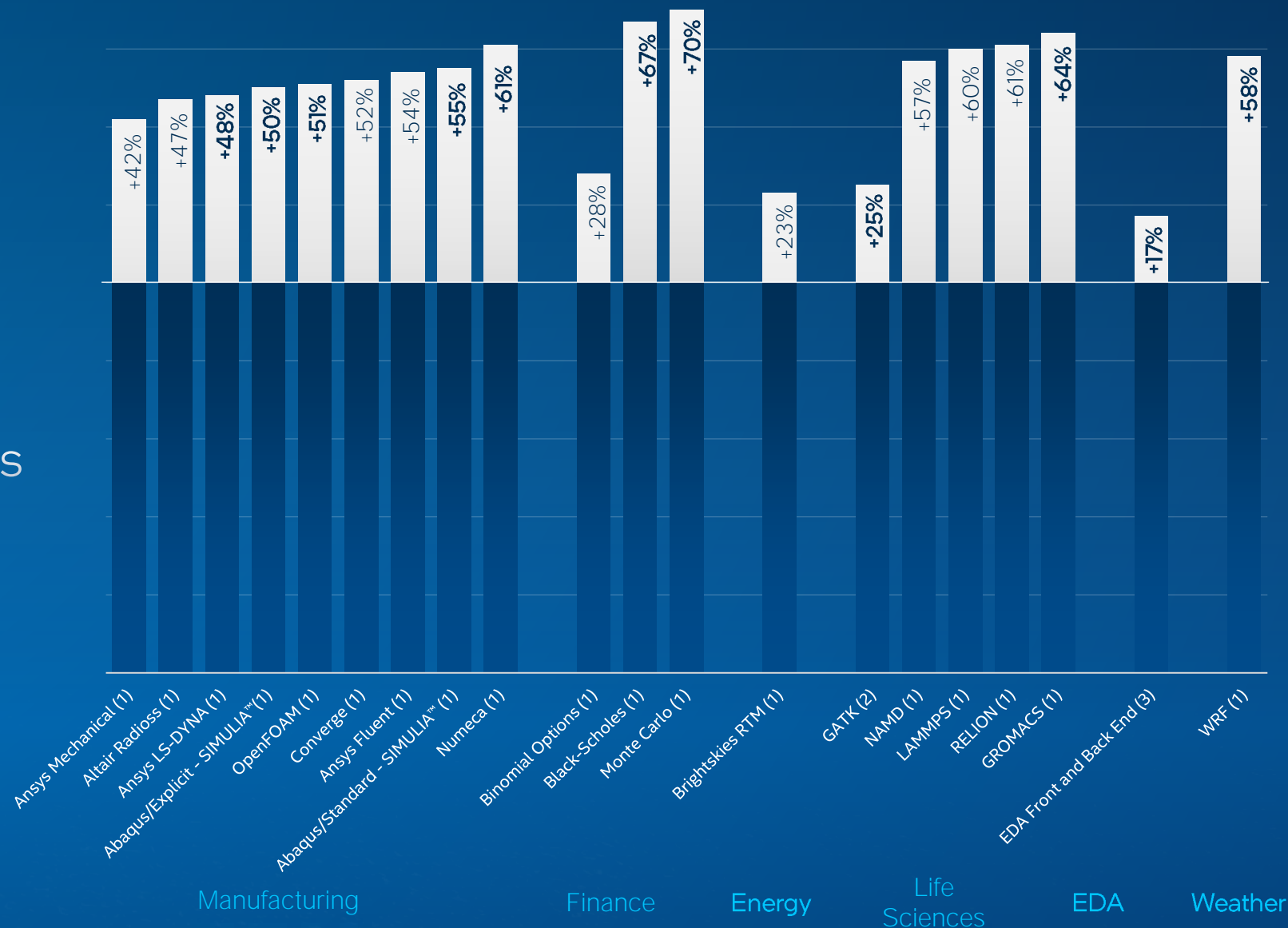
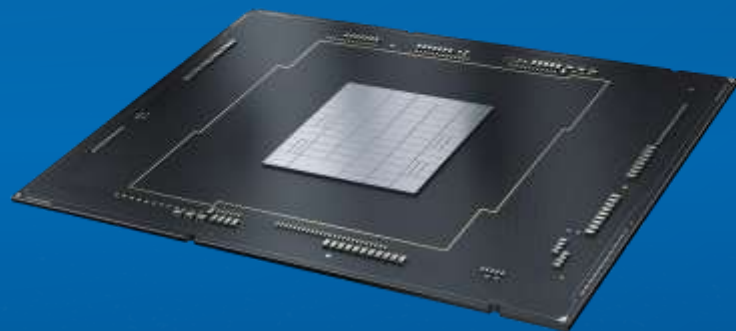


* Power control unit (PCU)

3rd Gen Xeon Processors

+49% Geomean

**Generational
performance leap**
for 20 critical HPC workloads



3rd Gen Xeon Processors

+32% Geomean

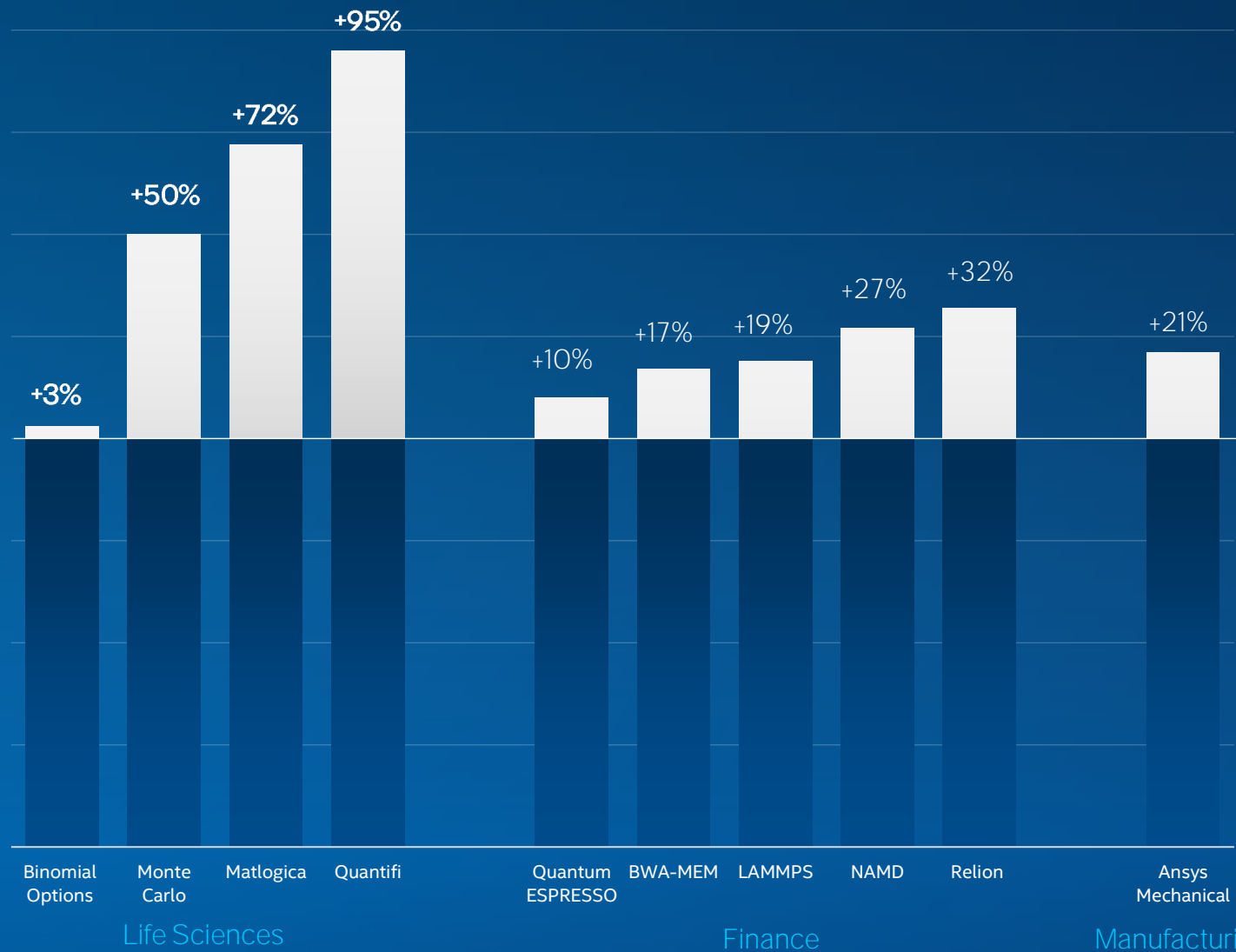
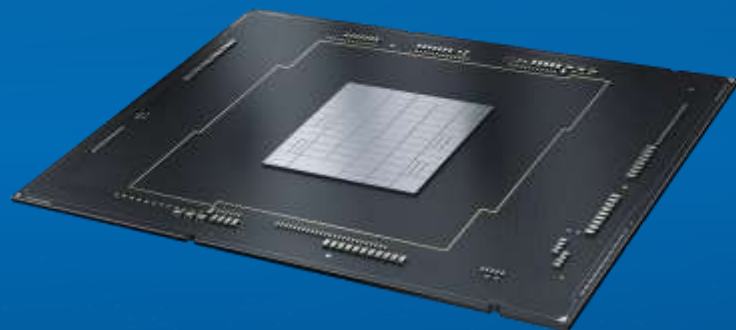
Better performance
across workloads that matter



