
intel®
XEON®
Accelerate with Xeon
January 10

CPU + Accelerators: Differentiated Performance On Real Workloads

4th Gen Intel® Xeon® Scalable processors					Intel® Xeon® CPU Max Series
General Purpose Compute	Artificial Intelligence	Network 5G vRAN	Networking & Storage	Data Analytics	HPC
53%	Up to 10x	Up to 2x	Up to 2x	Up to 3x	Up to 3.7x
average performance gain*	higher inference and training performance*	capacity for vRAN workloads at same power envelope*	higher data compression with 95% fewer cores*	higher performance*	on memory-bound workloads**

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.
*4th Gen Intel Scalable Processor vs. 3rd Gen Intel Xeon Scalable processors
** Intel Xeon CPU Max Series vs. Intel Xeon 8380

CPU + Accelerators: Groundbreaking Efficiency

Higher Performance per Watt

2.9x

average improvement of perf/watt with built-in accelerators*

Lower Power Bills

up to **70W**

power savings per CPU with Optimized Power Mode

Lower TCO More Sustainable

55%

lower TCO and power consumption while reducing 524K kg of CO_2 emissions*

AI Real Time Inferencing workload, ResNet50

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.
*for selected workloads.
*4th Gen Intel® Xeon® Scalable Processor vs. 3rd Gen Intel Xeon Scalable processors.

Intel's Most Feature Rich Server Platform

4th Gen Intel® Xeon® Scalable Processors and Intel® Xeon® CPU Max Series Processors

New Integrated IP Accelerators

Increased I/O Bandwidth with PCIe 5.0

Compute Express Link (CXL) 1.1

4th Gen Intel Xeon Scalable Processor

Intel Xeon CPU Max Series

Intel® In-Memory Analytics Accelerator (Intel® IAA)

Intel® Data Streaming Accelerator (Intel® DSA)

Intel® Dynamic Load Balancer (Intel® DLB)

Intel® QuickAssist Technology (Intel® QAT)

PCIe 5.0/4.0 Up to 80 Lanes

1 to 8 Socket Support

Intel® In-field Scan **

Intel® Seamless Firmware Update

High Bandwidth Memory Support

8 DMI 4.0 LINKS

Intel® C741 Chipset

Intel® Advanced Matrix Extensions (AMX)

Accelerating Interfacing Architecture (AiA)

Intel® Advanced Vector Extensions for vRAN

Enhanced Intel® Software Guard Extensions

8 Channel DDR5 Up To 4800 MT/s (1DPC)

Up To 4400 MT/s (2DPC)

16 DIMMS per Socket

New RAS features (Enhanced ECC, Error Check & Scrub)

Intel® Ultra Path Interconnect (Up to 4 Links @ 16 GT/s)

Optimized Power Mode

New Integrated ISA Accelerators

HW-Enhanced Security

Increased Memory Bandwidth with DDR5

Increased Inter-Socket Bandwidth with UPI 2.0

Intel® Optane™ Persistent Memory 300 Series

Intel® Agilex™ FPGAs

Intel® Data Center GPU Max Series
Intel® Data Center GPU Flex Series

Intel® Infrastructure Processing Unit (Intel® IPU)

intel XEON Accelerate with Xeon

* Enabled for WorkStation + vPro SKUs ONLY

** Intel In-Filed Scan is a new capability available through select providers in 2023

INTRODUCING

4th Gen Intel® Xeon® Scalable Processors

Workload-First Approach to Innovation, Design and Delivery

Most Built-In Accelerators of any CPU in the Market with Leading Performance in AI, Analytics, Networking, Storage, Security, and HPC

Intel's Most Sustainable Data Center Processor Ever

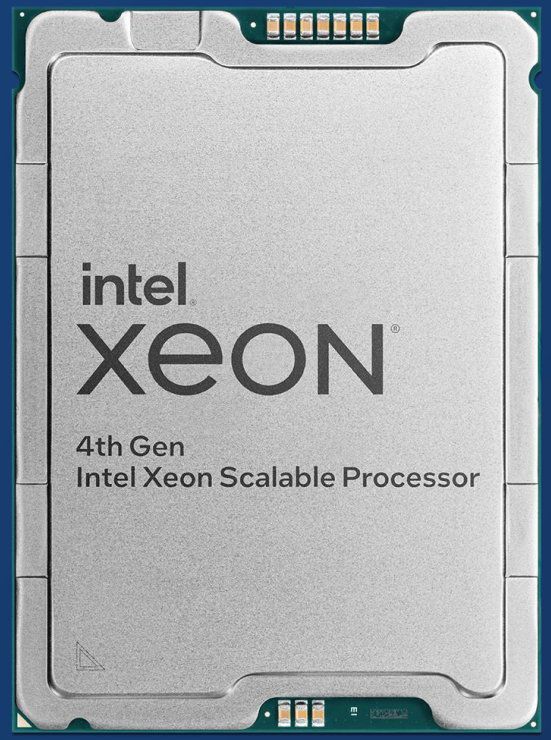
In customer hands and delivering real-world results today

intel XEON Accelerate with Xeon

Built for Better Node & Data Center Performance

Node Performance

- Scalar Performance
- Data Parallel Performance
- Cache & Memory Sub-System Architecture
- Intra/Inter Socket Scaling



Data Center Performance

- Infrastructure & Framework Overhead
- Consolidation & Orchestration
- Performance Consistency
- Elasticity & Efficient Data Center Utilization

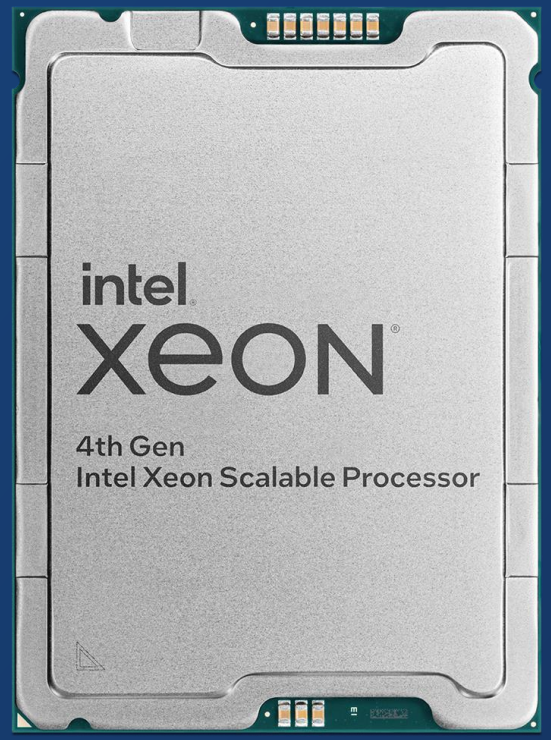
Security-by-Design Foundation

Delivering Better Node & Data Center Performance

New or Enhanced Capabilities

Node Performance

- Intel P-Core Microarchitecture
- 4-Tile & Monolithic Architecture
- ISA and IP Accelerators
- Accelerator Interface Architecture
- Low Jitter Architecture
- Workload Optimized SKUs



Data Center Performance

- Optimized Power Mode
- Platform Monitoring Technology
- Resource Director Technology
- In-field Scan**
- Intel® On Demand
- Seamless Firmware Update

Security-by-Design Foundation

Intel® Software Guard Extensions (Intel® SGX), Intel® Trust Domain Extension (Intel® TDX)*, Intel® QuickAssist Technology (Intel® QAT), Intel® Control-flow Enforcement Technology

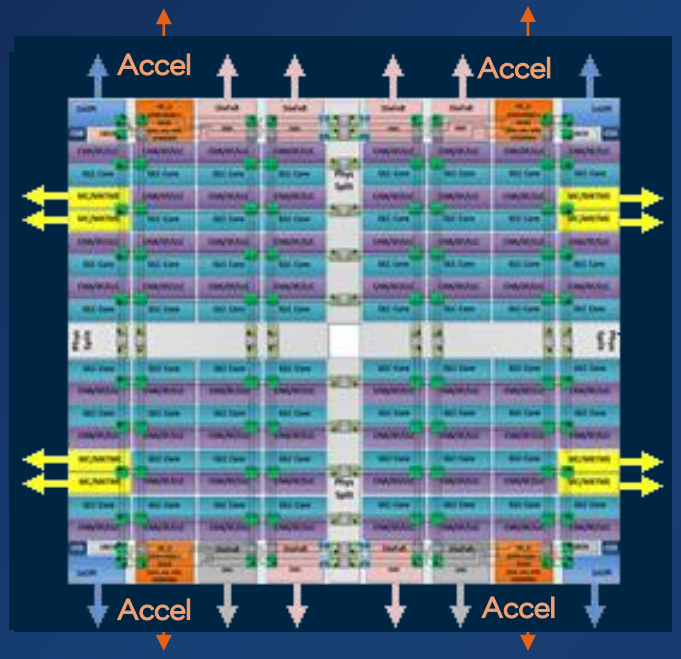
intel XEON Accelerate with Xeon

* Intel TDX is a new capability available through select cloud providers in 2023
** Intel In-Filed Scan is a new capability available through select providers in 2023

Unique Die Packages for Unique Market Needs

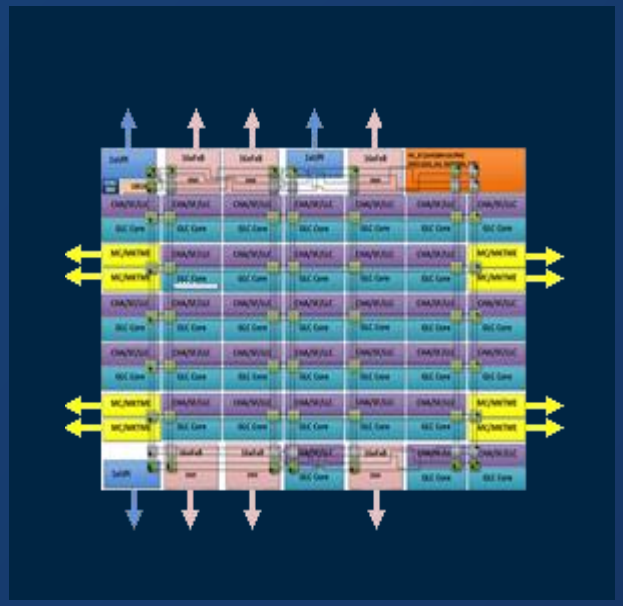
Intel® Xeon® Die Package Architecture

4th Gen Intel Xeon processor - XCC
4 Tile Architecture



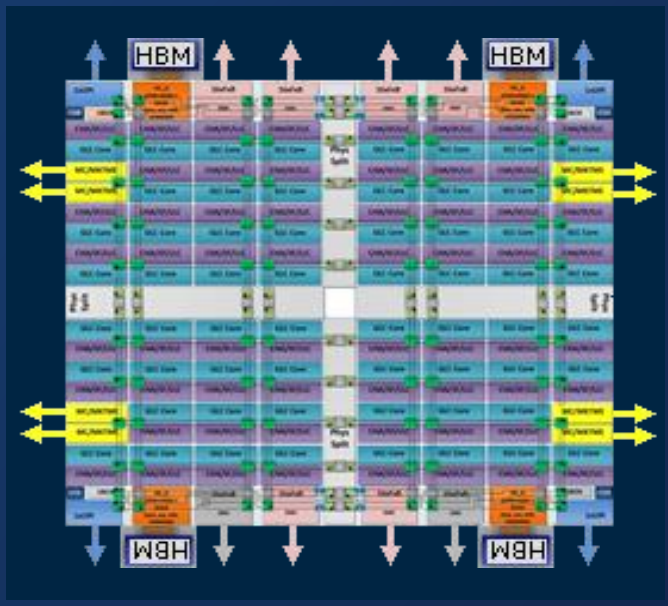
Socket Scalability,
Highest Core Count

4th Gen Intel Xeon processor - MCC
Monolithic Architecture



Mainstream Market, Higher Frequencies,
Lower Latency

Intel Xeon CPU Max Series
4 Tile Architecture



HPC and Memory Bandwidth
Bound Applications

Die Package Details

Features	4th Gen Intel® Xeon® Scalable Processors		Intel® Xeon® CPU Max Series
	Extreme Core Count (XCC)	Medium Core Count (MCC)	High Bandwidth Memory (HBM)
Die Construction	4 tiles connected using MDF over Intel Embedded Multi-die Interconnect Bridge (EMIB)	1 monolithic chip	4 tiles connected using MDF over Intel Embedded Multi-die Interconnect Bridge (EMIB)
Core Count	Up to 60 active cores	Up to 32 active cores	Up to 56 active cores
TDP Range	225 to 350W	125 to 350W	350W
Memory	DDR5 @ 4800 (1 DPC), 4400 (2DPC), 16 Gb DRAM, 8 Channels Intel® Optane™ PMem 300 (Crow Pass) @4400 MT/s		DDR5 @ 4800 (1 DPC), 4400 (2DPC), 8Channels 64 GB HBM2e memory with up to 1.14 GB/core
Intel UPI	UPI 2.0 @ 16 GT/s, up to 4 Ultra Path Interconnects	UPI 2.0 @ 16 GT/s, up to 3 Ultra Path Interconnects	UPI 2.0 @ 16 GT/s, up to 4 Ultra Path Interconnects
Scalability	1 Socket, 2 Socket, 4 Socket, 8 Socket	1 Socket, 2 Socket, 4 Socket	1 Socket, 2 Socket
PCIe/Compute Express Link	PCIe 5.0 (80 lanes), Up to 4 devices supported via Compute Express Link (CXL) 1.1		
Security	Intel® SGX Minimum Enclave Page Cache (EPC) size 256 MB		Intel® SGX (Flat mode only)
Integrated IP Accelerators	Intel® QAT, DLB, IAA, DSA (up to 4 devices each)	Intel® QAT, DLB (up to 2 devices each) Intel ® DSA, IAA (1 device each)	Intel® DSA (4 devices)

Driving Platform Innovation

Broad Ecosystem Readiness To Deploy Today

DDR5

- 8ch DDR5 (per CPU): up to 4800 MT/s
- 9x4 RDIMM support
- 3DS RDIMM support
- New RAS features
  - Enhanced ECC
  - Error Check and Scrub
- Both 2DPC and 1DPC today

Up to **1.5x**
higher memory bandwidth vs. DDR4

PCIe 5.0

- 80 lanes
- Improved DDIO and QoS capabilities
- X2 bifurcation @ Gen 4

Up to **2x**
Increased I/O Bandwidth vs. PCIe 4.0

CXL 1.1

- Next Gen I/O
- Up to 4 CXL devices supported per CPU
- Type 1 device: CXL.io and CXL.cache (e.g., SmartNIC)
- Type 2 device: CXL.io, CXL.cache and CXL.mem (e.g., GPU, ASIC, FPGA)

Type 1
and Type 2

UPI 2.0

- Up to 4 UPI links @ 16 GT/s
- New 8S-4 UPI Performance Optimized Topology

Up to **1.9x**
Increased inter-socket bandwidth vs. prior gen

Most Significant 4S/8S Upgrade in 5 Years

Architected for Cloud and Enterprise Customers with Large Scale up Applications

Customers Value 4S+ Systems

Customer Usages

- In-Memory DB
- Business Intelligence
- OLTP
- ERP
- CRM
- Data Warehousing/ Data Mining
- Data Visualization
- DC Consolidation

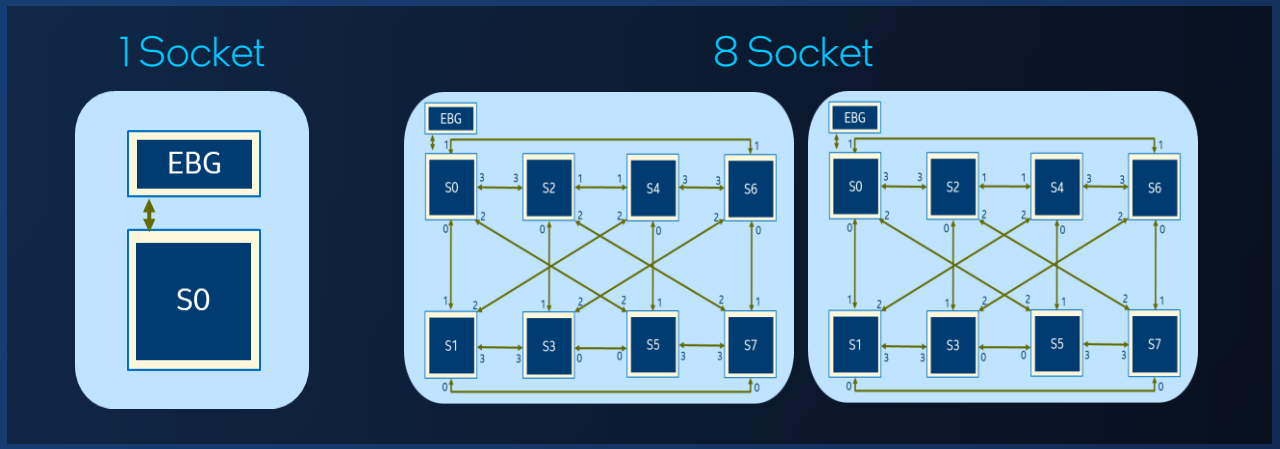
Customer Benefits

- Advanced RAS for Mission Critical apps
- High VM density on a single server node
- Reduces Total Cost of Ownership through maximum consolidation per server