3rd Gen Intel® Xeon® Scalable processors

vs.

AMD EPYC™ processors

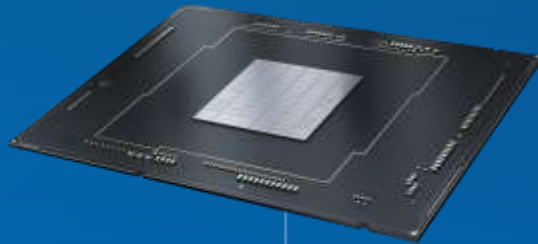


← **Better performance at fewer cores** →
Xeon 8380 (40 cores) vs AMD 7763 (64 cores)

Head-to-Head
Xeon 8358 vs AMD 7543
(both 32 cores)

3rd Gen Xeon Processors

Codenamed "Ice Lake"



2021

Agenda

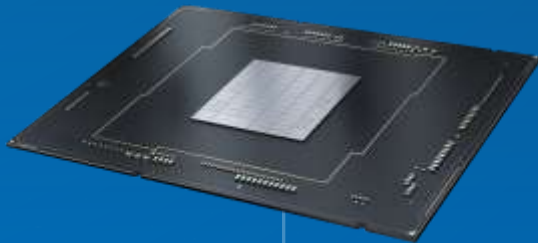
- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update

Next Gen Xeon Scalable Processors

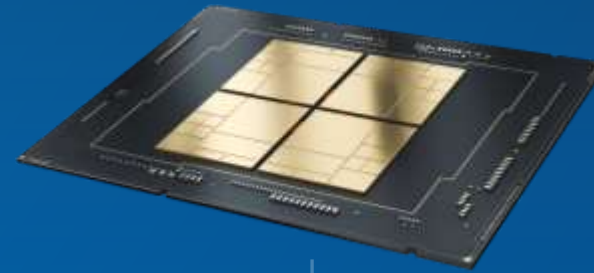
Codenamed Sapphire Rapids

3rd Gen Xeon Processors

Codenamed Ice Lake



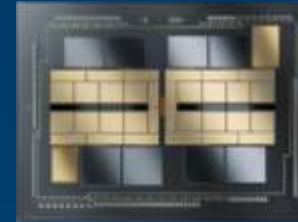
2021



2022

GPU

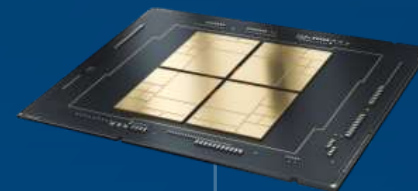
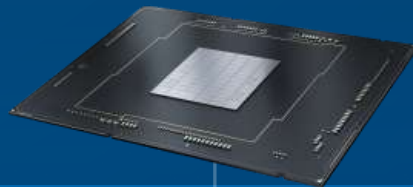
Ponte Vecchio



CPU

3rd Gen Xeon Processors

Next Gen Xeon Processors



2021

2022

GPU

Ponte Vecchio



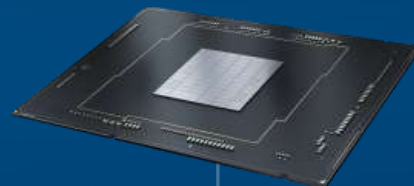
CPU + HBM

Next Gen Xeon Processors
+HBM

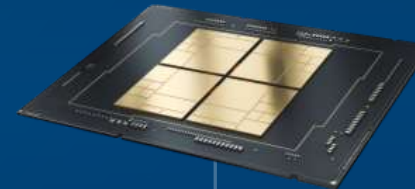


CPU

3rd Gen Xeon Processors



Next Gen Xeon Processors



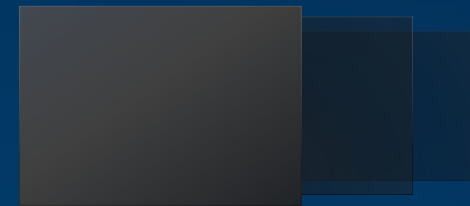
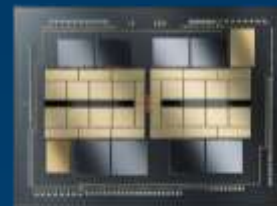
2021

2022

Intel Proprietary – OneAPI workshop with EuroCC/Castiell, February 16-17, 2022

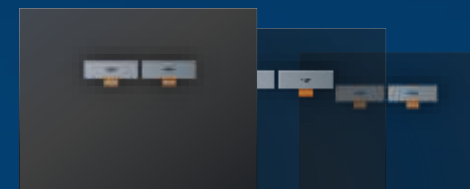
GPU

Ponte Vecchio



CPU + HBM

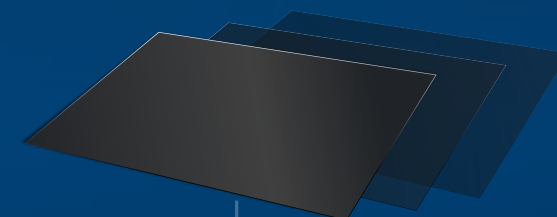
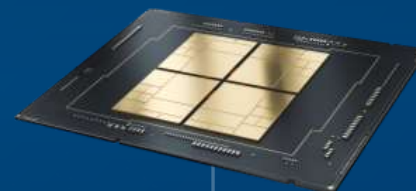
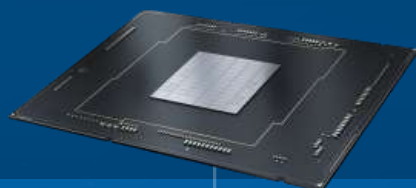
Next Gen Xeon Processors
+HBM