Customer Ecosystem

- Over 30 4s/8s server designs and cloud instances expected in market starting in Q2'23

1S to 8S Glueless Topologies



Highest x86 density of compute / memory in one server

Scale Compute Performance

Up to **480** cores per server (8S)

Expand Memory Capacity

Up to **32 TB** memory per server (8S)

Maximize Multi-Socket BW

8S-4L new perf opt. topology (UPI 2.0)

Maximize the Effectiveness of Every Core

New Integrated IP Acceleration Engines

Increase performance, power and cost efficiency by enabling offload of common mode tasks via seamlessly integrated acceleration engines

Native Dispatch, Signaling, and Synchronization from User Space

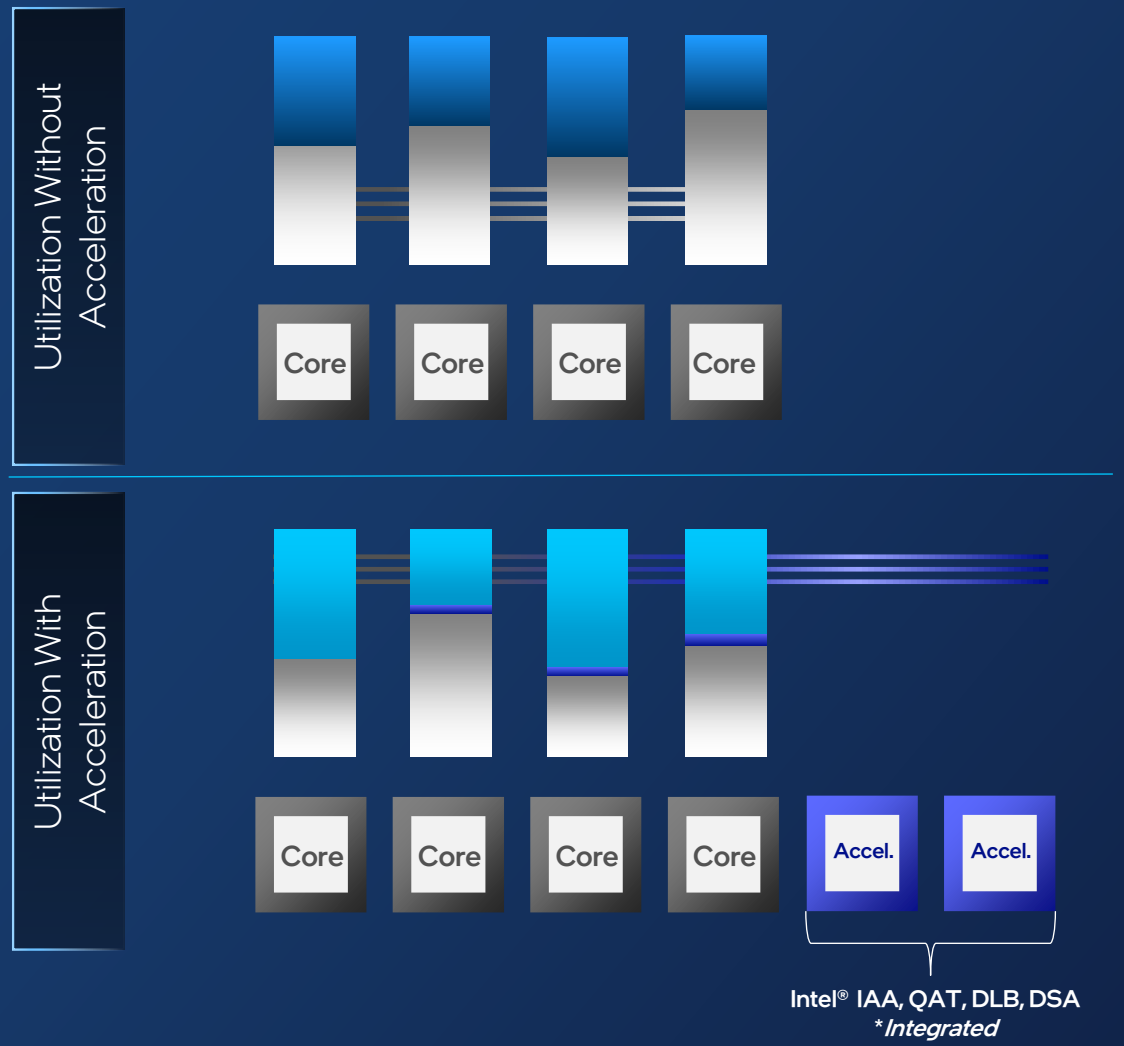
Accelerator Interfacing Architecture

Coherent, Shared Memory Space

Between Cores & Acceleration Engines

Concurrently Shareable

Processes, Containers, & VMs



intel XEON Accelerate with Xeon

Intel® Accelerator Engines

Most Built-in Accelerators of any CPU on the market providing customers with increased **performance, costs savings** and **sustainability** advantages for the biggest and fastest-growing workloads

Intel® AI Engines

- Intel® Advanced Matrix Extensions (Intel® AMX)
- Intel® Advanced Vector Extensions 512 (Intel® AVX-512)
- Intel® Deep Learning Boost (Intel® DL Boost)

Intel® Security Engines

- Intel® Control-Flow Enforcement Technology (Intel® CET)
- Intel® Crypto Acceleration
- Intel® Software Guard Extensions (Intel® SGX)
- Intel® Trust Domain Extensions (Intel® TDX)
- Intel® QuickAssist Technology (Intel® QAT)

Intel® HPC Engines

- Intel® Advanced Vector Extensions 512 (Intel® AVX-512)
- Intel® Advanced Matrix Extensions (Intel® AMX)
- Intel® Data Streaming Accelerator (Intel® DSA)
- Intel® QuickAssist Technology (Intel® QAT)

Intel® Network Engines

- Intel® QuickAssist Technology (Intel® QAT)
- Intel® Dynamic Load Balancer (Intel® DLB)
- Intel® Data Streaming Accelerator (Intel® DSA)
- Intel® Advanced Vector Extensions (Intel® AVX) for vRAN
- Intel® Speed Select Technology (Intel® SST)

Intel® Analytics Engines

- Intel® In-memory Analytics Accelerator (Intel® IAA)
- Intel® Data Streaming Accelerator (Intel® DSA)
- Intel® Advanced Vector Extensions 512 (Intel® AVX-512)
- Intel® QuickAssist Technology (Intel® QAT)

Intel® Storage Engines

- Intel® Data Streaming Accelerator (Intel® DSA)
- Intel® QuickAssist Technology (Intel® QAT)
- Intel® In-memory Analytics Accelerator (Intel® IAA)
- Intel® Data Direct I/O (Intel® DDIO)
- Intel® Advanced Vector Extensions 512 (Intel® AVX-512)
- Intel® Crypto Acceleration

intel XEON Accelerate with Xeon

Supported in 4th Gen Intel® Xeon® Scalable processors and Intel® Xeon® CPU Max Series processors

Intel Offers the Most Built-in Accelerators

	4th Gen Intel® Xeon® Scalable Processors	4th Gen AMD EPYC (Genoa)
Intel® Advanced Vector Extensions 512 (Intel® AVX-512)	Supported since 2015	New in 2022 – AVX 512
Intel® Crypto Acceleration	Supported since 2021	New in 2022 – Crypto NI
VNNI, BF16 (Intel® Deep Learning Boost)	Supported since 2019	New in 2022 – VNNI, BF16
Intel® Advanced Vector Extensions (Intel® AVX) for vRAN	✓	
Intel® Advanced Matrix Extensions (Intel® AMX)	✓	
Intel® Control-Flow Enforcement Technology (Intel® CET)	✓	Shadow Stack
Intel® Software Guard Extensions (Intel® SGX)	Supported since 2021	
Intel® Trust Domain Extensions (Intel® TDX)	✓ *Targeted CSP availability*	SEV-SNP
Intel® Speed Select Technology (Intel® SST)	Supported since 2019	
Intel® Data Direct I/O Technology (Intel® DDIO)	Supported since 2015	
Intel® Dynamic Load Balancer (Intel® DLB)	✓	
Intel® QuickAssist Technology (Intel® QAT) (integrated)	✓	
Intel® Data Streaming Accelerator (Intel® DSA)	✓	
Intel® In-Memory Analytics Accelerator (Intel® IAA)	✓	

A Higher Performance Server Architecture

Benefits of Intel® Accelerator Engines

Intel® Advanced Matrix Extensions (Intel® AMX)

Up to

8.6x

higher speech recognition inference performance with built-in AMX BF16 vs. FP32

Intel® Dynamic Load Balancer (Intel® DLB)

Up to

96%

lower latency at the same throughput for Istio-Envoy Ingress with Intel® DLB vs. software for Istio Ingress gateway

Intel® Data Streaming Accelerator (Intel® DSA)

Up to

1.7x

higher IOPs for SPDK-NVMe with built-in Intel® DSA vs. ISA-L software

Intel® In-Memory Analytics Accelerator (Intel® IAA)

Up to

2.1x

higher RocksDB performance with Intel® IAA vs Ztsd software

Intel® QuickAssist Technology (Intel® QAT)

Up to

84%

fewer cores to achieve same connections/s on NGINX with built-in QAT vs. out-of-box software

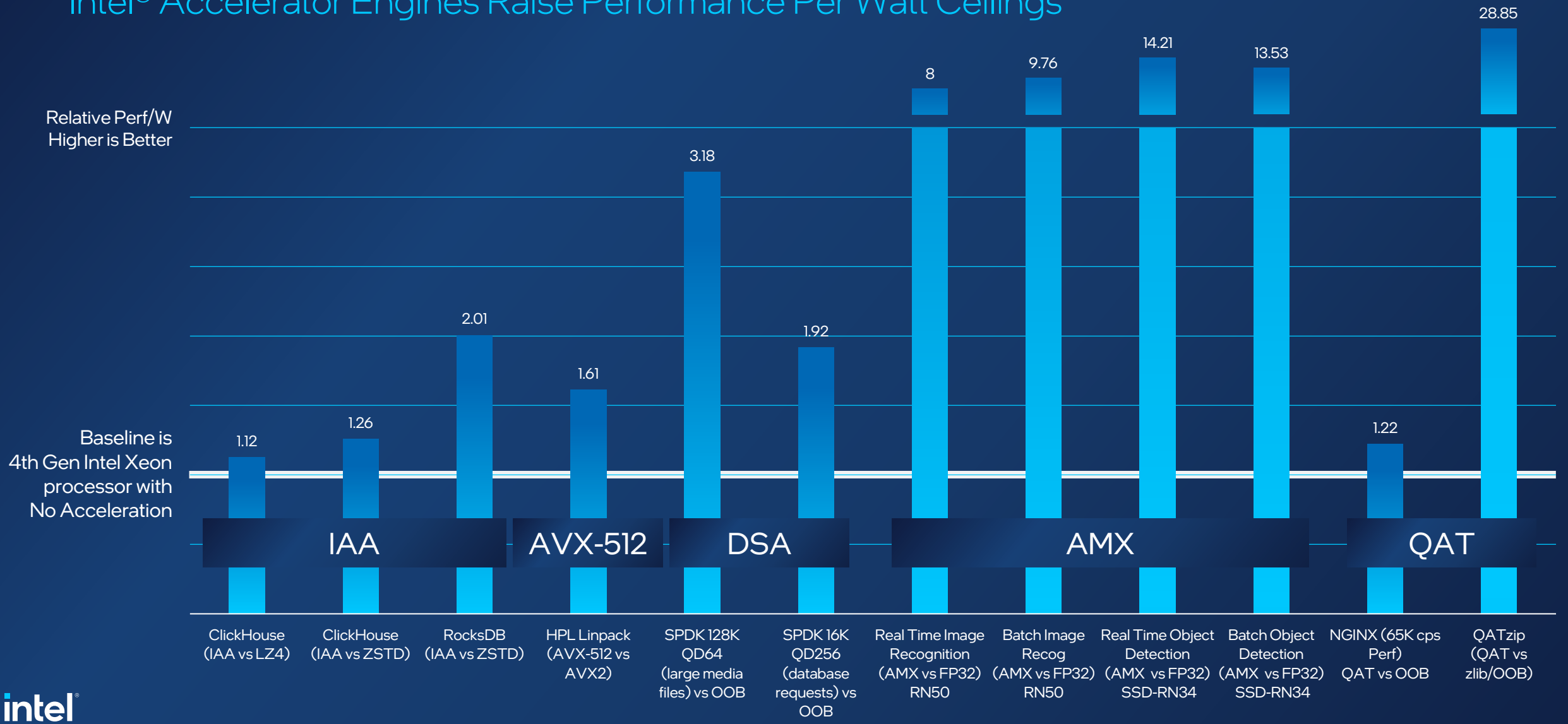
Accelerators Enable Step Function Performance Beyond Base Architecture

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

A More Energy Efficient Server Architecture

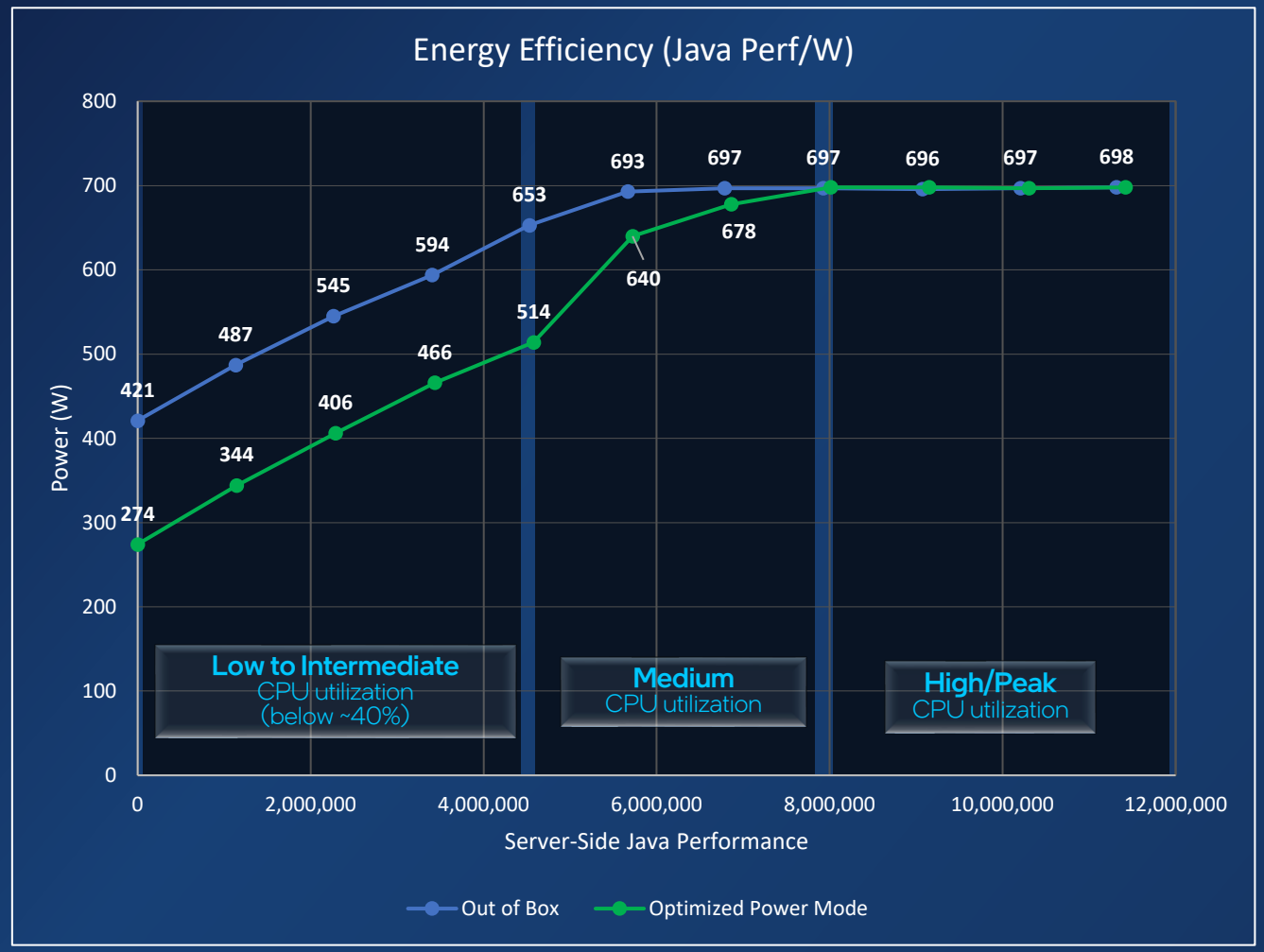
Intel® Accelerator Engines Raise Performance Per Watt Ceilings



intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

Built for More Sustainable Operations



Optimized Power Mode

Saves up to 20% CPU power at less than 5% performance impact for selected workloads

Saves power where customers tend to run (~30-40% utilization)

- up to 70W per socket at low utilization

Easy button BIOS option to enable (CSP, OxM, End User)

See backup for workloads and configurations. Results may vary.

Delivering Leading Performance for Customer Workloads

Artificial Intelligence

Networking 5G

Storage

HPC

Data Analytics

intel XEON Accelerate with Xeon

Architecting to Accelerate Customer Workloads

Leading Performance with most built-in accelerators

4th Gen Intel® Xeon® Scalable processors					Intel® Xeon® CPU Max Series
General Purpose Compute	Artificial Intelligence	Network 5G vRAN	Networking & Storage	Data Analytics	HPC
53%	Up to 10x	Up to 2x	Up to 2x	Up to 3x	Up to 3.7x
average performance gain*	higher inference and training performance*	capacity for vRAN workloads at same power envelope*	higher data compression with 95% fewer cores*	higher performance*	on memory-bound workloads**

Performance per Watt **2.9x** average improvement with built-in accelerators*

*4th Gen Intel® Xeon® Scalable Processor vs. 3rd Gen Intel Xeon Scalable processors.

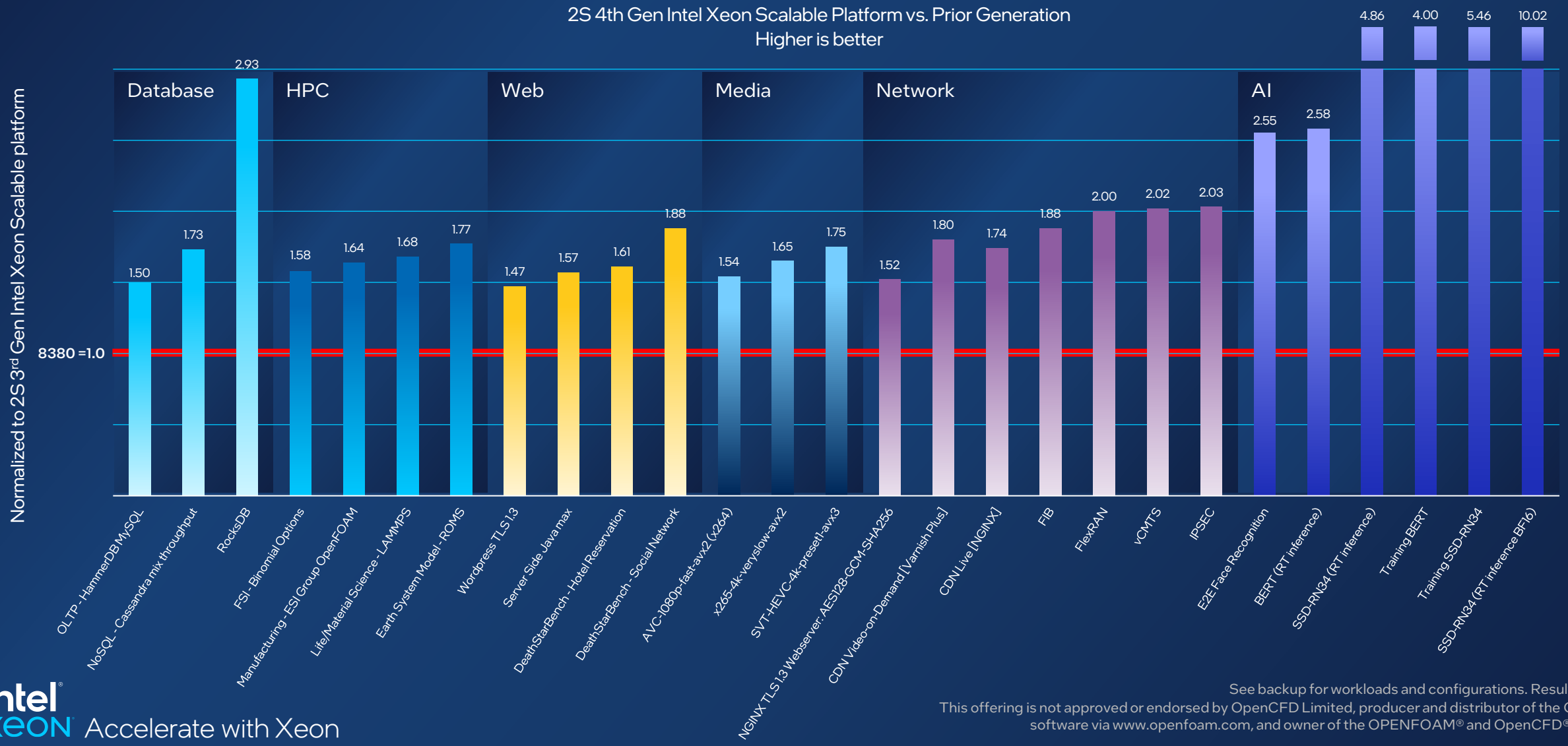
** Intel Xeon CPU Max Series vs. Intel Xeon 8380 processor.

See backup for workloads and configurations. Results may vary.

intel XEON Accelerate with Xeon

4th Gen Intel® Xeon® Scalable Processors

Significant Performance Across Broad Workloads



intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

Improving Performance When Response Time Matters

4th Gen Intel® Xeon® Scalable processors

Usages/Workload	Benchmark	Service Level Agreement (SLA) Requirement	Performance Gain*
Web Microservices	CloudXPRT	P95 latency < = 3sec	1.5x
Java Application Performance	Server-Side Java (Critical JOPs)	Geomean of 10ms, 25ms, 50 ms, 75ms and 100ms response times	1.6x
Cassandra Database	Cassandra Stress	P99 latency of <=20 ms	1.7x
Microservices – Social Networking	DeathStarBench	100 ms max SLA	1.8x
Image Classification (Real-time)	Resnet50 v1.5	15 ms max	3.0x
Object Detection (Real-time)	SSD-RN34	100 ms max	4.8x

*4th Gen Intel Xeon vs. previous gen.

See backup for workloads and configurations. Results may vary.

A More Cost-Efficient Server Architecture

Benefits of Workload Optimized Products

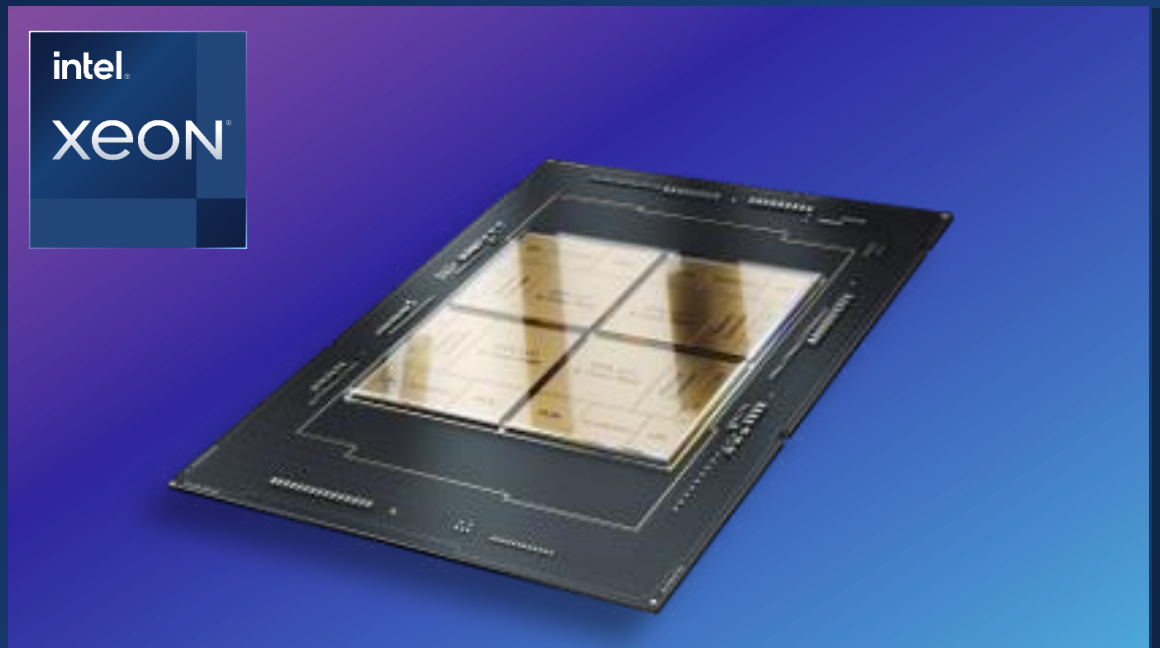
When considering new purchases for the data center, deploy fewer 4th Gen Intel® Xeon® processor-based servers or Intel® Xeon® CPU Max processor-based servers to meet the same performance requirement

Comparisons to deploying 50 servers with 3rd Gen Intel Xeon processor	Artificial Intelligence (Real time Inferencing, RSN50 w/ Intel® AMX)	Database (Rocks DB w/Intel® IAA)	HPC (OpenFOAM)
Number of Intel Xeon processor-based servers	17 servers with 4th Gen Intel® Xeon processors	18 servers with 4th Gen Intel® Xeon processors	16 servers with Intel® Xeon® CPU Max Series
Lower Fleet Power (kilowatts)	22.1 kW	15.4 kW	25.7 kW
Reduced CO2 emissions (kg)*	524,000 kg	366,000 kg	611,000 kg
TCO savings ($)*	$1.3M	$1.2M	$1.5M
	55% Lower TCO	**52% Lower TCO**	**66% Lower TCO**

* Estimated over 4 years
See backup for workloads and configurations. Results may vary.

This offering is not approved or endorsed by OpenCFD Limited, producer and distributor of the OpenFOAM software via www.openfoam.com, and owner of the OPENFOAM® and OpenCFD® trademark

4th Gen Intel® Xeon® Scalable Processors for HPC Accelerated Performance



- Built-in acceleration to boost application performance – Intel AVX 512, Intel DSA, Intel AMX
- New µarch, built on Intel 7, with up to 60 performance cores for added compute
- Increased memory bandwidth and higher speeds with DDR5
- Higher IO bandwidth with PCIe 5. and support for coherent interface with CXL 1.1

Intel Advanced Vector Extensions 512 (AVX-512)

Intel Advanced Matrix Extensions (AMX)

Intel Data Streaming Accelerator (DSA)

Up to 2.6x Performance On Real Workloads

2S 4th Gen Intel® Xeon® processor vs.
2S 3rd Gen Intel® Xeon® 8380 processor

See backup for workloads and configurations. Results may vary.
This offering is not approved or endorsed by OpenCFD Limited, producer and distributor of the OpenFOAM software via www.openfoam.com, and owner of the OPENFOAM® and OpenCFD® trademark
MLPerf™ HPC-AI v0.7 Training benchmark Performance. Result not verified by MLCommons Association. Unverified results have not been through an MLPerf™ review and may use measurement methodologies and/or workload implementations that are inconsistent with the MLPerf™ specification for verified results. The MLPerf™ name and logo are trademarks of MLCommons Association in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited. See www.mlcommons.org for more information

INTRODUCING

Intel® Xeon® CPU Max Series

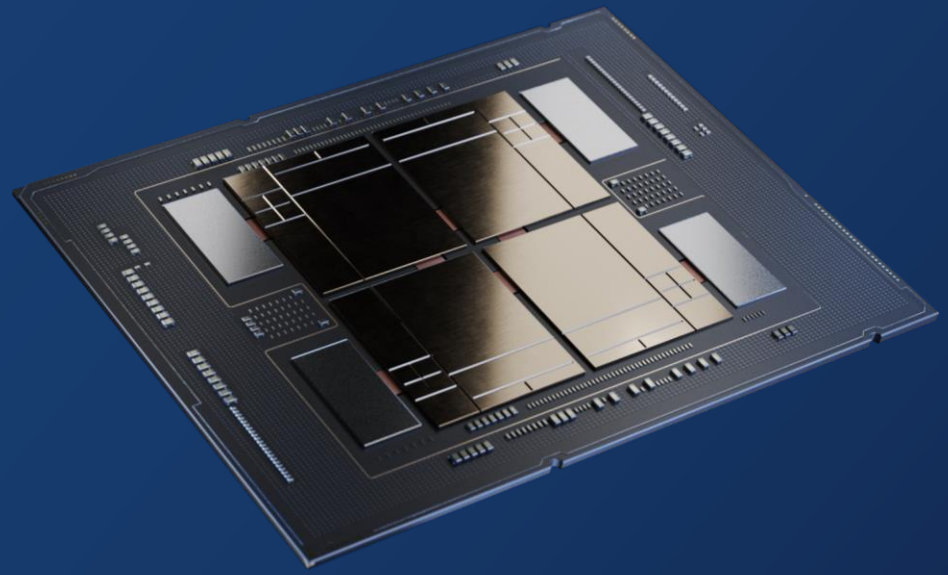
Designed for HPC, AI, Analytics and other memory bound Workloads

1st x86 CPU to integrate high bandwidth memory and accelerators onto the processor package

Improved TCO with reduced DDR dependency

intel XEON MAX SERIES

Only x86 CPU with High Bandwidth Memory (HBM)



HBM

64GB
HBM2e

up to
112.5MB
shared LLC

DDR5
8 channels per CPU @ 4800MTS (1DPC) / 16 DIMMs per socket

~1TB/s memory BW

>1GB/core HBM memory capacity

Memory modes

HBM Only
Workloads ≤ 64GB capacity

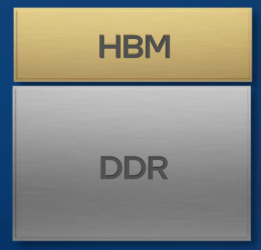
No code change
No DDR

System boots and operates with HBM only

HBM Flat Mode
Flat Mem Regions w/ HBM & DRAM
Workloads > 64GB capacity

Code change may be needed to optimize perf

Provides flexibility for applications that require large memory capacity



HBM Caching Mode
DRAM backed cache
Improved performance for workloads > 64GB capacity

No code change
HBM Caches DDR

Blend of both prior modes. Whole applications may fit in HBM cache
Blurs line between cache and memory

intel XEON Accelerate with Xeon

intel XEON MAX SERIES

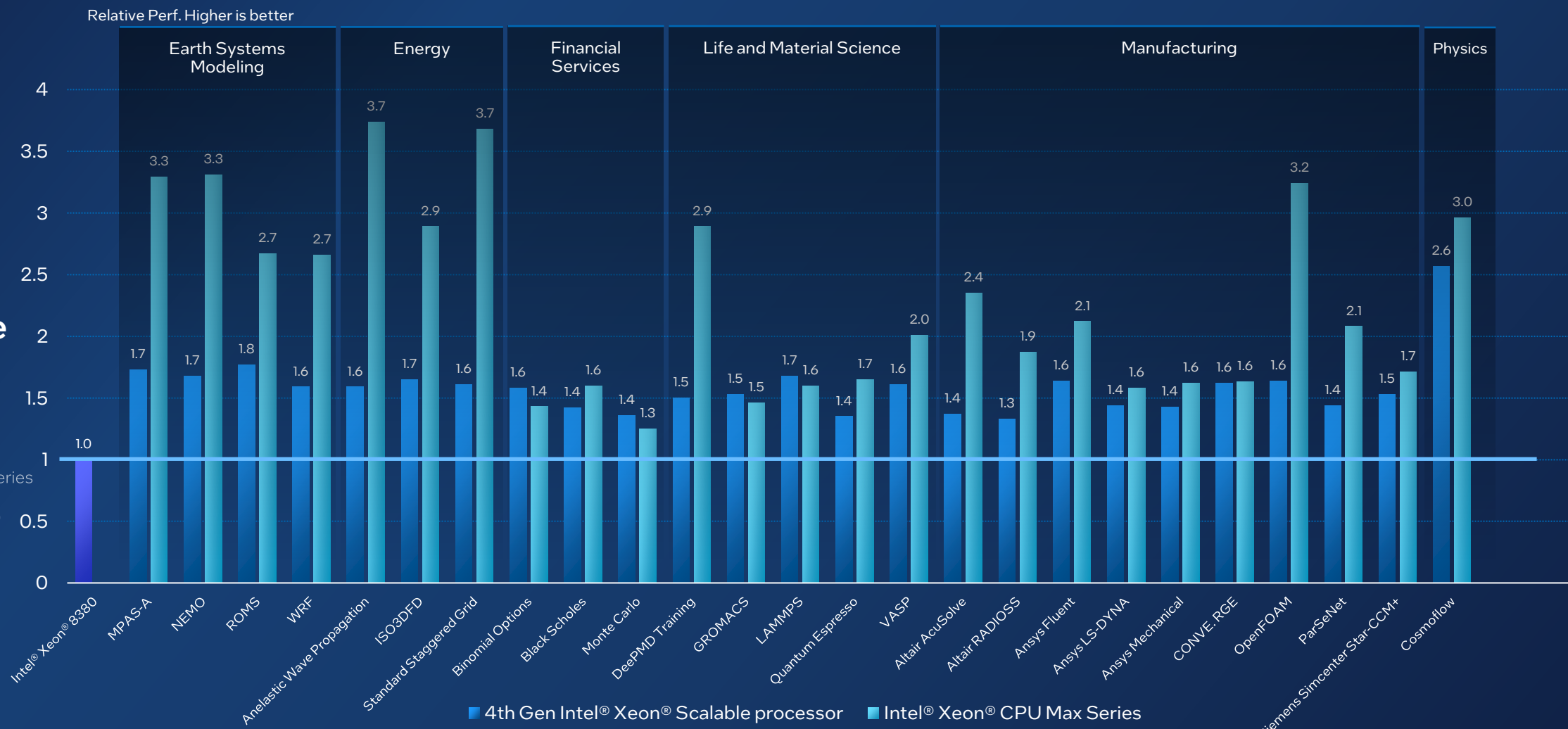
Up to

3.7x

performance

on real world workloads

2S Intel® Xeon® CPU Max Series
vs.
2S 3rd Gen Intel® Xeon® 8380 processor



See backup for workloads and configurations. Results may vary.

intel XEON Accelerate with Xeon

Bringing the Architecture to Life

4th Gen Intel® Xeon® Scalable Processors

Life Sciences

Get up to 53% faster results for life and material sciences for more effective research.