3rd Gen Xeon Processors

Next Gen Xeon Processors

CPU



2021

2022

2023+

Intel Proprietary – OneAPI workshop with EuroCC/Castiell, February 16-17, 2022

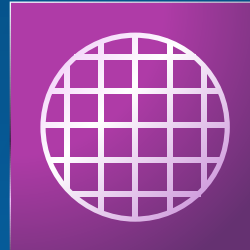
Scalar



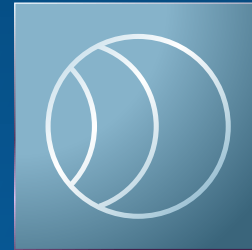
Vector



Matrix



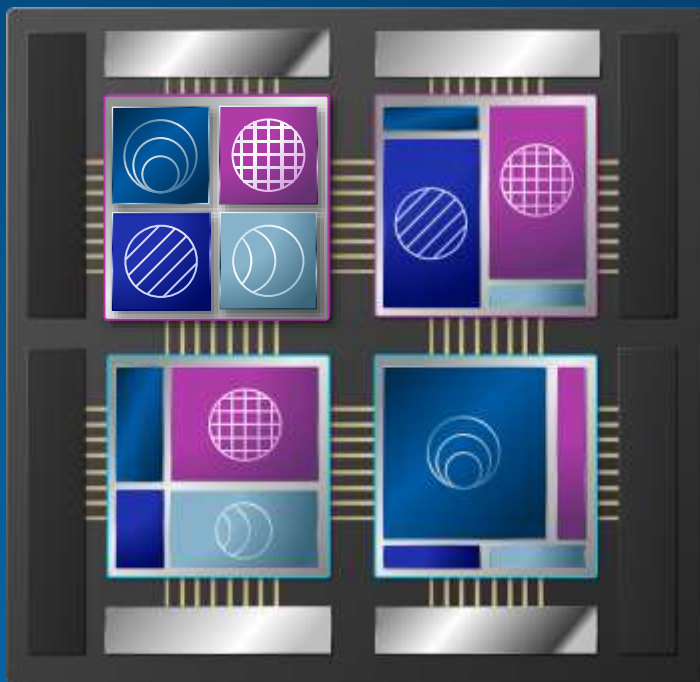
Spatial



Solving the Entire Workload



Solving the Entire Workload



Solving the Entire Workload

Next Gen Intel® Xeon® Scalable Processors
codenamed **Sapphire Rapids**

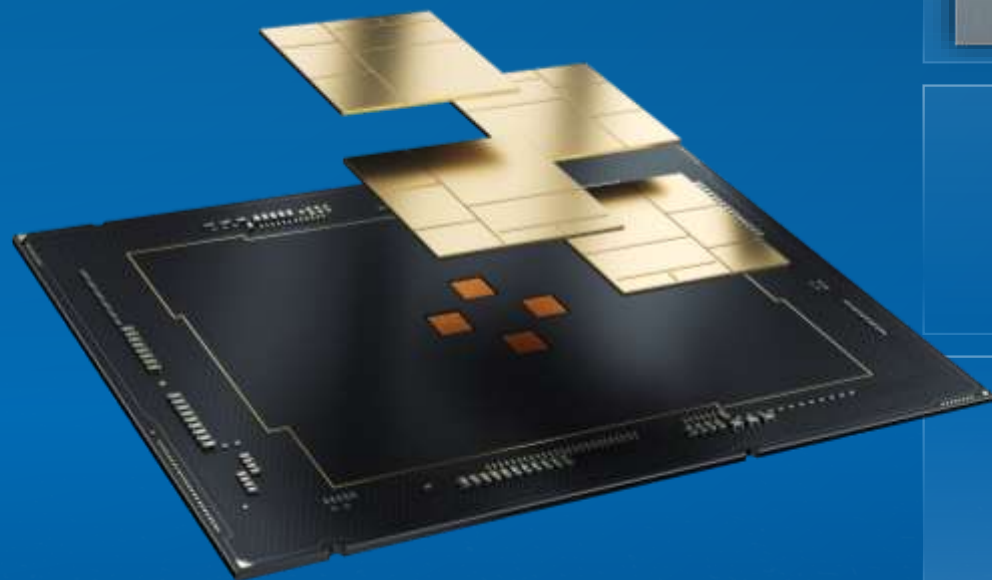
Xe HPC based GPU
Ponte Vecchio

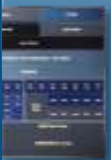





Next Gen Xeon®

Processors

Built for Supercomputing

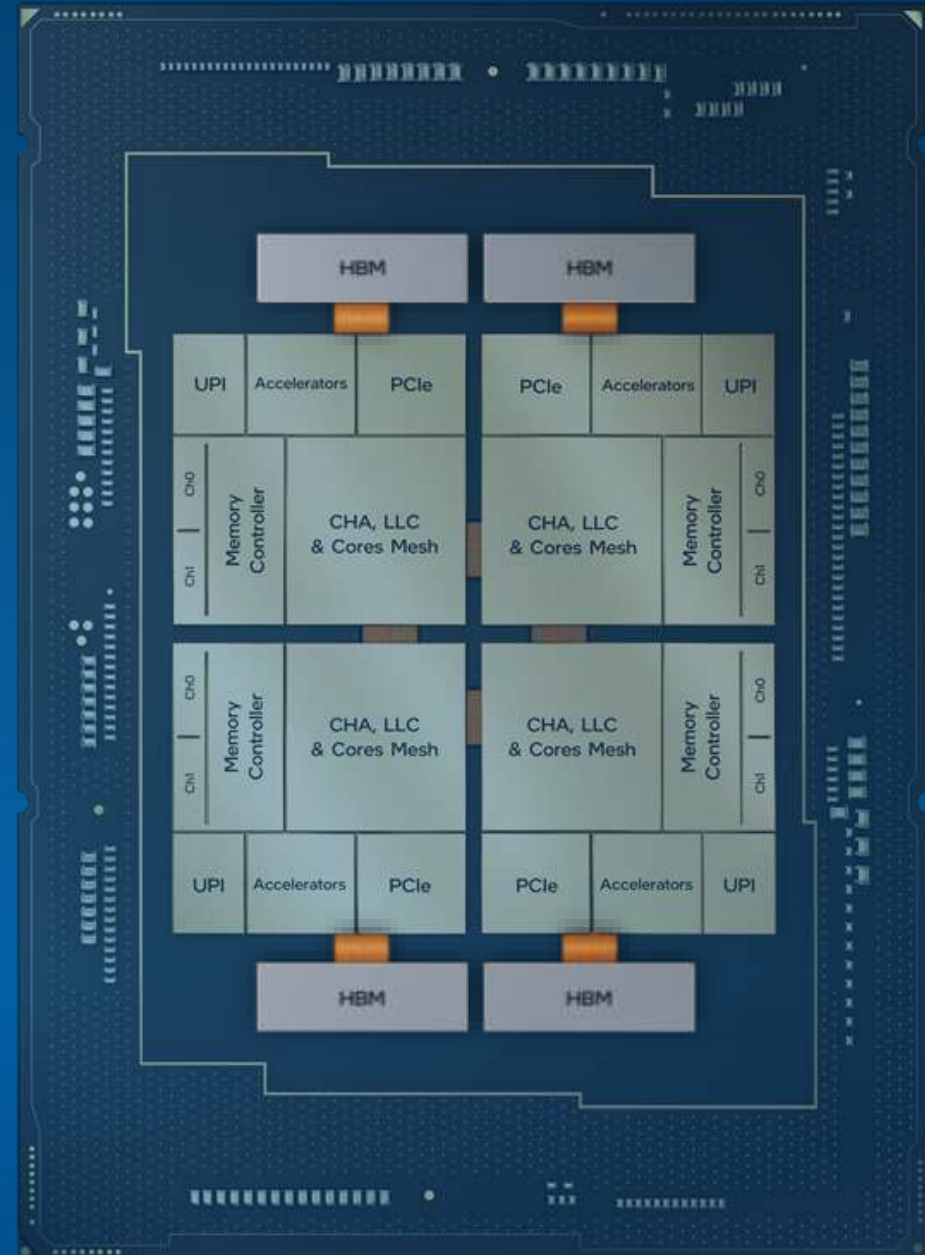


 Performance Cores	Advanced Matrix Extensions	Integrated Acceleration Engines	Compute
 HBM	>100MB Shared LLC	Optane™ & DDR5	Memory
 CXL Compute Express Link 1.1	PCIe Gen 5	UPI 2.0 Improved Multi-socket Scaling	I/O
 EMIB	Multi-Tile	Logically Monolithic	Technology

Next Gen Intel® Xeon® Processors

+HBM

Up to
64 GB
HBM2e

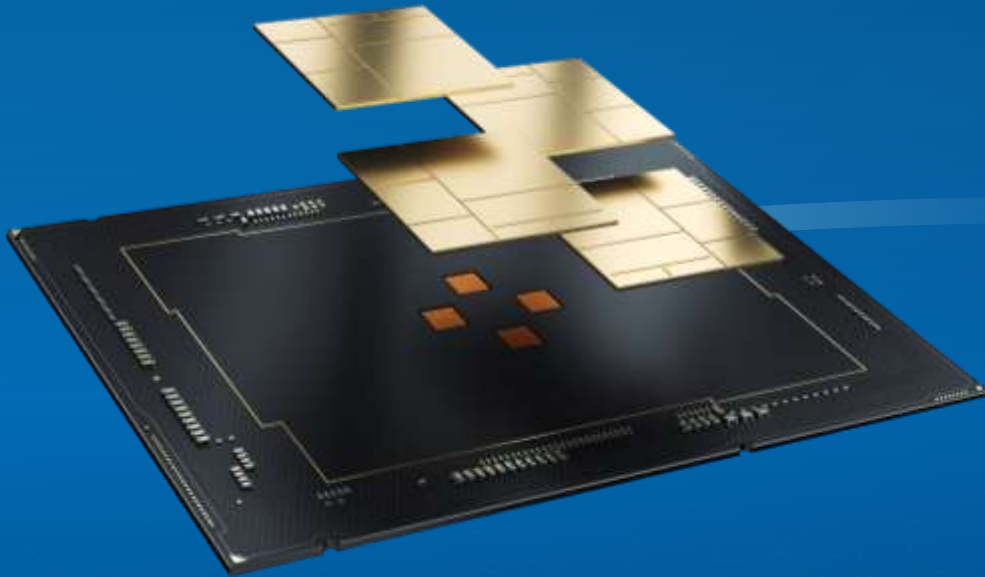


Agenda

- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update

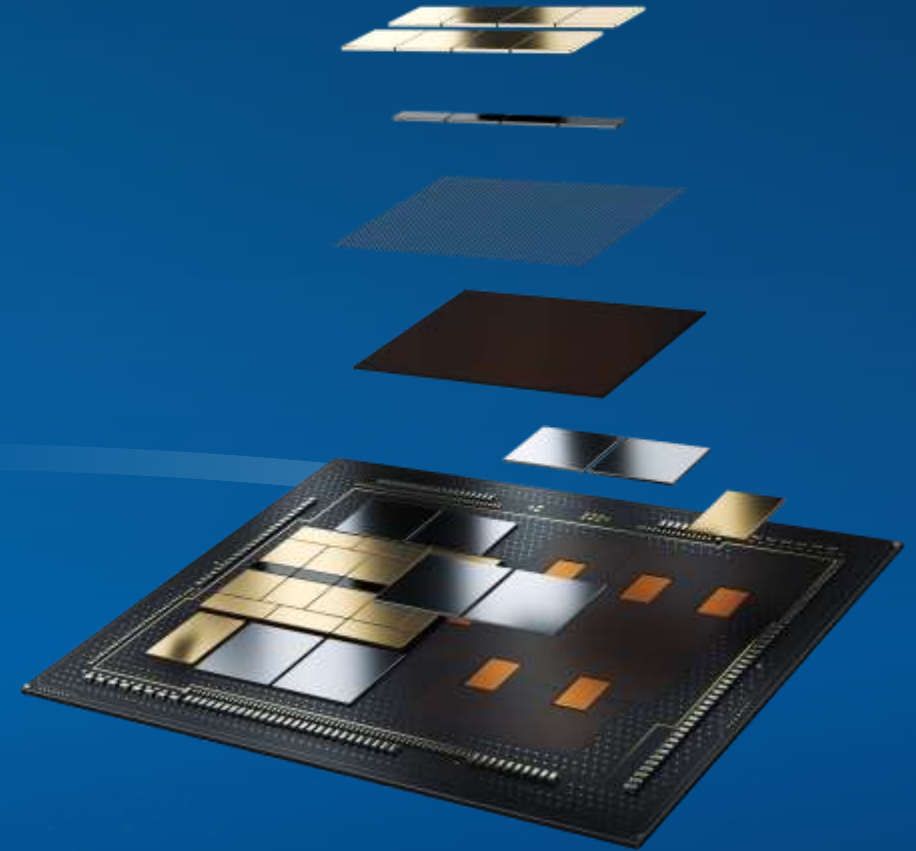
Sapphire Rapids

Next Gen Intel® Xeon® Scalable Processors







Ponte Vecchio

Xe HPC based GPU



Ponte Vecchio

X^e HPC based GPU

Compute	Up to 128 Ray tracing Units	Highest Compute Density socket & node	128 X^e Cores 
Memory	Up to 64MB L1 cache in 2 Stacks	Up to 408MB L2 Cache in 2 Stacks	HBM2e 
I/O	Up to 8 Fully Connected GPUs	PCIe Gen 5	X^e Link High-Speed Coherent Unified Fabric 
Technology	 EMIB	 Foveros	Intel 7 TSMC N5 TSMC N7