Digital Consumer Web Services

Run social network microservices up to 88% faster for better user experiences.

Financial Services

Meet tight timelines with up to 45% faster results for options pricing.

Retail

Offer personalized product recommendations up to 6.3x faster for smoother e-commerce.

See backup for workloads and configurations. Results may vary.

Architected to Accelerate Real World Workloads

Cloud

Live migration performance

Up to 89% performance increase with Intel® QAT vs. prior gen.[11]

inspur

Better user experience and cost efficiency

"We were pleased to observe a 20% increase in performance over the current generation C2 VMs from Google Cloud in testing with one of our key workloads."[12]

Google Cloud

Security

Helping Secure AI workloads

Intel® SGX performs up to 4.6x higher vs. prior gen.[13]

Fortanix

AI

Data purification

Up to 2.48x performance improvement with Intel® AMX vs. prior gen.[14]

Alibaba

Sentiment analysis

Up to 4x performance gain with Intel® AMX vs. prior gen.[15]

BRAINPOOL.AI FUJITSU

Graph neural network training

"Intel's [4th Gen Xeon processor] provides unprecedented levels of performance for critical graph intelligence tasks."

KATANA GRAPH

5G

Virtualized network

"It is not just a software, it is not interfaces, it is not only radio. It is how we can build all the pieces in our architecture."

Telefónica

HPC

High energy physics

Up to 4.3x performance improvement with Intel AMX® on Intel Xeon Max Series vs. prior gen.[16]

CERN openlab SURF

Climate simulation

Up to 8.57x performance improvement on Intel Xeon Max Series vs. Intel E5V4.[17]

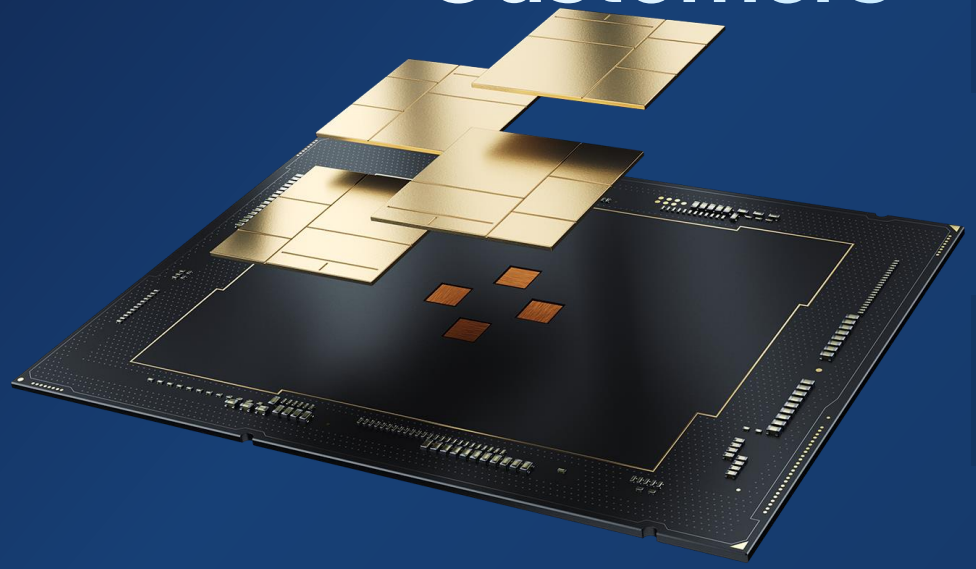
Los Alamos NATIONAL LABORATORY Hewlett Packard Enterprise

Sustainable liquid cooling

"The reason we use the 4th Gen Intel® Xeon® processor as the building block for immersion born systems is really because of its unrivaled power and efficiency."

Hypertec

An Architecture Influenced by Customers



Workload-first approach to innovation, design, and delivery

- Unique Die, SKU, features for Unique Market Needs
- DDR5, PCIe5 and CXL here today
- Cores + Accelerators deliver better value

Most built-in accelerators of any CPU on the market

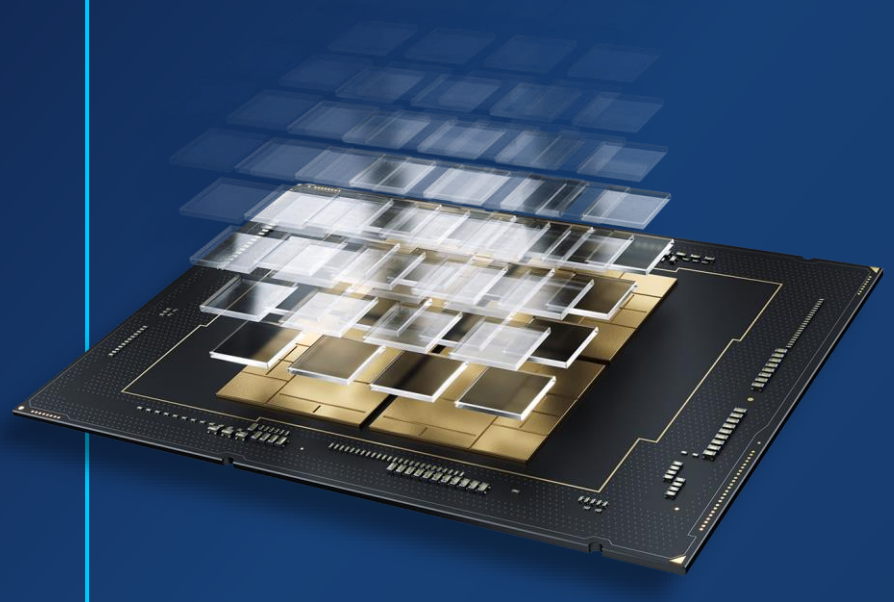
- Increased performance, power and cost efficiency
- Accelerating AI, Analytics, Networking, Storage, HPC
- 2.9x higher avg performance gains

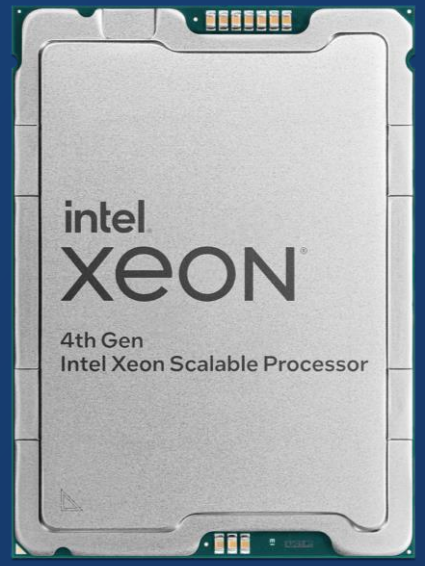
Leading performance and efficiency for our customers

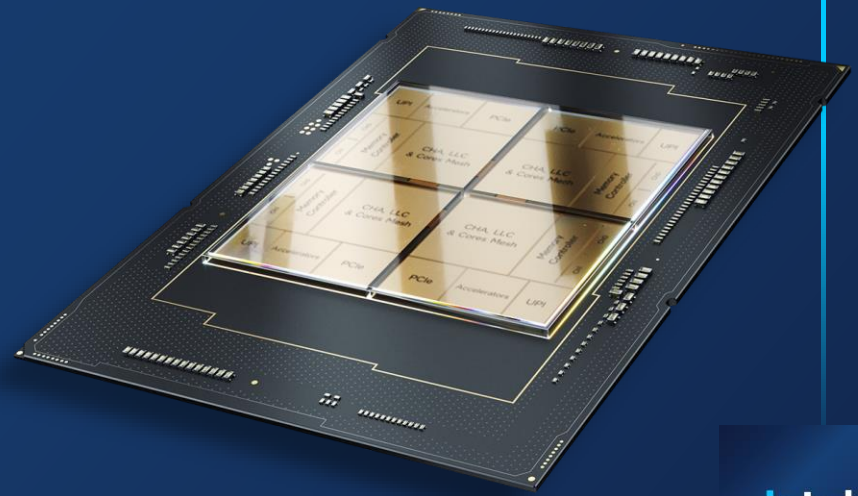
- Only x86 CPU to offer 4S and 8S scalability & HBM
- Up to 10X higher AI inference and training
- Lower TCO and power consumption

See backup for workloads and configurations. Results may vary.

Thank you!







intel XEON


intel

Intel's Data Center Evolution

Intel® Xeon® Scalable

Most Built-In Accelerators in the Market

Intel® Max CPU + GPU

Breakthrough Memory Bandwidth and Performance

Unrivaled Software Ecosystem

90% of developers are using software developed or optimized by Intel*

* Global Development Survey conducted by Evans Data Corp., 2021.

Intel Accelerator Engines by Processor Generation

Intel® Xeon® architecture has included purpose-built workload acceleration across Xeon generations

	Intel® Xeon® Scalable processors (Sky Lake)	2nd Gen Intel® Xeon® Scalable processors (Cascade Lake)	3rd Gen Intel® Xeon® Scalable processors (Ice Lake)	4th Gen Intel® Xeon® Scalable processors (Sapphire Rapids)
Intel® Advanced Vector Extensions 512 (Intel® AVX-512)	X	X	X	X
Intel® Crypto Acceleration			X	X
VNNI, BF16 (Intel® Deep Learning Boost)		X	X	X
Intel® Advanced Vector Extensions (Intel® AVX) for vRAN				X
Intel® Advanced Matrix Extensions (Intel® AMX)				X
Intel® Control-Flow Enforcement Technology (Intel® CET)				X
Intel® Software Guard Extensions (Intel® SGX)			X	X
Intel® Trust Domain Extensions (Intel® TDX)				Limited
Intel® Speed Select Technology (Intel® SST)		X	X	X
Intel® Data Direct I/O Technology (Intel® DDIO)	X	X	X	X
Intel® Dynamic Load Balancer (Intel® DLB)				X
Intel® QuickAssist Technology (Intel® QAT) (integrated)				X
Intel® Data Streaming Accelerator (Intel® DSA)				X
Intel® In-Memory Analytics Accelerator (Intel® IAA)				X

intel XEON Accelerate with Xeon

Codenames are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks

Intel® Quick Assist Technology

Acceleration Engine

Function
- Accelerated cryptography and data de/compression

Business Value
- Accelerated compression/decompression offloading leads to greater CPU efficiency
- More encrypted connections and web secure connections between devices with less overhead

Software Support
- Intel® QAT Engine for acceleration of cryptographic operations

Use Cases
- Distributed storage systems, file systems, RocksDB, Data lakes, Apache Spark, Hadoop, NGINX, IPSec

Performance gains vs not using these accelerators

Network Secure Gateway

Up to

84%

fewer cores to achieve same connections/s on NGINX with built-in QAT vs. out-of-the-box software

Performance gains vs prior generation products

Enterprise Storage and Data Analytics

Up to

95%

fewer cores and

2x

higher level 1 compression throughput leveraging integrated QAT vs. prior generation

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

Intel® Data Streaming Accelerator

Acceleration Engine

Function
- Optimizing streaming data movement and transformation operations

Business Value
- Accelerated data protection for NVMe/TCP improving efficiency for data storage applications via CPU offload

Software Support
- Intel® Data Mover Library

Use Cases
- Virtualization, fast replication across non-transparent bridge, ERP, In-Memory Databases

Performance gains vs not using these accelerators

Data Integrity (Throughput)

Up to

1.7x

higher IOPs for large packet sequential reads with built-in Intel® DSA vs. ISA-L software

Performance gains vs prior generation products

Data Integrity (Throughput & Latency)

Up to

1.6x

higher IOPs and

37%

latency reduction for large packet sequential reads with built-in Intel® DSA vs. prior generation

See backup for workloads and configurations. Results may vary.

Intel® Dynamic Load Balancer

Acceleration Engine

Function

- Dynamic redistribution of data load across cores when static NIC distribution causes a load-imbalance

Business Value

- Improves system performance related to handling network data on multi-core Intel® Xeon® Scalable processors
- Improved performance for distributed processing, dynamic load balancing and dynamic network processing reordering

Software Support

- Intel® Data Mover Library

Use Cases

- IPSec security gateway, VPP router, UPF, vSwitch, Streaming data processing, Elephant flow handling

Performance gains vs not using these accelerators

Microservices

Up to

96%

lower latency at the same throughput with built-in Intel® DLB vs. software for Istio ingress gateway

Performance gains vs prior generation products

Microservices

Up to

89%

lower latency and

57%

lower CPU utilization at same core count with built-in Intel® DLB vs. prior generation

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

Intel® Advanced Matrix Extensions

Acceleration Engine

Function

- Provides extensive hardware and software optimizations to enhance AI acceleration

Business Value

- Significant performance increases for AI/Deep Learning inference and training workloads
- Delivers common applications faster through hardware acceleration

Software Support

- Market relevant frameworks, toolkits and libraries (PyTorch, TensorFlow), Intel® oneAPI Deep Neural Network Library (oneDNN)

Use Cases

- Image recognition, recommendation systems, machine/language translation, NLP, media processing, and delivery

Performance gains vs prior generation products

Speech Recognition Inference

Up to

8.6x

higher speech recognition inference performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)

Performance gains vs prior generation products

PyTorch Training and Inference

Up to

10x

higher PyTorch for both real-time inference and training performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)

See backup for workloads and configurations. Results may vary.

Intel® In-Memory Advanced Analytics Accelerator

Acceleration Engine

Function
- Integrated accelerator IP accelerating analytics primitives, CRC calculations, compression, and decompression

Business Value
- Increases query throughput for in-memory databases and analytics workloads
- Decreases memory and bandwidth footprint for analytics workloads, freeing up space on CPU

Software Support
- Intel® Query Processing Library, Intel® Data Mover Library

Use Cases
- Commercial in-memory databases, open-source in-memory databases (RocksDB, Redis, Cassandra, MySQL, MongoDB), columnar formats for big data analytics

Performance gains vs not using these accelerators

Embedded Databases

Up to

2.1x

Higher RocksDB performance with built-in Intel® IAA vs. Zstd software

Performance gains vs prior generation products

Embedded Databases

Up to

3x

higher RocksDB performance with

66%

latency reduction with built-in Intel® IAA vs. prior generation

intel XEON Accelerate with Xeon

See backup for workloads and configurations. Results may vary.

Notices and Disclaimers

- Performance varies by use, configuration and other factors. Learn more on the Performance Index site.
- 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.
- Your costs and results may vary.
- Intel technologies may require enabled hardware, software or service activation.
- Intel contributes to the development of benchmarks by participating in, sponsoring, and/or contributing technical support to various benchmarking groups, including the BenchmarkXPRT Development Community administered by Principled Technologies.
- Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.
- Statements in this document that refer to future plans or expectations are forward-looking statements. These statements are based on current expectations and involve many risks and uncertainties that could cause actual results to differ materially from those expressed or implied in such statements. For more information on the factors that could cause actual results to differ materially, see our most recent earnings release and SEC filings at www.intc.com.
- Availability of accelerators varies depending on SKU. Visit the Intel Product Specifications page for additional product details.
-

Resources and Configurations

Geomean of HP Linpack, Stream Triad, SPECrate2017_fp_base est, SPECrate2017_int_base est. See [G2, G4, G6] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable.

Up to 10x higher PyTorch real-time inference performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)
PyTorch geomean of ResNet50, Bert-Large, MaskRCNN, SSD-ResNet34, RNN-T, Resnext101.

Up to 10x higher PyTorch training performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)
PyTorch geomean of ResNet50, Bert-Large, DLRM, MaskRCNN, SSD-ResNet34, RNN-T.

Estimated as of 8/30/2022 based on 4th generation Intel® Xeon® Scalable processor architecture improvements vs 3rd generation Intel® Xeon® Scalable processor at similar core count, socket power and frequency on a test scenario using FlexRAN™ software. Results may vary.

Up to 95% fewer cores and 2x higher level 1 compression throughput with 4th Gen Intel Xeon Platinum 8490H using integrated Intel QAT vs. prior generation.
8490H: 1-node, pre-production platform with 2x 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), with Total 1024GB (16x64 GB) DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, QAT v20.I.0.9.1 , QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel September 2022.