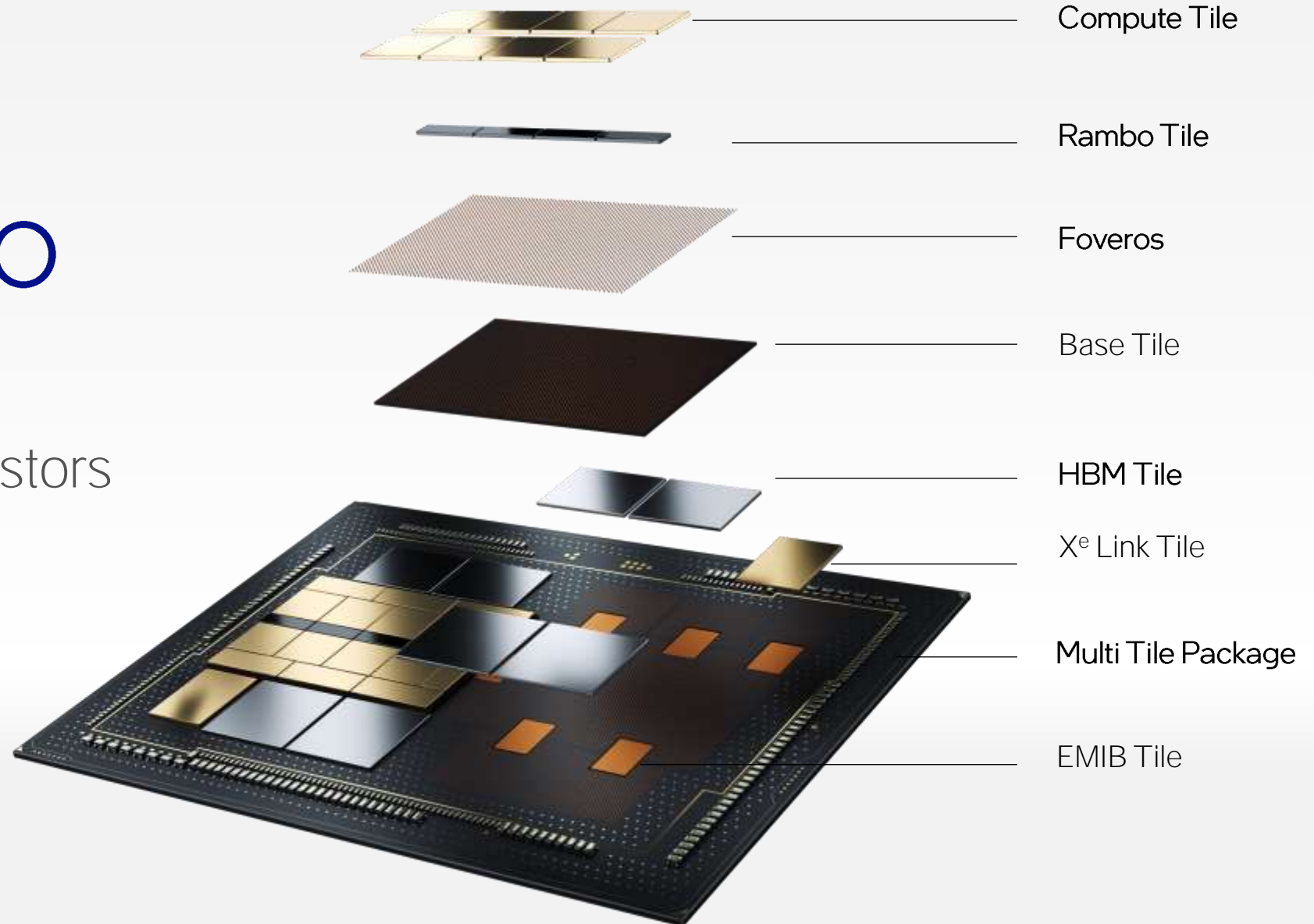


Ponte Vecchio soc

>100 Billion Transistors

47 Active Tiles

5 Process Nodes



Deep Dive Inside the Xe-HPC Micro-Architecture

Built on four hierarchical building blocks

Xe-core

Xe-HPC
Slice

Xe-HPC
Stack

Xe Link

The Xe Core

New ultra-efficient GPU building block for scaling the architecture

8
Vector Engines

512 bit
per engine

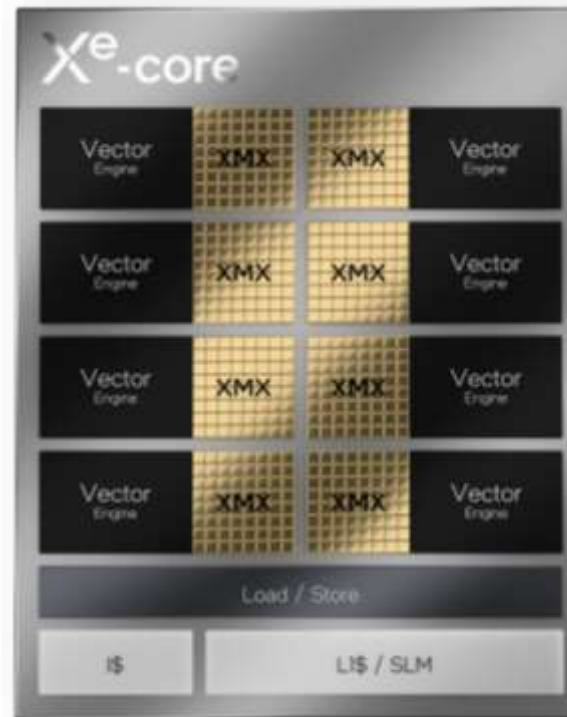
Load / Store
512 B/CLK

Cache
L1\$/SLM (512KB), I\$

8
Xe Matrix Engines
(XMX)

4096 bit
per engine

Driving compelling Op/CLK for critical data formats -
Essential for HPC and AI



Vector
Engine

256 FP64 ops/clk/Xe-core

256 FP32 ops/clk/Xe-core

512 FP16 ops/clk/Xe-core

Matrix
Engine

2048 TF32 ops/clk/Xe-core

4096 FP16 ops/clk/Xe-core

4096 BF16 ops/clk/Xe-core

8192 INT8 ops/clk/Xe-core

The Xe Slice

16 Xe-cores
totaling 8MB of L1 Cache

16 Ray Tracing
Units

that provide fixed-function
computation for Ray
Traversal, Bounding Box
Intersection, and Triangle
Intersection

1 Hardware
Context

that enables execution of
multiple applications
concurrently without
expensive software-based
context switches



The Xe Stack

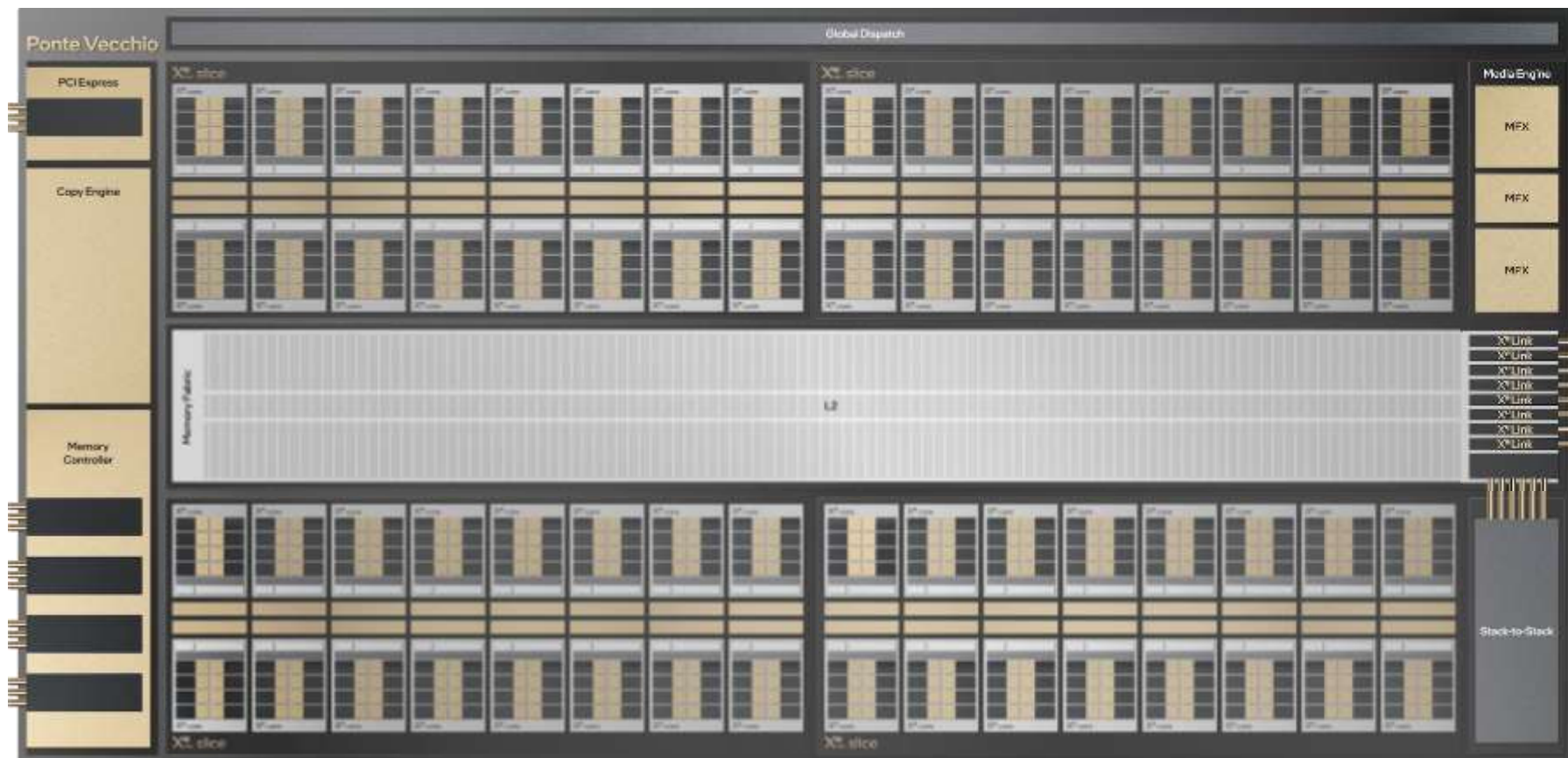
Contains:

4 Slices
which adds up to 64 Xe
Cores, 64 Ray Tracing Units,
and 4 hardware contexts

Massive L2 Cache
(up to 204MB)