8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01,QAT v1.7.I.4.16, QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel October 2022.

Up to 3x higher RocksDB performance with 4th Gen Intel Xeon Platinum 8490H using integrated Intel IAA vs. prior generation.

8490H: 1-node, pre-production Intel platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0xf000380, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel September 2022.

8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel October 2022.

Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, YASK v3.05.07

Intel® Xeon® CPU Max Series: Test by Intel as of ww36'22. 1-node, 2x Intel® Xeon® CPU Max SeriesHT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, YASK v3.05.07.

Resources and Configurations

Lower performance per watt 2.9x

Geomean of following workloads: RocksDB (IAA vs ZTD), ClickHouse (IAA vs ZTD), SPDK large media and database request proxies (DSA vs out of box), Image Classification ResNet-50 (AMX vs VNNI), Object Detection SSD-ResNet-34 (AMX vs VNNI), QATzip (QAT vs zlib)

RocksDB

New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), on pre-production Intel platform and software, HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.2.1,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022.

ClickHouse

New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), Number of IAA device utilized=8(2 sockets active), on pre-production Intel platform and software, HT On, Turbo On, SNC off, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21, accel-config-v3.4.6.4, gcc 11.2, Clickhouse 21.12, Star Schema Benchmark, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, gcc 11.2, Clickhouse 21.12, Star Schema Benchmark, tested by Intel November 2022.

SPDK

New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR, DDR4 memory total 1024GB (16x64 GB), HT On, Turbo On, SNC Off, microcode 0xd000375, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x 1.92TB Intel SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel November 2022.

ResNet-50

New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), on pre-production Supermicro SYS-221H-TNR with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 FP32 8 cores/instance (max. 15ms SLA), BS1 INT8 2 cores/instance (max. 15ms SLA), BS1 AMX 1 core/instance (max. 15ms SLA), BS16 FP32 5 cores/instance, BS16 INT8 5 cores/instance, BS16 AMX 5 cores/instance, using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 FP32 8 cores/instance (max. 15ms SLA), BS1 INT8 2 cores/instance (max. 15ms SLA), BS16 FP32 5 cores/instance, BS16 INT8 5 cores/instance, using physical cores, tested by Intel November 2022.

SSD-ResNet-34

New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), Intel platform with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=SSD-ResNet34, best scores achieved: BS1 FP32 60 cores/instance (max. 100ms SLA), BS1 INT8 4 cores/instance (max. 100ms SLA), BS1 AMX 4 core/instance (max. 100ms SLA), BS8 FP32 8 cores/instance, BS2 INT8 1 cores/instance, BS2 AMX 1 cores/instance, using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model= SSD-ResNet34, best scores achieved: BS1 FP32 40 cores/instance (max. 100ms SLA), BS1 INT8 10 cores/instance (max. 100ms SLA), BS16 FP32 4 cores/instance, BS16 INT8 1 cores/instance, using physical cores, tested by Intel November 2022.

QAT.zip

New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), on pre-production Intel platform and software with DDR5 memory Total 1024GB (16x64 GB), microcode 0x2b0000a1, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, QAT v20.l.0.9.1, QATzip v1.0.9, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, QAT v20.l.0.9.1, QATzip v1.0.9, tested by Intel November 2022.

Lower Power Bills up to 70%

1-node, Intel Reference Validation Platform, 2x Intel® Xeon 8480+ (56C, 2GHz, 350W TDP), HT On, Turbo ON, Total Memory: 1 TB (16 slots/ 64GB/ 4800 MHz), 1x P4510 3.84TB NVMe PCIe Gen4 drive, BIOS: 0091.D05, (ucode:0x2b0000c0), CentOS Stream 8, 5.15.0-spr.bkc.pc.10.4.11.x86_64, Java Perf/Watt w/ openjdk-11+28_linux-x64_bin, 112 instances, 1550MB Initial/Max heap size, Tested by Intel as of Oct 2022.

Lower TCO/More Sustainable, 55%

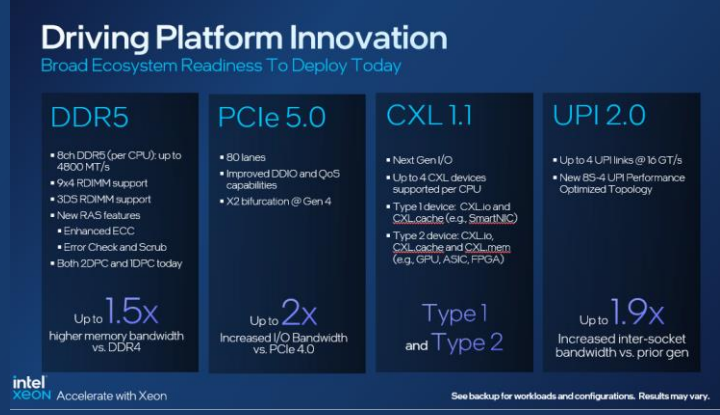
ResNet50 Image Classification

New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable 8490H processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), on pre-production SuperMicro SYS-221H-TNR with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 AMX 1 core/instance (max. 15ms SLA), using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable 8380 Processor (40 cores) on SuperMicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 INT8 2 cores/instance (max. 15ms SLA), using physical cores, tested by Intel November 2022.

For a 50 server fleet of 3rd Gen Xeon 8380 (RN50 w/DLBoost), estimated as of November 2022:
CapEx costs: $1.64M
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $739.9K
Energy use in kWh (4 year, per server): 44627, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

For a 17 server fleet of 4th Gen Xeon 8490H (RN50 w/AMX), estimated as of November 2022:
CapEx costs: $799.4K
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $275.3K
Energy use in kWh (4 year, per server): 58581, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

Resources and Configurations

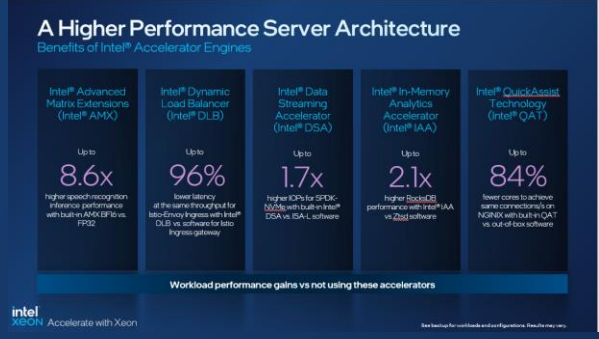


Driving Platform Innovation

Broad Ecosystem Readiness To Deploy Today

- **Up to 1.5x higher memory bandwidth vs. DDR44** - 8490H:1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8490H on Archer City with GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000070, HT off, Turbo on, Ubuntu 22.04 LTS, 5.15.0-47-generic, 1x INTEL SSDPF2KX076TZ, STREAM Triad version 5.10, OneAPI 2022.1, test by Intel on 9/9/2022. 8380: 1-node, 2x Intel(R) Xeon(R) Platinum 8380 on Whitley with GB (16 slots/ 32GB/ DDR4 3200) total memory, ucode 0xd000363, HT off, Turbo on, Ubuntu 22.04 LTS, 5.15.0- 39-generic, 1x INTEL SSDPF21Q016TB, STREAM Triad version 5.10, OneAPI 2022.1, test by Intel on 8/13/2022.
- **1Up to 2x Increased I/O Bandwidth vs. PCIe 4.0** - Results have been estimated or simulated. Comparing x16 bandwidth on 4th Gen Intel Xeon Scalable processor with PCIe Gen5 running at up to 128 GB/s vs 3rd Gen Intel Xeon Scalable processor with PCIe Gen4 running at up to 64 GB/s
- **Up to 2x Increased I/O Bandwidth vs. PCIe 4.0** - SPR (UPI 2.0, 4 UPI links @16 GT/s) compared to prior gen is ICX (UPI 1,0, 3 UPI links @11.2 GT/s) - raw platform enhancement gen-gen.

Resources and Configurations



A Higher Performance Server Architecture

Benefits of Intel® Accelerator Engines

- **Up to 8.6x higher speech recognition inference performance with built-in AMX BF16 vs. FP3220** - 8480+: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), Resnext101 32x16d, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, amx bf16=1,64, amx int8=1,116, ImageNet, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.; 8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, Resnext101 32x16d, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, int8=1,116, ImageNet, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

- **Up to 96% lower latency at the same throughput for Istio-Envoy Ingress with Intel® DLB vs. software for Istio Ingress gateway** -8480+: 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ with Intel DLB on Intel ArcherCity with GB (32 slots/ 32GB/ DDR5 4800) total memory, ucode 0x2b0000a1, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-40-generic, 1x 54.9G INTEL SSDPEK1A058GA, 4x Ethernet Controller E810-C for QSFP, 2x Ethernet Controller XXV710 for 25GbE SFP28, 1x Ethernet Controller I225-LM, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 10/27/2022.; 8360Y: 1-node, 2x Intel(R) Xeon(R) Platinum 8360Y on Intel M50CYP2SBSTD with GB (32 slots/ 16GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-50-generic, 2x 1.8T INTEL SSDPE2KX020T8, 1x Ethernet Controller E810-C for QSFP, 2x BCM57416 NetXtreme-E Dual-Media 10G RDMA Ethernet Controller, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 11/3/2022.

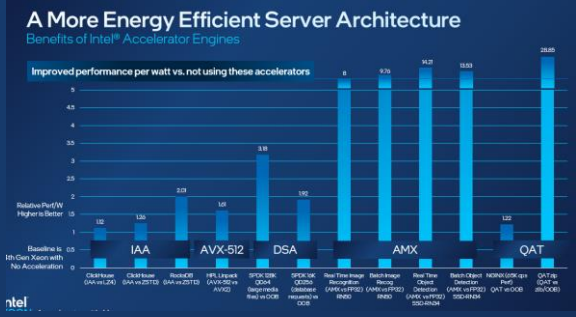
- **Up to 1.7x higher IOPs for SPDK-NVMe with built-in Intel® DSA vs. ISA-L software -** 8490H: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0xf000380, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel September 2022.; 8380: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel October 2022.

- **Up to 2.1x higher RocksDB performance with Intel® IAA vs Ztsd software-** 8490H: 1-node, pre-production Intel platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0xf000380, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel September 2022; 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel October 2022

- **Up to 84% fewer cores to achieve same connections/s on NGINIX with built-in QAT vs. out-of-box software** - QAT Configuration HW/SW on 8490H: 1-node, pre-production platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel QuickAssist Accelerator (Intel QAT), Number of QAT device utilized=4(1 socket active), 1024GB (16x64 GB) total DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, QAT engine v0.6.14, QAT v20.l.0.9.1, NGINX 1.20.1, OpenSSL 1.1.1l, IPP crypto v2021_5, IPSec v1.1 , TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel September 2022.; QAT Configuration SW on 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, QAT engine v0.6.14, NGINX 1.20.1, OpenSSL 1.1.1l, IPP crypto v2021_5, IPSec v1.1 , TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel October 2022.
  - OOB Configurations:
  - 8490H: 1-node, pre-production platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores), with 1024GB (16x64 GB) total DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, NGINX 1.20.1, OpenSSL 1.1.1l, TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel September 2022.
  - 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, NGINX 1.20.1, OpenSSL 1.1.1l, TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel October 2022

intel
XEON Accelerate with Xeon

Resources and Configurations



A More Energy Efficient Server Architecture

Up to 1.12x and 1.26x higher performance/W using 4th Gen Xeon Scalable w/Intel Analytics Accelerator vs LZ4 and Zstd on ClickHouse
1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), Number of IAA device utilized=8(2 sockets active), on pre-production Intel platform and software, HT On, Turbo On, SNC off, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21, accel-config-v3.4.6.4, gcc 11.2, Clickhouse 21.12, Star Schema Benchmark, tested by Intel November 2022.

Up to 2.01x higher performance/W using 4th Gen Xeon Scalable w/Intel Analytics Accelerator vs Zstd on RocksDB
1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), on pre-production Intel platform and software, HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.2.1,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022.

Up to 1.61 higher performance/W using 4th Gen Xeon Scalable w/AVX-512 vs AVX2 on Linpack
1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core), on pre-production Supermicro SYS-221H-TNR and software with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC 4, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, One API BaseKit 2022.2.0.262, One API HPC 2022.2.0.191, Linpack ver 2.3, tested by Intel November 2022.

Up to 3.18x and 1.92x higher performance/W using 4th Gen Xeon Scalable w/Data Streaming Accelerator vs out-of-box OS software on SPDK NVMe TCP
1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel November 2022.

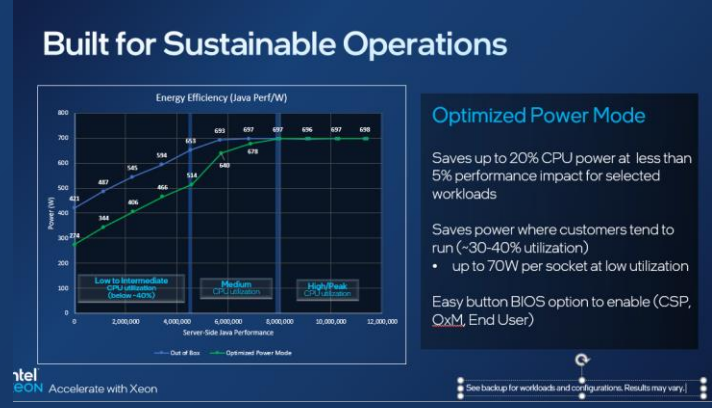
Up to 8x and 9.76x higher performance/W using 4th Gen Xeon Scalable w/Advanced Matrix Extensions using AMX vs VNNI instructions on ResNet50 Image Processing
1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), on pre-production Supermicro SYS-221H-TNR with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 FP32 8 cores/instance (max. 15ms SLA), BS1 INT8 2 cores/instance (max. 15ms SLA), BS1 AMX 1 core/instance (max. 15ms SLA), BS16 FP32 5 cores/instance, BS16 INT8 5 cores/instance, BS16 AMX 5 cores/instance, using physical cores, tested by Intel November 2022.

Up to 14.21x and 13.53x higher performance/W using 4th Gen Intel Xeon Scalable w/Advanced Matrix Extensions using AMX vs VNNI instructions on SSD-ResNet34 on Object Detection
1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), Intel platform with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=SSD-ResNet34, best scores achieved: BS1 FP32 60 cores/instance (max. 100ms SLA), BS1 INT8 4 cores/instance (max. 100ms SLA), BS1 AMX 4 core/instance (max. 100ms SLA), BS8 FP32 8 cores/instance, BS2 INT8 1 cores/instance, BS2 AMX 1 cores/instance, using physical cores, tested by Intel November 2022.

Up to 1.22x higher performance/W using 4th Gen Intel Xeon Scalable w/QuickAssist Accelerator vs out-of-box software on NGINX TLS Handshake.
QAT Accelerator: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel QuickAssist Accelerator (Intel QAT), Number of QAT device utilized=4(1 socket active), on pre-production Intel platform and software with DDR5 memory total 1024GB (16x64 GB), microcode 0x2b0000a1, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, QAT engine v0.6.14, QAT v20.l.0.9.1, NGINX 1.20.1, OpenSSL 1.1.1l, IPP crypto v2021_5, IPSec v1.1 , TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, 65K CPS target SLA, tested by Intel November 2022. Out of box configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel QuickAssist Accelerator (Intel QAT), Number of QAT device utilized=0, on pre-production Intel platform and software with DDR5 memory total 1024GB (16x64 GB), microcode 0x2b0000a1, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, NGINX 1.20.1, OpenSSL 1.1.1l, TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, 65K CPS target SLA, tested by Intel November 2022.

Up to 28.85x higher performance/W using 4th Gen Intel Xeon Scalable w/QuickAssist Accelerator vs out-of-box zlib on QATzip compression
1-node, 2x pre-production 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), on pre-production Intel platform and software with DDR5 memory Total 1024GB (16x64 GB), microcode 0x2b0000a1, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, QAT v20.l.0.9.1 , QATzip v1.0.9, tested by Intel November 2022.

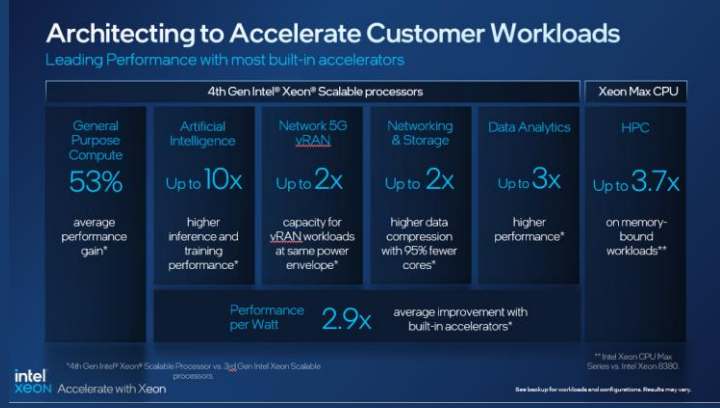
Resources and Configurations



Built for Sustainable Operations

1-node, Intel Reference Validation Platform, 2x Intel® Xeon 8480+ (56C, 2GHz, 350W TDP), HT On, Turbo ON, Total Memory: 1 TB (16 slots/ 64GB/ 4800 MHz), 1x P4510 3.84TB NVMe PCIe Gen4 drive, BIOS: 0091.D05, (ucode:0x2b0000c0), CentOS Stream 8, 5.15.0-spr.bkc.pc.10.4.11.x86_64, Java Perf/Watt w/ openjdk-11+28_linux-x64_bin, 112 instances, 1550MB Initial/Max heap size, Tested by Intel as of Oct 2022.