1 Media Engine

8 Xe-Links (90GB/s
per link)



Scalable to 2 stacks allowing for multi-stack design

Ponte Vecchio

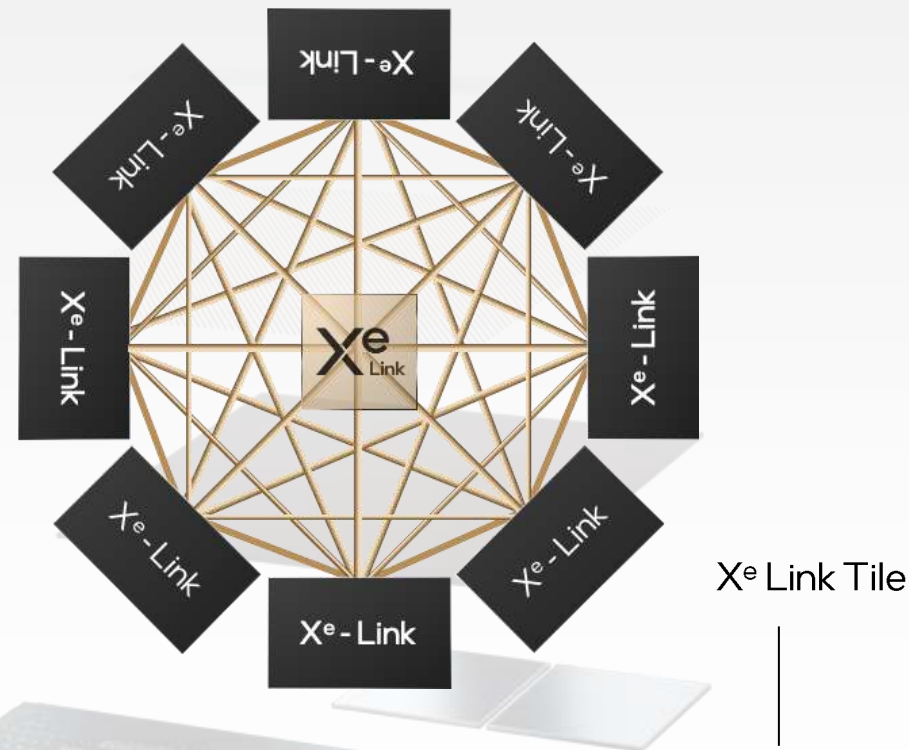
Xe Link Tile

Per Tile
8 Xe Links

8 ports
**Embedded
Switch**

Built on
TSMC N7

Up to
90G
Serdes



Ponte Vecchio

Execution Progress

A0 Silicon Current Status

> 45 TFLOPS

FP32 Throughput

> 5 TBps

Memory Fabric
Bandwidth

> 2 TBps

Connectivity
Bandwidth



Agenda

- Client CPUs, focus on Alder Lake
- Server CPUs, focus on Ice Lake-SP and Sapphire Rapids processors
- Intel HW Discrete Graphics Accelerators, focus on Ponte Vecchio
- Intel FPGA update

Introducing Intel® Agilex™ FPGAs

Intel innovation for ultimate
agility and flexibility

intel.
AGILEX™

High-Performance Compute

10nm Process

Memory-Coherent Intel® Xeon®
Processor Acceleration

Massive Bandwidth

Any-to-Any Integration

Memory, Analog, Logic

Any Node, Supplier, IP

Rapid eASIC Optimization

Any Developer

Intel® Quartus® Prime
Design Tool for
Hardware Developers

OneAPI for Software
Developers

Intel® Agilex™ FPGA – Strong Out of the Gate



~2X Better fabric performance per watt compared to competing 7 nm FPGAs

Up to **400 Gbps Ethernet**
Industry's highest data rate transceivers (116 Gbps)

intel
AGILEX™

intel
Quartus® Prime
Design Software

intel
INTEL-AGILEX™
QWMA35K06R2H4ZE2V
1M19P03A07

intel
Agilex™ FPGA

Shipping in Production Today

~2X better fabric performance per watt*
compared to competing 7nm FPGAs enabling flexible, energy-efficient designs for the data center and beyond

Intel® Agilex™ FPGAs

Industry-leading FPGAs offering customers the adaptability and agility to innovate for diverse workload transitions in 5G, Network, Cloud, and Edge

*See [Intel.com/performanceindex](https://www.intel.com/performanceindex) for configuration details

Intel® Agilex™ FPGA Series

F - SERIES

For a wide range of applications

Up to 58G transceivers

PCIe Gen4

DDR4

Quad-Core Arm Cortex-A53
SoC Option

In Production Now

I - SERIES

For high-performance
processor interface and
bandwidth-intensive applications

Up to 116G transceivers

PCIe Gen5

DDR4

Quad-Core Arm Cortex-A53
SoC Option

Coherent attach to Intel® Xeon®
Scalable Processor option

**Sampling Now, New Family Members
Early Access in 2H 2022**

M - SERIES

For compute-intensive applications

Up to 116G transceivers

PCIe Gen5

DDR4, DDR5, LPDDR5, and Intel®
Optane™ persistent memory support

Quad-Core Arm Cortex-A53
SoC Option

Coherent attach to Intel® Xeon®
Scalable Processor option

HBM Option

Early Access Summer 2022

Thank you



Configuration Details: Generational Performance, slide 8 (1 of 2)

For the following applications (8280 vs 8380): NAMD, LAMMPS, RELION, GROMACS, Monte Carlo, Black-Scholes, Binomial, LS-Dyna, Fluent, OpenFOAM, RADIOSS, Converge, Numeca, WRF: See [108] at www.intel.com/3gen-xeon-config. Results may vary.

Ansys Mechanical: Intel 8280; Config Date: 1/21/2021; Platform: Wolf Pass; CPU Details: 2 CPUs per node; Stepping: 6; 28c @ 2.7Ghz; # CPU Sockets: 2; # CPU Cores: 28; CPU Base Frequency: 2.7 GHz; CPU Max Frequency: 4.0 GHz; CPU Base TDP: 205 W; RAM: 192GB, 12x16GB 2933MHz DDR4, Micron 18ASF2G72PDZ-2G9E1; Hard Drive: 480GB Model: ATA INTEL SSDSC2KG48 (scsi); Cluster File System: OPA based Lustre; BIOS: SE5C620.86B.02.01.0012.070720200218; BIOS Settings: HT=on TURBO=ON; Microcode: 0x4002f01; Intel Management: 04.01.04.381; BMC: 2.48. Operating System: CentOS Linux 8.3.2011; Kernel: 4.18.0-240.1.1.el8_3.crt1.x86_64; OFED stack: OPA 10.10.3.1.1, Lustre 2.10.8-default.App Version: 2021 R1; Data collected: 2/2/2021; Build notes: Hyper-Threading Enabled, Turbo Enabled, Intel FORTRAN Compiler 19.0.0; Intel C/C++ Compiler 19.0.0; Intel Math Kernel Library 2020.0.0; Intel MPI Library 2018 Update 3; One thread per core, 52ppn. Intel 8380; Config Date: 1/21/2021; Platform: Coyote Pass; CPU Details: 2 CPUs per node; Stepping: 6; 40c @ 2.3Ghz; # CPU Sockets: 2; # CPU Cores: 40; CPU Base Frequency: 2.3 GHz; CPU Max Frequency: 2.3 GHz; CPU Base TDP: 250 W; RAM: 256GB 16*16GB 3200MT/s DDR4, Hynix HMA82GR7CJR8N-XN; Hard Drive: SSDSC2KG96 960GB; Cluster File System: HDR based Lustre; BIOS: SE5C6200.86B.2021.D40.2103100308; BIOS Settings: HT=on TURBO=ON; Microcode: 0x8d055260; Intel Management: 04.04.03.249; BMC: 2.66; Operating System: CentOS Linux release 8.3.2011; Kernel: 4.18.0-240.1.1.el8_3.crt1.x86_64; OFED stack: Mellanox mlnx-5.1-2.5.8.0-default. App Version: 2021 R1; Data collected: 4/8/2021; Build notes: Intel(R) FORTRAN Compiler 19.0.0; Intel(R) C/C++ Compiler 19.0.0; Intel(R) Math Kernel Library 2020.0.0; Intel(R) MPI Library 2018 Update; cores: 80ppn. Tested by Intel as of October 2021.

Simulia Abaqus/Explicit: Intel Xeon Platinum 8380/2/6 (40c) ; 2.3/2.1/1.8 . RAM 256GB 16*16GB 3200MT/s DDR4, Hynix HMA82GR7CJR8N-XN; SSDSC2KG96 960GB; HDR based Lustre; BIOS SE5C6200.86B.0020.P23.2103261309; HT=ON; TURBO=ON; Microcode 0xd000270; Intel Management Engine: 04.04.04.053; BMC 2.78; OFED stack Mellanox mlnx-5.1-2.5.8.0-default. Intel Xeon Platinum 8280L 2/6 (2.7GHz) 2.7/2.2/1.8. RAM 192GB, 12x16GB 2933MHz DDR4, Hynix HMA82GR7CJR8N-WM; hard drive 480GB Model: ATA INTEL SSDSC2KG48 (scsi); OPA based Lustre; BIOS SE5C620.86B.02.01.0012.070720200218; HT=ON; TURBO=ON; Microcode 0x4002f01. Abaqus2021HF4. MKL 2020 Update 2. Tested by Intel as of October 2021.

Simulia Abaqus/Standard: Intel Xeon Platinum 8380/2/6 (40c) ; 2.3/2.1/1.8 . RAM 256GB 16*16GB 3200MT/s DDR4, Hynix HMA82GR7CJR8N-XN; SSDSC2KG96 960GB; HDR based Lustre; BIOS SE5C6200.86B.0020.P23.2103261309; HT=ON; TURBO=ON; Microcode 0xd000270; Intel Management Engine: 04.04.04.053; BMC 2.78; OFED stack Mellanox mlnx-5.1-2.5.8.0-default. Intel Xeon Platinum 8280L 2/6 (2.7GHz) 2.7/2.2/1.8. RAM 192GB, 12x16GB 2933MHz DDR4, Hynix HMA82GR7CJR8N-WM; hard drive 480GB Model: ATA INTEL SSDSC2KG48 (scsi); OPA based Lustre; BIOS SE5C620.86B.02.01.0012.070720200218; HT=ON; TURBO=ON; Microcode 0x4002f01. Abaqus2021HF4. MKL 2020 Update 2. Tested by Intel as of October 2021.

Brightskies Reverse Time Migration: Intel Xeon Platinum 8280L CPU @ 2.70GHz; Platform/Motherboard: S2600WFD (Wolf Pass) J46732-801; BIOS: SE5C620.86B.0X.02.0001.051420190324; RAM: 12x16GB (192GB) DDR4 2666; OS: Ubuntu 20.04.3 LTS; Kernel: 5.4.0-89-generic; Microcode: 0x5003102. Intel Xeon Platinum 8380 CPU @ 2.30GHz; Platform/Motherboard: M50CYP2SB2U (Coyote Pass) K88091-301; BIOS: SE5C6200.86B.0022.D64.2105220049; RAM: 16x32GB (512GB) DDR4 3200; OS: Ubuntu 20.04.3 LTS; Kernel: 5.4.0-89-generic; Microcode: 0xd0002b1. Application version 4.0.0. Build: Intel® oneAPI 2021.4. CBS = x (best score). Tested by Intel as of October 2021.

Configuration Details: Generational Performance, slide 8 (2 of 2)

GATK whole genome sequencing:

- Config 1: Test by Intel as of 11/14/2019. 1 application node and 4 compute nodes. Application node configuration: 2x Intel Xeon Gold 6252 processor (24 cores, 2.10 GHz); 1x Intel® Server Board S2600WFT; 192 GB (12x 16 GB DDR4 2666 MHz, 1DC); 1x 960 GB Intel® SSD D3-S4510 Series (2.5 in SATA 6 Gb/s, 3D2,TLC); 1x 1.6 TB Intel® SSD DC P4610 Series (2.5 in PCIe 3.1 x4, 3D2, TLC); Microcode: 0x500002c, BIOS: SE5C620.86B.02.01.0009.092820190230; CentOS Linux Installation ISO (minimal or full) 7.7 build 1910; Intel® Cluster Runtimes 2019.4; Intel® Cluster Checker 2019.3.5; Intel® Select HPC Solution for RPM packages for EL7 2018.0; OpenHPC 1.3.8. 4x compute nodes configuration: 2x Intel Xeon Gold 6252 processor (24 cores, 2.10 GHz); 1x Intel Server Board S2600WFT; 384 GB (12x 32 GB DDR4 2933MHz); 1x 960 GB Intel SSD D3-S4510 Series (2.5 in SATA 6 Gb/s, 3D2, TLC); 1x 1.6 TB Intel SSD DC P4610 Series (2.5 in PCIe 3.1 x4, 3D2, TLC); Network devices: 1x Intel® C620 Series Chipset Ethernet Connection; Intel® Ethernet Adapter X722 onboard 10 GbE; Microcode: 0x500002c, BIOS: SE5C620.86B.02.01.0009.092820190230; 1x distributed 10 GB Lustre 2.10 ZFS system, 6 OST, 3 OSS, Lnet Router with single 10 GB link for all I/O traffic clients to Lustre servers plus 1X 1.6 TB Intel SSD DC P4610 Series (2.5 in PCIe 3.1 x4, 3D2,TLC).
- Config 2: Test by Intel as of 8/8/2021. 1 front-end node and 4 compute nodes, all using Intel® Server Board M50CYP2SB-003; Front-end node configuration: 2x Intel® Xeon® Gold 6348 processor (28 cores) 2.90 GHz, Intel® Hyper-Threading Technology = ON, Intel® Turbo Boost Technology = ON, total memory 256 GB (16 slots/16 GB/3200 MHz); BIOS version: 22D08; BMC 2.66, SDR 0.31, CPLD 3p0; uCode: 0x0b000280; CentOS Linux installation ISO (minimal or full) 8 build 2011; storage – boot drive 1x Intel® SSD P4610 1.6 TB (3D NAND PCIe 3.1 x4, 3D1, TLC); high-performance network: 1x Intel® Ethernet Converged Network Adapter X550-T2, model X550T2. Compute node configuration: 2x Intel Xeon Gold 6348 processor (28 cores @2.60 GHz,) Intel Hyper-Threading Technology = ON, Intel Turbo Boost Technology = ON, total memory 512 GB (16 slots/32 GB/3200 MHz); BIOS version: 22D08; BMC 2.66, SDR 0.31, CPLD 3p0; uCode: 0x0b000280; storage – Scratch drive: 1x Intel SSD P4610 1.6 TB (3D NAND PCIe 3.1 x4, 3D1, TLC); high-performance network: 1x Intel Ethernet Converged Network Adapter X550-T2 (10 GbE), model X550T2.

EDA front end and back end (7 workloads): Xeon 6248R (24 cores): 3.0GHz/4.0GHz/3.6GHz; HDD 2x 1.2TB SAS; BIOS 3.3-IN001, 02/12/2020; OS SLES12 SP5; kernel 4.12.14-122.7-default. Turbo enabled, Hyperthreading disabled. Xeon 6342 (24 cores): 2.8GHz/3.5GHz/3.3GHz; HDD 2x 2.4TB SAS; BIOS 1.0-IN001, 03/12/2021; OS SLES12 SP5; kernel 4.12.14-122.60-default. Turbo enabled, Hyperthreading disabled. Tested by Intel as of April 2021.

Footnotes for page 16

1. Xeon 8380 vs Xeon 8280
2. Xeon 6348 vs Xeon 6252
3. Xeon 6342 vs 6248R

Configuration Details: Competitive Performance, slide 9 (1 of 2)

Binomial Options: 2S Xeon® Platinum 8380 (40C, 2.3GHz, 270W): Intel “Coyote Pass”; 256GB, 16x16GB 3200MHz DDR4; CentOS Linux 8.3.2011, 4.18.0-240.1.1.el8_3.crt1.x86_64, SE5C6200.86B.0021.D40.2101090208, 0x261. App Version: v1.0; Build notes: Tools: Intel C Compiler 2020u4, Intel Threading Building Blocks ; threads/core: 2; Turbo: used; Build knobs: -O3 -xCORE-AVX512 -qopt-zmm-usage=high -fimf-domain-exclusion=31 -fimf-accuracy-bits=11 -no-prec-div -no-prec-sqrt. 2S AMD EPYC 7763 (64C, 2.45GHz, 280W): GIGABYTE R282-Z92 server; 512GB, 16x32GB 3200MHz DDR4; NPS4, CTDp=280W, Determinism=Power; Red Hat Enterprise Linux 8.3, 4.18, ucode 0xa001114. App Version: v1.0; Build notes: Tools: Intel C Compiler 2020u4, Intel Threading Building Blocks ; threads/core: 2; Turbo: used; Build knobs: -O3 -march=core-avx2 -fimf-domain-exclusion=31 -fimf-accuracy-bits=11 -no-prec-div -no-prec-sqrt. Workloads tested by Intel and results as of May 2021.

Monte Carlo: See [37] at www.intel.com/3gen-xeon-config. Results may vary.