Resources and Configurations

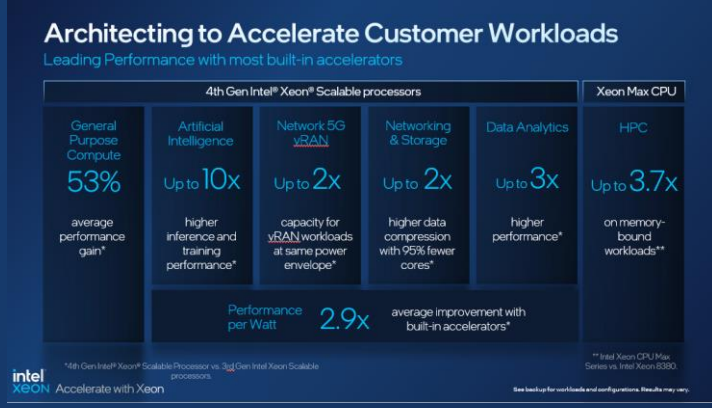


Architecting to Accelerate Customer Workloads (1 of 2)

Leading Performance with the most built – in accelerators

- **General purpose compute 53% average performance gain** - Geomean of HP Linpack, Stream Triad, SPECrate2017_fp_base est, SPECrate2017_int_base est. See [G2, G4, G6] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable.
- **Up to 10X higher inference and training performance** - 5.7x to 10x higher PyTorch real-time inference performance on 4th Gen Intel Xeon Scalable processor with built in Intel AMX (BF16) vs. prior generation (FP32);5.7-10x & 7x: PyTorch geomean of ResNet50, Bert-Large, MaskRCNN, SSD-ResNet34, RNN-T, Resnext101; 5.8x to 9.6x higher PyTorch batch inference performance on 4th Gen Intel Xeon Scalable processor with built in Intel AMX (BF16) vs. prior generation (FP32); 5.8-9.6x & 7x: PyTorch geomean of ResNet50, Bert-Large, MaskRCNN, SSD-ResNet34, RNN-T, Resnext101, DLRM
- **Up to 2x capacity for vRAN workloads at same power envelope** - Estimated as of 8/30/2022 based on 4th generation Intel® Xeon® Scalable processor architecture improvements vs 3rd generation Intel® Xeon® Scalable processor at similar core count, socket power and frequency on a test scenario using FlexRAN™ software. Results may vary.
- **Up to 2x higher data compression with 95% fewer cores** - 90H: 1-node, pre-production platform with 2x 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), with Total 1024GB (16x64 GB) DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, QAT v20.l.0.9.1 , QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel September 2022. ; 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01,QAT v1.7.l.4.16, QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel October 2022.
- **Up to 3x higher performance** - 8490H: 1-node, pre-production Intel platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0xf000380, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel September 2022; 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel October 2022
- **Up to 3.7x on memory-bound workloads** - Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Stream v5.10; Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® CPU Max Series, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Stream v5.10

Resources and Configurations

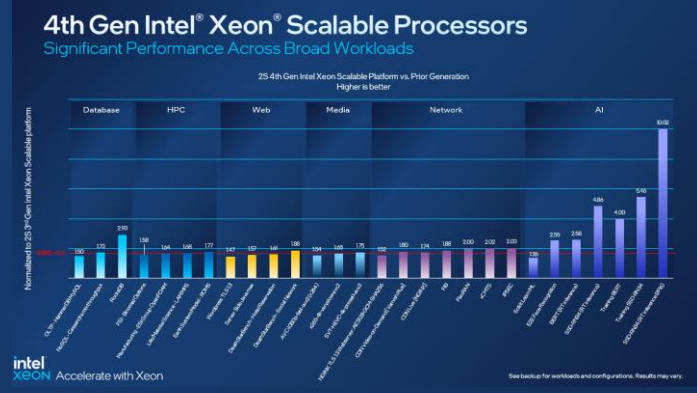


Architecting to Accelerate Customer Workloads (2 of 2)

Leading Performance per Watt with the most built – in accelerators

- Geomean of following workloads: RocksDB (IAA vs ZTD), ClickHouse (IAA vs ZTD), SPDK large media and database request proxies (DSA vs out of box), Image Classification ResNet-50 (AMX vs VNNI), Object Detection SSD-ResNet-34 (AMX vs VNNI), QATzip (QAT vs zlib)
- RocksDB
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), on pre-production Intel platform and software, HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.2.1,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022.
- ClickHouse
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), Number of IAA device utilized=8(2 sockets active), on pre-production Intel platform and software, HT On, Turbo On, SNC off, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21, accel-config-v3.4.6.4, gcc 11.2, Clickhouse 21.12, Star Schema Benchmark, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, gcc 11.2, Clickhouse 21.12, Star Schema Benchmark, tested by Intel November 2022.
- SPDK
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), HT On, Turbo On, SNC Off, microcode 0xd000375, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x 1.92TB Intel SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel November 2022.
- ResNet-50
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), on pre-production Supermicro SYS-221H-TNR with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 FP32 8 cores/instance (max. 15ms SLA), BS1 INT8 2 cores/instance (max. 15ms SLA), BS1 AMX 1 core/instance (max. 15ms SLA), BS16 FP32 5 cores/instance, BS16 INT8 5 cores/instance, BS16 AMX 5 cores/instance, using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 FP32 8 cores/instance (max. 15ms SLA), BS1 INT8 2 cores/instance (max. 15ms SLA), BS16 FP32 5 cores/instance, BS16 INT8 5 cores/instance, using physical cores, tested by Intel November 2022.
- SSD-ResNet-34
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), Intel platform with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000a1, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=SSD-ResNet34, best scores achieved: BS1 FP32 60 cores/instance (max. 100ms SLA), BS1 INT8 4 cores/instance (max. 100ms SLA), BS1 AMX 4 core/instance (max. 100ms SLA), BS8 FP32 8 cores/instance, BS2 INT8 1 cores/instance, BS2 AMX 1 cores/instance, using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model= SSD-ResNet34, best scores achieved: BS1 FP32 40 cores/instance (max. 100ms SLA), BS1 INT8 10 cores/instance (max. 100ms SLA), BS16 FP32 4 cores/instance, BS16 INT8 1 cores/instance, using physical cores, tested by Intel November 2022.
- QAT.zip
  - New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), on pre-production Intel platform and software with DDR5 memory Total 1024GB (16x64 GB), microcode 0x2b0000a1, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, QAT v20.l.0.9.1, QATzip v1.0.9, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Supermicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, QAT v20.l.0.9.1, QATzip v1.0.9, tested by Intel November 2022.

Resources and Configurations



Significant Performance Across Broad Workloads (1-6)

- 1.50x OLTP - HammerDB MySQL – **8490H:** 1-node, pre-production platform with 2(1 used)x Intel(R) Xeon(R) Platinum 8490H on ArcherCity with GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, RHEL 8.6 (Ootpa), 4.18.0-372.9.1.el8.x86_64, 1x 894.3G INTEL SSDSC2KG96, 5x 1.5T INTEL SSDPF21Q016TB, 1x Ethernet Controller I225-LM, HammerDB 4.4, MySQL 8.0.30, test by Intel on 10/12/2022. **8380:** 1-node, 2x(1 socket used) Intel(R) Xeon(R) Platinum 8380 on M50CYP2SBSTD with GB (16 slots/ 32GB/ DDR4 3200) total memory, ucode 0xd000363, HT on, Turbo on, RHEL 8.6 (Ootpa), 4.18.0-372.9.1.el8.x86_64, 1x 894.3G INTEL SSDSC2KG96, 1x 1.5T INTEL SSDPF21Q016TB, 4x 2.9T INTEL SSDPF21Q032TB, 2x Ethernet Controller X710 for 10GBASE-T, HammerDB 4.4, MySQL 8.0.30, test by Intel on 10/12/2022.

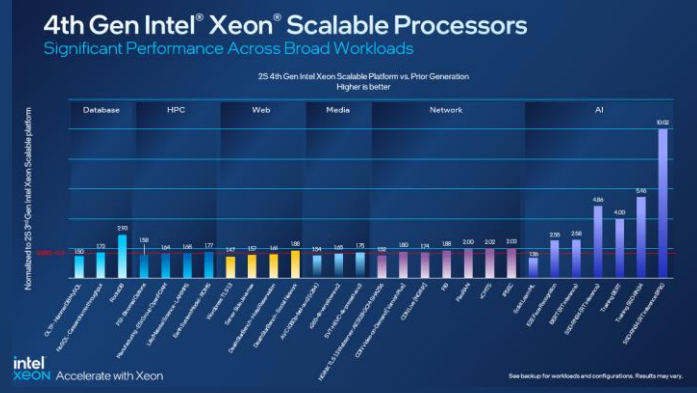
- 1.73x NoSQL - Cassandra mix throughput - **8490H:** Test by Intel as of 10/18/22. 1-node, 2x Intel(R) Xeon(R) Platinum 8490H, 60 cores, HT On, Turbo On, Total Memory 512GB (16x32GB 4800 MT/s [4800 MT/s]), BIOS EGSDCRB1.FEI.0088.D16.2209090804, microcode 0xababc0a0, 1x Ethernet Controller I225-LM, 4x Ethernet Controller X710/X557-AT 10GBASE-T, 1x 1.5T INTEL SSDSC2BB01, 8x 1.5T INTEL SSDPF21Q016TB, Ubuntu 22.04.1 LTS, 5.19.10-051910-generic, compiler gcc version 11.2.0, cassandra-stress version 4.0.6, jdk version 14 build 14+36-1461. **8380:** Test by Intel as of 10/17/22, 1-node, 2x Intel(R) Xeon(R) Platinum 8380 CPU @ 2.30GHz, 40 cores, HT On, Turbo On, Total Memory Installed 512GB (16x32GB DDR4 3200 MT/s [3200 MT/s]), Total Memory Used 256GB as 256GB blocked, BIOS SE5C620.86B.01.01.0005.2202160810, microcode 0xd000375, 2x Ethernet Controller X710 for 10GBASE-T, 1x 223.6G KINGSTON SA400M8240G, 4x 1.5T INTEL SSDPF21Q016TB, Ubuntu 22.04.1 LTS, 5.19.10-051910-generic, compiler gcc version 11.2.0, cassandra-stress version 4.0.6, jdk version 14 build 14+36-146.

- 2.93x RocksDB – **8490H:** 1-node, 2x pre-production 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), on pre-production Intel platform and software, HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.2.1,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022. **8380:** 1-node, 2x 3rd Gen Intel Xeon Scalable Processors (40 cores) on Supermicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022.

- 1.58x FSI - Binomial Options – **8480+:** Test by Intel as of 10/7/2022. 1-node, 2x Intel Xeon Platinum 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Binomial Options v1.1, Black Scholes v1.4, Monte Carlo v1.2. **8380:** Test by Intel as of 10/7/2022. 1-node, 2x Intel Xeon Platinum 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Binomial Options v1.1, Black Scholes v1.4, Monte Carlo v1.2.

- 1.64x Manufacturing - ESI Group OpenFOAM – **8480+:** Test by Intel as of 9/2/2022. 1-node, 2x Intel Xeon Platinum 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode revision=0xaa0000a0, CentOS Stream 8, Linux version 4.18.0-365.el8.x86_64, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations. **8380:** Intel Xeon Platinum 8380: Test by Intel as of 9/2/2022. 1-node, 2x Intel Xeon Platinum 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode revision=0xd000270, Rocky Linux 8.6, Linux version 4.18.0-372.19.1.el8_6.crt1.x86_64, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations.

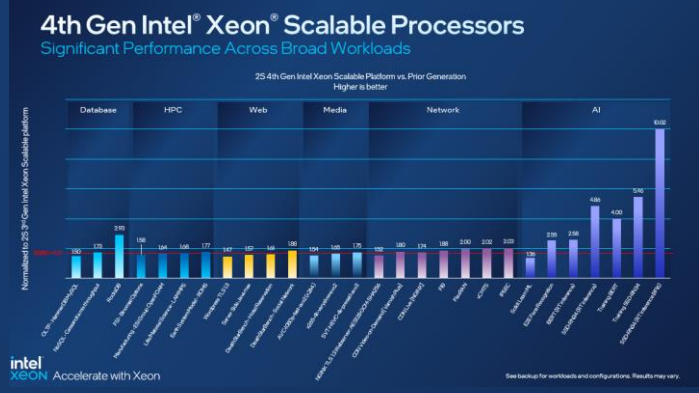
Resources and Configurations



Significant Performance Across Broad Workloads (2-6)

- 1.68x Life/Material Science – LAMMPS – **8480+:** Test by Intel as of 9/29/2022. 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LAMMPS v2021-09-29 cmkl:2022.1.0, icc:2021.6.0, impi:2021.6.0, tbb:2021.6.0; threads/core:; Turbo:off; BuildKnobs:-O3 -ip -xCORE-AVX512 -g -debug inline-debug-info -qopt-zmm-usage=high; **8380:** Test by Intel as of 10/11/2022. 1-node, 2x Intel Xeon Platinum 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LAMMPS v2021-09-29 cmkl:2022.1.0, icc:2021.6.0, impi:2021.6.0, tbb:2021.6.0; threads/core:; Turbo:on; BuildKnobs:-O3 -ip -xCORE-AVX512 -g -debug inline-debug-info -qopt-zmm-usage=high;
- 1.77x Earth System Model – ROMS – **8480+:** Test by Intel as of 10/12/2022. 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+, HT On, Turbo On, NUMA configuration SNC4, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, ROMS V4 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -heap-arrays -xCORE-AVX512 -qopt-zmm-usage=high -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", ROMS V4. **8380:** Test by Intel as of 10/12/2022. 1-node, 2x Intel Xeon Platinum 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, ROMS V4 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -heap-arrays -xCORE-AVX512 -qopt-zmm-usage=high -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", ROMS V4.
- 1.47x WordPress TLS 1.3 – **8490H:** 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8490H on ArcherCity with 1024 GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000070, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.5T INTEL SSDPF21Q016TB, 1x Ethernet Controller I225-LM, Wordpress 5.6, PHP 8.0.18 (fpm-fcgi), mysqld Ver 10.3.37-MariaDB for Linux on x86_64, Siege (docker image) SIEGE 2.78,hhvm-perf (docker image) v2.0.0 nginx (docker image) 1.20.1, test by Intel on 9/23/2022. **8380:** 1-node, 2x Intel(R) Xeon(R) Platinum 8380 on Whitley with 512 GB (16 slots/ 32GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 3.5T INTEL SSDPF2KX038TZ, 2x Ethernet Controller X710 for 10GBASE-T, Wordpress 5.6, PHP 8.0.18 (fpm-fcgi), mysqld Ver 10.3.37-MariaDB for Linux on x86_64, Siege (docker image) v2.78,hhvm-perf (docker image) v2.0.0 nginx (docker image) v1.20.1, test by Intel on 9/20/2022.
- 1.57x Server-side Java* Max JOPS – **8490H:** 1-node, pre-production platform with 2x Intel® Xeon® Platinum 8490H on ArcherCity with 1024 GB (16 slots/ 64GB/ DDR5 4800[4800]) total memory, ucode 0x2b0000a1, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-52-generic, 1x 7T INTEL SSDPF2KX076TZ, Server-side Java* (Max JOPS), JDK17, test by Intel on 10/30/2022. **8380:** 1-node, 2x Intel® Xeon® Platinum 8380 on Intel Rack Mount Chassis with 512 GB (16 slots/ 32GB/ DDR4 3200[3200]) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, Server-side Java* (Max JOPS), JDK17, test by Intel on 11/2/2022.