MatLogica Provided Vector Accelerator Library XVA pricing benchmark: 2S Intel® Xeon® Platinum 8380 @ 2.30GHz 40 cores on Intel platform with 512 GB DDR4 memory (8(1DPC)/32GB/SK Hynix/1.2v/3200 MT/s), HT on, Turbo on, CentOS Linux 8.4, internal Bios SE5C6200.86B.0020.P23.2103261309 Release Date: 03/16/2021, CentOS Linux release 8.4.2105, Metalogical Libraries AADC-demo-2021-10-01-cd0737f-M9s6, Run Instructions: taskset -c \$EXEFILE \$CONFIG \$INPUT file (for Intel® XEON® avx512 CONFIG=512 used), export OMP_NUM_THREADS=1, 2 thread/core, optimized with AVX512 and Metalogical AADC libraries, test by Intel on 10/14/2021. 2S AMD EPYC™ 7763 @ 2.45GHz 64-Core Processor on GIGABYTE R282-Z92-00 with 512 GB DDR4 memory (8(1DPC)/32GB/SK Hynix/1.2v/3200 MT/s), HT on, Turbo on, Bios: GIGABYTE M07 Release Date: 09/03/2021, CentOS Linux release 8.4.2105, Metalogical Libraries AADC-demo-2021-10-01-cd0737f-M9s6, Run Instructions: taskset -c \$EXEFILE \$CONFIG \$INPUT file (for AMD EPYC avx2 CONFIG=256 used), export OMP_NUM_THREADS=1, 2 threads/core, optimized with AVX2 and Metalogical AADC libraries, test by Intel on 10/14/2021.

Quantifi Credit Option Pricing AI Inference: 2S Intel Xeon 8380 CPU @ 2.30 GHz (40 cores/processor), Turbo ON, HT ON. 512GB DDR4-3200. CentOS Linux Version 8. BIOS Version: SE5C6200.86B.0022.D08.2103221623. BIOS Release Date: 03/22/2021. 2S AMD EPYC 7763 @ 2.45 GHz (64 cores/processor). Turbo ON, HT ON. 512GB DDR4-3200. CentOS Linux Version 8. BIOS Version: M06. BIOS Release Date: 07/10/2021. Tested by Intel. Libraries: python 3.8.11, intel-tensorflow 2.6.0. Turbo=ON. All platforms: python Quantifi_Inference.py. 2 threads/core (all platforms). Tested by Intel as of 10-12-2021.

Quantum ESPRESSO AUSURF112, PSIWAT: 2S Xeon® Platinum 8380 (40C, 2.3GHz, 270W): Intel “Coyote Pass”; 256GB, 16x16GB 3200MHz DDR4; CentOS Linux 8.3.2011, 4.18.0-240.1.1.el8_3.crt1.x86_64, SE5C6200.86B.0021.D40.2101090208, 0x261. App Version: 6.3; Build notes: Tools: Intel Fortran Compiler 2021.2, Intel MKL 2021.2, Intel MPI 2021.2, ELPA 2020.11; threads/core: 2; Turbo: used; Build knobs: -O2 -align array64byte -threads -heap-arrays 4096 -xCORE-AVX512 -qopt-zmm-usage=high -fp-model fast=2 -complex-limited-range -assume byterecl -qopenmp. 2S AMD EPYC 7763 (64C, 2.45GHz, 280W): GIGABYTE R282-Z92 server; 512GB, 16x32GB 3200MHz DDR4; NPS4, CTDp=280W, Determinism=Power; Red Hat Enterprise Linux 8.3, 4.18, ucode 0xa001114.. App Version: 6.3; Build notes: Tools: Intel Fortran Compiler 2021.2, Intel MKL 2021.2, Intel MPI 2021.2, ELPA 2020.11; threads/core: 2; Turbo: used; Build knobs: -O2 -align array64byte -threads -heap-arrays 4096 -march=core-avx2 -fp-model fast=2 -complex-limited-range -assume byterecl -qopenmp. Workloads tested by Intel and results as of May 2021.

BWA-MEM2 on Intel Xeon vs AMD EPYC: 2S Intel® Xeon® Platinum 8380 (40C, 2.3GHz, 270W). Intel “Coyote Pass” - 256GB, 16x16GB 3200MHz DDR4; CentOS Linux release 8.4.2105, 4.18.0-240.22.1.el8_3.crt4.x86_64, ucode 0xd000270. 2S AMD EPYC 7763 (64C, 2.45GHz, 280W): GIGABYTE R282-Z92 server - 512GB; 16x32GB 3200MHz DDR4; CentOS Linux release 8.4.2105, 4.18.0-240.22.1.el8_3.crt4.x86_64, ucode 0xa00111d; NPS=1 yields best performance for EPYC. Baseline code: BWA-MEM, v0.7.17; optimized code (OpenOmics): BWA-MEM2, v2.2.1. Datasets used: Reference sequence: version GRCh38; Read dataset: 10M reads from SRR7733443 (151bp each) and HG004 (250bp each). Tested by Intel as of September 2021.

Configuration Details: Competitive Performance, slide 9 (2 of 2)

LAMMPS (Polyethylene, Stillinger-Weber, Tersoff, Water):): 2S Xeon® Platinum 8380 (40C, 2.3GHz, 270W): Intel “Coyote Pass”; 256GB, 16x16GB 3200MHz DDR4; CentOS Linux 8.3.2011, 4.18.0-240.1.1.el8_3.crt1.x86_64, SE5C6200.86B.0021.D40.2101090208, 0x261. App Version: v2020-10-29; Build notes: Tools: Intel MKL 2020u4, Intel C Compiler 2020u4, Intel Threading Building Blocks 2020u4, Intel MPI 2019u8; threads/core: 2; Turbo: used; Build knobs: -O3 -ip -xCORE-AVX512 -qopt-zmm-usage=high. 2S AMD EPYC 7763 (64C, 2.45GHz, 280W): GIGABYTE R282-Z92 server; 512GB, 16x32GB 3200MHz DDR4; NPS4, CTDp=280W, Determinism=Power; Red Hat Enterprise Linux 8.3, 4.18, ucode 0xa001114. App Version: v2020-10-29; Build notes: Tools: Intel MKL 2020u4, Intel C Compiler 2020u4, Intel Threading Building Blocks 2020u4, Intel MPI 2019u8; threads/core: 2; Turbo: used; Build knobs: -O3 -ip -march=core-avx2. Workloads tested by Intel and results as of May 2021.

NAMD: See [36] at www.intel.com/3gen-xeon-config. Results may vary.

RELION: See [38] at www.intel.com/3gen-xeon-config. Results may vary.

Ansys Mechanical 2021 R2 on Endeavor (HDR Fabric): 21% improvement across all SP cases (V21sp-1, V21sp-2, V21sp-3, V21sp-4). Intel® Xeon® Platinum 8358 Processor (32 cores, 48M Cache, 2.60 GHz). RAM 256 GB, 16*16GB 3200 MT/s DDR4, Hynix HMA82GR7CJR8N-XN. HDR based Lustre. BIOS SE5C6200.86B.0020.P23.2103261309. BIOS settings HT=on, Turbo Boost ON, SNC 2. Microcode 0xd000270. Intel Management Engine 04.04.04.053. BMC 2.78. Cent OS 8.4.2105. Kernel 4.18.0-240.22.1.el8_3.crt2.x86_64. OFED stack Mellanox mlnx-5.1-2.5.8.0- default. AMD EPYC™ 7543 (32 cores, Total L3 Cache 256MB, 2.8 GHz base clock). RAM 256 GB, 16*16GB 3200 MT/s DDR4, SK Hynix HMA82GR7CJR8N-XN. 240GB ATA Intel SSDSC2KB24. HDR based Lustre. BIOS Ver 2.1 Rev 5.22. Default settings, Multi-threading on, Boost on. NPS4 . Microcode 0xa00111d. BMC 3.1. Cent OS 8.4.2105. Kernel 4.18.0-240.22.1.el8_3.crt2.x86_64. OFED stack Mellanox mlnx-5.1-2.5.8.0- default. Intel MPI 2019u9 used throughout. 2018u3 used where needed. Intel MKL used for Intel platforms: Intel(R) Math Kernel Library Version 2020.0.0 Product Build 20191122. AMD BLIS used for AMD platforms. Intel Compiler Version: 2019.0.0. Intel(R) FORTRAN Compiler Version 19.0.0 (Build: 20190206). Intel(R) C/C++ Compiler Version 19.0.0 (Build: 20190206). Tested by Intel as of October 2021.