Resources and Configurations



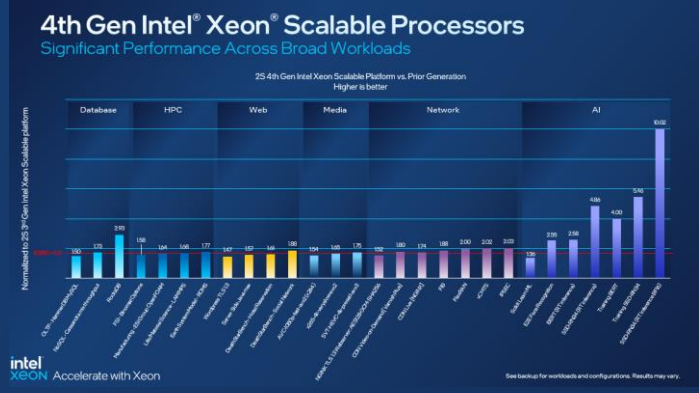
Significant Performance Across Broad Workloads (3-6)

- 1.61x DeathStarBench – Hotel Reservation: **8480+:** 4 (1master, 3worker)-node, each-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ on Intel ArcherCity with GB (32 slots/ 32GB/ DDR5 4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-52-generic with intel_iommu=off, 1x 54.9G INTEL SSDPEK1A058GA, 4x Ethernet Controller E810-C for QSFP, 2x Ethernet Controller XXV710 for 25GbE SFP28, DeathStarBench hotelReservation 1.0 [lianhao/dsbpp_hotel_reserve:1.0], Golang 1.17.3, GNU C Library 2.31-13+deb11u2, ice 5.15.0-52-generic, Kubernetes 1.23.6, Containerd 1.6.6 CRI-RM 0.7.0, Cilium: 1.11.7, gRPC-go 1.1, Consul 1.9.2, Memcached 1.6.8, MongoDB 4.4.3, Traffic generator open loop wrk2 included in DSB: mixed-workload_type_1.lua, 4 instance, 6 replica/instance, 4 wrk2 instance, 48 wrk2 thread/instance, 1920 wrk2 connection/instance, 70k wrk2 input rate/instance., test by Intel on 11/9/2022 . **8360Y:** 4 (1master, 3worker)-node, each-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8360Y on Intel M50CYP2SBSTD with GB (32 slots/ 16GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-52-generic with intel_iommu=off, 1x 223.6G KINGSTON SA400M8, 8x 2.9T INTEL SSDPE2KE032T8, 2x Ethernet Controller E810-C for QSFP, 2x BCM57416 NetXtreme-E Dual-Media 10G RDMA Ethernet Controller, DeathStarBench hotelReservation 1.0 [lianhao/dsbpp_hotel_reserve:1.0], Golang 1.17.3, GNU C Library 2.31-13+deb11u2, ice 5.15.0-52-generic, Kubernetes 1.23.6, Containerd 1.6.6 CRI-RM 0.7.0, Cilium: 1.11.7, gRPC-go 1.1, Consul 1.9.2, Memcached 1.6.8, MongoDB 4.4.3, Traffic generator open loop wrk2 included in DSB: mixed-workload_type_1.lua, 4 instance, 6 replica/instance, 4 wrk2 instance, 32 wrk2 thread/instance, 1920 wrk2 connection/instance, 44.5k wrk2 input rate/instance., test by Intel on 11/9/2022. https://github.com/delimitrou/DeathStarBench#publications

- 1.88x DeathStarBench – Social Network: **8480+:** 4 (1master, 3worker)-node, each-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ on QuantaGrid D54Q-2U with GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000081 , HT on, Turbo on, CentOS Linux release 8.4.2105, 6.0.6, 1x 2.9T INTEL SSDPE2KE032T7, 1x 893.8G AVAGO JBOD, 2x Ethernet Controller X710 for 10GBASE-T, 2x Ethernet Controller E810-C for QSFP, DeathStarBench Social Network, wrk2 - load generator, ICE driver (CVL): 6.0.6, Cilium CNI - 1.11.4, Kubernetes - 1.21.14, ContainerD - 1.4.12, deathstarbench/social-network-microservices:0.0.8, nginx-thrift: yg397/openresty-thrift:xenial, memcached:1.6.7, mongo:4.4.6, redis 7.0.5, dataset: DeathStarBench/socialNetwork/datasets/social-graph/socfb-Reed98/ , test by Intel on 11/2/2022. **8360Y:** 4 (1master, 3worker)-node, each-node, 2x Intel(R) Xeon(R) Platinum 8360Y on Intel Whitley with GB (16 slots/ 32GB/DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, CentOS Linux release 8.4.2105, 6.0.6, 1x 894.3G INTEL SSDSC2KG96, 2x Ethernet Controller X710 for 10GBASE-T, 1x Ethernet Controller E810-C for QSFP, DeathStarBench Social Network, wrk2 - load generator, ICE driver (CVL): 6.0.6, Cilium CNI - 1.11.4, Kubernetes - 1.21.14, ContainerD - 1.4.12, deathstarbench/social-network-microservices:0.0.8, nginx- thrift: yg397/openresty-thrift:xenial, memcached:1.6.7, mongo:4.4.6, redis 7.0.5, dataset: DeathStarBench/socialNetwork/datasets/social- graph/socfb- Reed98/ , test by Intel on 11/2/2022. https://github.com/delimitrou/DeathStarBench#publications

- 1.54x for AVC-1080p-fast-avx2 (x264)
  1.65x for x265-4k-veryslow-avx2
  1.75x for SVT-HEVC-4k-preset1-avx3
  8490H: 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8490H on Rack Mount Chassis with 1024 GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000070, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-48-generic, 1x 7T INTEL SSDPF2KX076TZ, 1x 1.5T INTEL SSDPF21Q016TB, 1x Ethernet Controller I225-LM, FFmpeg Result (fps), x264 Version=0.164.x, x265 Version=Release_3.5, SVT-HEVC Version=v1.3.0, SVT-AV1 Version=v0.8.7, test by on 09/30/2022. **8380:** 1-node, 2x Intel(R) Xeon(R) Platinum 8380 on Rack Mount Chassis with 512 GB (16 slots/ 32GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 7T INTEL SSDPF2KX076TZ, 1x 1.5T INTEL SSDPF21Q016TB, 1x Ethernet Controller I225-LM, FFmpeg Result (fps), x264 Version=0.164.x, x265 Version=Release_3.5, SVT-HEVC Version=v1.3.0, SVT-AV1 Version=v0.8.7, test by on 09/20/2022.

Resources and Configurations



Significant Performance Across Broad Workloads (4-6)

- 1.52x NGINX TLS 1.3 Webserver: **6338N:** Test by Intel on 10/18/2022. 1-node, 2x Intel® Xeon® Gold 6338N CPU @ 2.20GHz, 32 cores on Supermicro SYS-740GP-TNRT, HT Off, Turbo On, Total Memory 256GB (16x16GB DDR4 3200 MT/s [2666 MT/s]), BIOS 1.4, microcode 0xd000375, 4x Ethernet Controller E810-C for QSFP, 2x Ethernet Controller 10G X550T, 1x 223.6G INTEL SSDSC2KB240G8, Ubuntu 22.04 LTS, 5.15.0-27-generic. **6428N:** CPS with and without QAT: Test by Intel on 10/18/2022. 1-node, pre-production platform with 2x Intel® Xeon® Gold 6428N, 32 cores on Archer City, HT On, Turbo On, Total Memory 512GB (16x32GB DDR5 4800 MT/s [4000 MT/s]), BIOS EGSDCRB1.86B.8612.P03.2208120625, microcode 0xab000060, 1x Ethernet Controller I225-LM, 6x Ethernet Controller E810-C for QSFP, 1x 223.6G INTEL SSDSC2BB240G4, 1x 223.6G INTEL SSDSC2KB240G8, 1x 240M Disk, Ubuntu 22.04 LTS, 5.15.0-27-generic, NGINX (async mode nginx 0.4.7), GCC 11.2.0, openssl 1.1.1m, qatengine v0.6.14 (Optimized SW and QAT HW), IPsecmb v1.2 (Optimized SW), IPP-Crypto ipp-crypto_2021_5 (Optimized SW), QAT driver QAT.20.L.0.9.5.

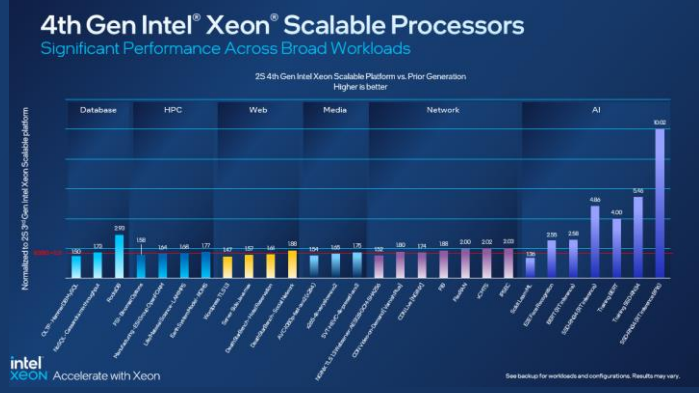
- 1.80x CDN Video-on-Demand [Varnish Plus]: **6448N:** Test by Intel as of 10/14/22. 1-node, pre-production platform with 2x Intel® Xeon® Gold 6438N Processor, 32 cores, HT On, Turbo On, Total Memory 256 GB (16 slots/ 16 GB/ 4800 MT/s), 16x Intel® P5510, 4x Intel® E810-2CQDA2, BIOS EGSDCRB1.SYS.0087.D13.2208261709 (ucode 0x2b000070), RHEL 8.6, kernel 4.18.0-372.26.1.el8_6.x86_64, gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-10), OpenSSL 1.1.1k FIPS 25 Mar 2021, varnish-plus-6.0.10r3 revision 4f67b6ec0d63f04560913cc7e195a3919bdf0366, wrk master 02/07/2021(keep alive OR connection: close, 400 OR 4000 OR 20000 total connections) Throughput measured with 100% Transport Layer Security (TLS) traffic with 93.3% cache hit ratio. **6338N:** Test by Intel as of 10/14/22. 1-node, 2x Intel® Xeon® Gold 6338N Processor, 32 cores, HT On, Turbo On, Total Memory 256 GB (16 slots/ 16 GB/ 3200 MT/s @ 2666 MT/s), 12x Intel® P5510, 4x Mellanox MCX516A-CDAT, BIOS 1.4 (ucode 0xd000375), RHEL 8.6, kernel 4.18.0-372.26.1.el8_6.x86_64, gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-10), OpenSSL 1.1.1k FIPS 25 Mar 2021, varnish-plus-6.0.10r3 revision 4f67b6ec0d63f04560913cc7e195a3919bdf0366, wrk master 02/07/2021 (keep alive OR connection: close, 400 OR 4000 OR 20000 total connections) Throughput measured with 100% Transport Layer Security (TLS) traffic with 93.3% cache hit ratio.

- 1.74X CDN live Nginx: **6438N:** Test by Intel as of 10/10/22. 1-node, pre-production platform with 2x Intel® Xeon® Gold 6438N Processor, 32 cores, HT On, Turbo On Total Memory 256 GB (16 slots/ 16 GB/ 4800 MT/s), Total Persistent Memory 2048 GB (16 slots/ 128 GB/ 4400 MT/s, App-Direct-Interleaved), 4x Intel® E810-2CQDA2, BIOS EGSDCRB1.SYS.0087.D13.2208261709 (ucode 0x2b000070), Ubuntu 22.04, kernel 5.15.0-48-generic, gcc (Ubuntu 11.2.0-19ubuntu1) 11.2.0, OpenSSL 3.0.2 15 Mar 2022 (Library: OpenSSL 3.0.2 15 Mar 2022), NGINX 1.22.0, wrk master 02/07/2021 (keep alive OR connection: close, 400 OR 4000 OR 20000 total connections) Throughput measured with 100% Transport Layer Security (TLS) traffic with 93.3% cache hit ratio. **6338N:** Test by Intel as of 10/10/22. 1-node, 2x Intel® Xeon® Gold 6338N Processor, 32 cores, HT On, Turbo On, Total Memory 256 GB (16 slots/ 16 GB/ 3200 MT/s), Total Persistent Memory 2048 GB (16 slots/ 128 GB/ 3200 MT/s, App-Direct-Interleaved), 4x Mellanox MCX516A-CDAT, BIOS 1.4 (ucode 0xd000375), Ubuntu 22.04, kernel 5.15.0-48-generic, gcc (Ubuntu 11.2.0-19ubuntu1) 11.2.0, OpenSSL 3.0.2 15 Mar 2022 (Library: OpenSSL 3.0.2 15 Mar 2022), NGINX 1.22.0, wrk master 02/07/2021 (keep alive OR connection: close, 400 OR 4000 OR 20000 total connections) Throughput measured with 100% Transport Layer Security (TLS) traffic with 93.3% cache hit ratio.

- 1.88x FIB [64B] higher VPP IPv4 FIB throughput **8470N:** 1-node, pre-production platform with 2(1 used)x Intel Xeon Platinum 8470N on Archer City with 512 GB (16 slots/ 32GB/ DDR5-4800) total memory, ucode 0xab0000c0, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.35, 1x INTEL SSDSC2KB240G8, x 5x Gen4 x16 PCIe NICs Intel Ethernet Controller E810-2CQDA2 (total 1000Gb/s), VPP 22.06-release FIB ipv4 router, GCC 9.4, Dataset size 64B / 512B, Ix Network 9.00.1900.17, test by Intel on 9/30/2022. **6338N:** 1-node, 2(1 used)x Intel Xeon Gold 6338N on Wilson City with 256 GB (16 slots/ 16GB/ DDR4-3200[2666]) total memory, ucode 0xd000375, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.35, 1x INTEL SSDSC2KB240G8, x 4x Gen4 x16 PCIe NICs Intel Ethernet Controller E810-2CQDA2 (total 800Gb/s), VPP 22.06-release FIB ipv4 router, GCC 9.4, Dataset size 64B / 512B, Ix Network 9.00.1900.17, test by Intel on 10/5/2022.

- 2.00x capacity for FlexRAN : 4th Gen Intel Xeon Scalable Processor delivers up to twice the capacity at the same power envelope for vRAN workloads vs. 3rd Gen Intel Xeon Scalable processors, enabling Communications Service Providers up to double the performance-per-watt to meet their critical performance, scaling and energy efficiency requirements. Estimated as of 8/30/2022 based on 4th generation Intel® Xeon® Scalable processor architecture improvements vs 3rd generation Intel® Xeon® Scalable processor at similar core count, socket power and frequency on a test scenario using FlexRAN™ software. Results may vary.

Resources and Configurations



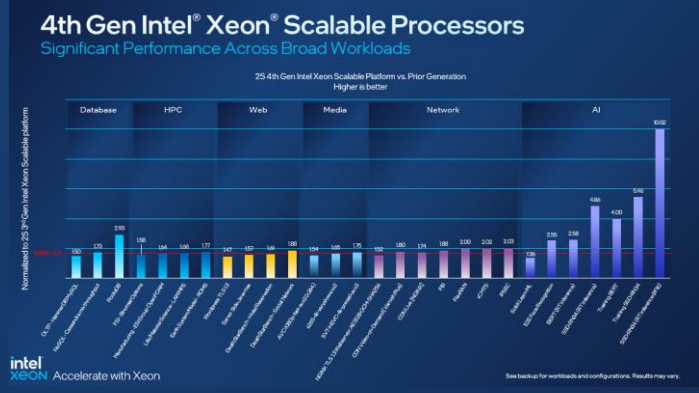
Significant Performance Across Broad Workloads (5-6)

- 2.02x better vCMTS - **8470N:** 1-node, pre-production platform with 2(1 used)x Intel Xeon Platinum 8470N on Archer City with 512 GB (16 slots/ 32GB/ DDR5-4800[4800]) total memory, ucode 0xab000080, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KB240G8, 1x Ethernet Controller E810-C for QSFP, vCMTS 22.10 beta, DPDK 22.03, GCC 11.2.0, DPDK 22.03, Collectd 5.12.0, Grafana 8.5.3, Prometheus 2.0.0, test by Intel on 9/20/2022. **6338N:** 1-node, 2(1 used)x Intel Xeon Gold 6338N on Supermicro X12DPG-QT6 with 512 GB (16 slots/ 32GB/ DDR4-3200[2666]) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KB240G8, 1x Ethernet Controller E810-C for QSFP, vCMTS 22.10 beta, DPDK 22.03, GCC 11.2.0, DPDK 22.03, Collected 5.12.0, Grafana 8.5.3, Prometheus 2.0.0, test by Intel on 10/10/2022.

- 2.03x higher IPsec throughput – **8470N:** 1-node, pre-production platform with 2(1 used)x Intel Xeon Platinum 8470N on Archer City with 512 GB (16 slots/ 32GB/ DDR5-4800) total memory, ucode 0xab0000c0, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.35, 1x INTEL SSDSC2KB240G8, x 5x Gen4 x16 PCIe NICs Intel Ethernet Controller E810-2CQDA2 (total 1000Gb/s), VPP 22.06-release 1 tunnel per VF port, GCC 9.4, Intel-ipsec-mb libraries v1.1 , Dataset size 512B / 1420B, IxNetwork 9.00.1900.17, test by Intel on 9/30/2022.
6338N: 1-node, 2(1 used)x Intel Xeon Gold 6338N on Wilson City with 256 GB (16 slots/ 16GB/ DDR4-3200[2666]) total memory, ucode 0xd000375, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.35, 1x INTEL SSDSC2KB240G8, x 4x Gen4 x16 PCIe NICs Intel Ethernet Controller E810-2CQDA2 (total 800Gb/s), VPP 22.06-release 1 tunnel per VF port, GCC 9.4, Intel-ipsec-mb libraries v1.1 , Dataset size 512B / 1420B, Ix Network 9.00.1900.17, test by Intel on 10/5/2022.

- 2.55x E2E Face Recognition – **8480+:** 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ on Intel Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b000041 , HT off, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-48-generic, INTEL SSDSC2KG011.7T , E2E Face Recognition with SSD-MobileNet + ResNet50_v1.5 inference , FP32/BF16/INT8, BS1, maintain 30fps per stream, Intel-tensorflow-avx512 2.10.0, ssd-mobilenet, Resnet50_v1.5, oneDNN v2.6.0, Dataset: Chelsea_celebration.mp4 (7969 frames), test by Intel on 10/21/2022. **8380:** 1-node, 2x Intel(R) Xeon(R) Platinum 8380 CPU on Intel WHITLEY with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT off, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-48-generic, INTEL SSDSC2BA80, E2E Face Recognition with SSD-MobileNet + ResNet50_v1.5 inference , FP32/BF16/INT8, BS1, maintain 30fps per stream, Intel-tensorflow-avx512 2.10.0, ssd-mobilenet, Resnet50_v1.5, oneDNN v2.6.0, Dataset: Chelsea_celebration.mp4 (7969 frames), test by Intel on 10/21/2022.

Resources and Configurations



Significant Performance Across Broad Workloads (6-6)

- 2.58x BERT (RT Inference)
  4.00x Training BERT
  8480+: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), Bert Large, Inf: SQuAD1.1 (seq len=384), bs=1 [4cores/instance], bs=n [1socket/instance], Inference: bs: fp32=1,56, amx bf16=1,16, amx int8=1,56, Trg: Wikipedia 2020/01/01 (seq len =512), bs:fp32=28, amx bf16=56 [1 instance, 1socket], Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.
  8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, Bert Large, Inf: SQuAD1.1 (seq len=384), bs=1 [4cores/instance], bs=n [1socket/instance], Inference: bs: fp32=1,56, int8=1,56, Trg: Wikipedia 2020/01/01 (seq len =512), bs:fp32=28, amx bf16=56 [1 instance, 1socket], Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

- 4.86x SSD-RN34 (RT inference) - **8480+:** 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, amx bf16=1,112, amx int8=1,112, Training bs:fp32/amx bf16=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.
  8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, int8=1,112, Training bs:fp32=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

- 5.46x Training SSD- RN34 (BF16)
  10.02x SSD-RN34(RT inference BF 16)
  8480+: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, amx bf16=1,112, amx int8=1,112, Training bs:fp32/amx bf16=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.
  8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, int8=1,112, Training bs:fp32=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

Resources and Configurations

Improving Performance When Response Time Matters (1-2)

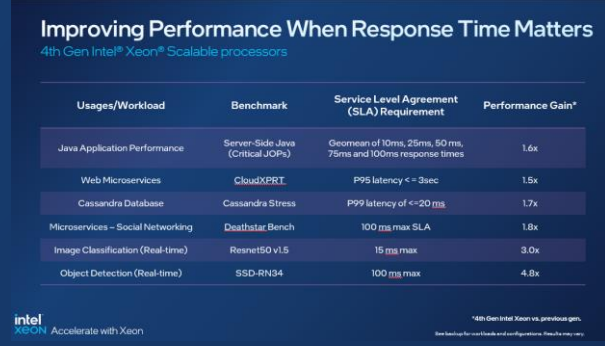
- 1.6x – Server-side Java* 1-node, on pre-production platform with 2x Intel® Xeon® Platinum 8490H on ArcherCity with 1024 GB (16 slots/ 64GB/ DDR5 4800[4800]) total memory, ucode 0x2b0000a1, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-52-generic, Server-side Java* (Critical JOPS), JDK17, test by Intel on 10/31/2022.
  1-node, 2x Intel® Xeon® Platinum 8380 on Intel Rack Mount Chassis with 512 GB (16 slots/ 32GB/ DDR4 3200[3200]) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-52-generic, Server-side Java* (Critical JOPS), JDK17, test by Intel on 11/2/2022.

- 1.5x – CloudXPRT… 8490H: 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8490 H on Archer City with GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000070, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x INTEL SSDPF21Q016TB, CloudXPRT v1.20, kubernetes v1.16.3, Golang v1.13.1, Ubuntu (Docker image) 20.04, Redis (Docker image) 5.0.8-buster, Cassandra (Docker image) 3.11.6, Nginx (Docker image) v1.17, test by Intel on 9/19/2022.
  8380: 1-node, 2x Intel(R) Xeon(R) Platinum 8380 on Whitley with GB (16 slots/ 32GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x INTEL SSDPF21Q016TB, CloudXPRT v1.20, kubernetes v1.16.3, Golang v1.13.1, Ubuntu (Docker image) 20.04, Redis (Docker image) 5.0.8-buster, Cassandra (Docker image) 3.11.6, Nginx (Docker image) v1.17, test by Intel on 8/25/2022.
  8180: 1-node, 2x Intel(R) Xeon(R) Platinum 8180 on Purley with GB (12 slots/ 32GB/ DDR4 2666) total memory, ucode 0x2006d05, HT on, Turbo on, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x INTEL SSDPF2KX038TZ, CloudXPRT v1.20, kubernetes v1.16.3, Golang v1.13.1, Ubuntu (Docker image) 20.04, Redis (Docker image) 5.0.8-buster, Cassandra (Docker image) 3.11.6, Nginx (Docker image) v1.17, test by Intel on 4/21/2022. Intel contributes to the development of benchmarks by participating in, sponsoring, and/or contributing technical support to various benchmarking groups, including the BenchmarkXPRT Development Community administered by Principled Technologies.

- 1.7x – Cassandra Stress. 8490H: Test by Intel as of 10/18/22. 1-node, 2x Intel(R) Xeon(R) Platinum 8490H, 60 cores, HT On, Turbo On, Total Memory 512GB (16x32GB 4800 MT/s [4800 MT/s]), BIOS EGSDCRB1.FEI.0088.D16.2209090804, microcode 0xababc0a0, 1x Ethernet Controller I225-LM, 4x Ethernet Controller X710/X557-AT 10GBASE-T, 1x 1.5T INTEL SSDSC2BB01, 8x 1.5T INTEL SSDPF21Q016TB, Ubuntu 22.04.1 LTS, 5.19.10-051910-generic, compiler gcc version 11.2.0, cassandra-stress version 4.0.6, jdk version 14 build 14+36-1461
  8380: Test by Intel as of 10/17/22, 1-node, 2x Intel(R) Xeon(R) Platinum 8380 CPU @ 2.30GHz, 40 cores, HT On, Turbo On, Total Memory Installed 512GB (16x32GB DDR4 3200 MT/s [3200 MT/s]), Total Memory Used 256GB as 256GB blocked, BIOS SE5C620.86B.01.01.0005.2202160810, microcode 0xd000375, 2x Ethernet Controller X710 for 10GBASE-T, 1x 223.6G KINGSTON SA400M8240G, 4x 1.5T INTEL SSDPF21Q016TB, Ubuntu 22.04.1 LTS, 5.19.10-051910-generic, compiler gcc version 11.2.0, cassandra-stress version 4.0.6, jdk version 14 build 14+36-146

- 1.8X - DeathStar Bench. 8480+: 4 (1master, 3worker)-node, each-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum **8480+** on QuantaGrid D54Q-2U with GB (16 slots/ 64GB/ DDR5 4800) total memory, ucode 0x2b000081 , HT on, Turbo on, CentOS Linux release 8.4.2105, 6.0.6, 1x 2.9T INTEL SSDPE2KE032T7, 1x 893.8G AVAGO JBOD, 2x Ethernet Controller X710 for 10GBASE-T, 2x Ethernet Controller E810-C for QSFP, DeathStarBench Social Network, wrk2 - load generator, ICE driver (CVL): 6.0.6, Cilium CNI - 1.11.4, Kubernetes - 1.21.14, ContainerD - 1.4.12, deathstarbench/social-network-microservices:0.0.8, nginx-thrift: yg397/openresty-thrift:xenial, memcached:1.6.7, mongo:4.4.6, redis 7.0.5, dataset: DeathStarBench/socialNetwork/datasets/social-graph/socfb-Reed98/ , test by Intel on 11/2/2022.
  8360Y: 4 (1master, 3worker)-node, each-node, 2x Intel(R) Xeon(R) Platinum 8360Y on Intel Whitley with GB (16 slots/ 32GB/DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo on, CentOS Linux release 8.4.2105, 6.0.6, 1x 894.3G INTEL SSDSC2KG96, 2x Ethernet Controller X710 for 10GBASE-T, 1x Ethernet Controller E810-C for QSFP, DeathStarBench Social Network, wrk2 - load generator, ICE driver (CVL): 6.0.6, Cilium CNI - 1.11.4, Kubernetes - 1.21.14, ContainerD - 1.4.12, deathstarbench/social-network-microservices:0.0.8, nginx- thrift: yg397/openresty-thrift:xenial, memcached:1.6.7, mongo:4.4.6, redis 7.0.5, dataset: DeathStarBench/socialNetwork/datasets/social- graph/socfb- Reed98/ , test by Intel on 11/2/2022. https://github.com/delimitrou/DeathStarBench#publications

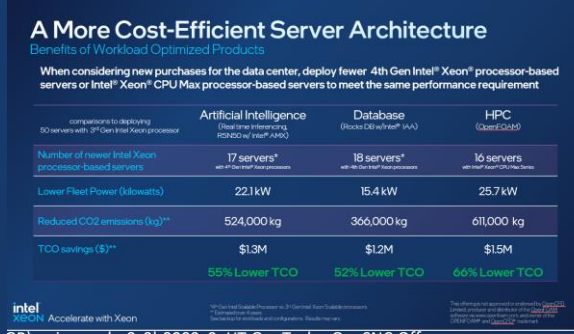
Resources and Configurations -



Improving Performance When Response Time Matters (2-2)

- 3.0x –Resnet50 V1.5(I used the TF config) **8480+**: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), ResNet50, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, amx bf16=1,80, amx int8=1,116, Training bs:fp32=1024 [1 instance, 1socket], Framework: https://github.com/intel-innersource/frameworks.ai.tensorflow.private-tensorflow/tree/spr_ww42_2022; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, TF: 2.11, OneDNN: v2.7, test by Intel on 10/24/2022.
  8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, ResNet50, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, int8=1,116, Training bs:fp32=1024 [1 instance, 1socket], Framework: https://github.com/intel-innersource/frameworks.ai.tensorflow.private-tensorflow/tree/spr_ww42_2022; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, TF: 2.11, OneDNN: v2.7, test by Intel on 10/24/2022.
- 4.8X- SSD-RN34 image classification (Real-time): **8480+**: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, amx bf16=1,112, amx int8=1,112, Training bs:fp32/amx bf16=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.
  8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, SSD-ResNet34, Inference: bs=n [1socket/instance], bs: fp32=1,112, int8=1,112, Training bs:fp32=224 [1 instance, 1socket], Coco 2017, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

Resources and Configurations



A More Cost-Efficient Server Architecture

ResNet50 Image Classification

New Configuration: 1-node, 2x pre-production 4th Gen Intel® Xeon® Scalable 8490H processor (60 core) with Intel® Advanced Matrix Extensions (Intel AMX), on pre-production SuperMicro SYS-221H-TNR with 1024GB DDR5 memory (16x64 GB), microcode 0x2b0000c0, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 AMX 1 core/instance (max. 15ms SLA), using physical cores, tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable 8380 Processor (40 cores) on SuperMicro SYS-220U-TNR , DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, CentOS Stream 8, 5.19.16-301.fc37.x86_64, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Intel TF 2.10, AI Model=Resnet 50 v1_5, best scores achieved: BS1 INT8 2 cores/instance (max. 15ms SLA), using physical cores, tested by Intel November 2022.

For a 50 server fleet of 3rd Gen Xeon 8380 (RN50 w/DLBoost), estimated as of November 2022:
CapEx costs: $1.64M
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $739.9K
Energy use in kWh (4 year, per server): 44627, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

For a 17 server fleet of 4th Gen Xeon 8490H (RN50 w/AMX), estimated as of November 2022:
CapEx costs: $799.4K
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $275.3K
Energy use in kWh (4 year, per server): 58581, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

RocksDB

New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable 8490H Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), on pre-production Intel platform and software, HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0x2b0000a1, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.2.1,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable 8380 Processor (40 cores) on SuperMicro SYS-220U-TNR , HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x3.84TB P5510 NVMe, 10GbE x540-AT2, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel November 2022.

For a 50 server fleet of 3rd Gen Xeon 8380 (RocksDB), estimated as of November 2022:
CapEx costs: $1.64M
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $677.7K
Energy use in kWh (4 year, per server): 32181, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

For a 18 server fleet of 4th Gen Xeon 8490H (RockDB w/IAA), estimated as of November 2022:
CapEx costs: $846.4K
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $260.6K
Energy use in kWh (4 year, per server): 41444, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

OpenFOAM
New Configuration: 1-node, 2x pre-production 4th Gen Intel Xeon CPU Max Series (56 cores) on pre-production Intel platform and software, HT On, Turbo On, SNC4 mode, Total Memory 128 GB (8x16GB HBM2 3200MT/s), microcode 0x2c000020, 1x3.5TB INTEL SSDPF2KX038TZ NVMe, CentOS Stream 8, 5.19.0-rc6.0712.intel_next.1.x86_64+server, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations, Tools: ifort:2021.6.0, icc:2021.6.0, impi:2021.6.0, tested by Intel December 2022. Baseline: 1-node, 2x production 3rd Gen Intel Xeon Scalable 8380 Processor (40 cores) on SuperMicro SYS-220U-TNR, HT On, Turbo On, 512GB (16x32GB DDR4 3200 MT/s), microcode 0xd000375, 1x2.9TB INTEL SSDPE2KE032T8 NVMe, CentOS Stream 8, 4.18.0-408.el8.x86_64, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations, Tools: ifort:2021.6.0, icc:2021.6.0, impi:2021.6.0, tested by Intel December 2022

For a 50 server fleet of 3rd Gen Xeon 8380 (OpenFOAM), estimated as of December 2022:
CapEx costs: $1.50M
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $780.3K
Energy use in kWh (4 year, per server): 52700, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

For a 16 server fleet of Intel Xeon CPU Max Series 56 core, estimated as of December 2022:
CapEx costs: $507.2K
OpEx costs (4 year, includes power and cooling utility costs, infrastructure and hardware maintenance costs): $274.9K
Energy use in kWh (4 year, per server): 74621, PUE 1.6
Other assumptions: utility cost $0.1/kWh, kWh to kg CO2 factor 0.42394

Resources and Configuration (1 of 7 for HPC)

HPCG
- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, HPCG from MKL_v2022.1.0
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, DDR5 Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Ubuntu 22.04.1 LTS, Linux version 5.15.0-50-generic, HPCG from MKL_v2022.1.0
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, HPCG from MKL_v2022.1.0

HPL
- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, HPL from MKL_v2022.1.0, 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, HPL v2.3_BLIS-3.0_AMD_OFFICIAL
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, DDR5 Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Ubuntu 22.04.1 LTS, Linux version 5.15.0-50-generic, HPL from MKL_v2022.1.0
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, HPL from MKL_v2022.1.0

Stream Triad
- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Stream v5.10
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Stream v5.10
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Stream v5.10

MPAS-A (MPAS-A V7.3 60-km dynamical core)
- Intel® Xeon® 8380: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, MPAS-A V7.3 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-O3 -march=core-avx2 -convert big_endian -free -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", MPAS-A V7.3
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, NUMA configuration SNC4, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, MPAS-A V7.3 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-O3 -march=core-avx2 -convert big_endian -free -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", MPAS-A V7.3
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/12/22. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, MPAS-A V7.3 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-O3 -march=core-avx2 -convert big_endian -free -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", MPAS-A V7.3

Resources and Configuration (2 of 7 for HPC)

NEMO (GYRE_PISCES_25, BENCH ORCA-1)

- Intel® Xeon® 8380: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, NEMO v4.2 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-i4 -r8 -O3 -fno-alias -march=core-avx2 -fp-model fast=2 -no-prec-div -no-prec-sqrt -align array64byte -fimf-use-svml=true"
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, NUMA configuration SNC4, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, NEMO v4.2 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-i4 -r8 -O3 -fno-alias -march=core-avx2 -fp-model fast=2 -no-prec-div -no-prec-sqrt -align array64byte -fimf-use-svml=true".
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, NEMO v4.2 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-i4 -r8 -O3 -fno-alias -march=core-avx2 -fp-model fast=2 -no-prec-div -no-prec-sqrt -align array64byte -fimf-use-svml=true"

ROMS (benchmark3 (2048x256x30), benchmark3 (8192x256x30))

- Intel® Xeon® 8380: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, ROMS V4 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -heap-arrays -xCORE-AVX512 -qopt-zmm-usage=high -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", ROMS V4
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, NUMA configuration SNC4, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, ROMS V4 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -heap-arrays -xCORE-AVX512 -qopt-zmm-usage=high -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", ROMS V4
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, ROMS V4 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -heap-arrays -xCORE-AVX512 -qopt-zmm-usage=high -align array64byte -fimf-use-svml=true -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low", ROMS V4

WRF (CONUS 2.5KM)

- Intel® Xeon® 8380: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, WRF v3.9.1.1 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -xCORE-AVX512 -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low -w -ftz -align array64byte -fno-alias -fimf-use-svml=true -inline-max-size=12000 -inline-max-total-size=30000 -vec-threshold0 -qno-opt-dynamic-align "
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/12/2022. 1-node, 4th Gen Intel® Xeon® Scalable Processor, HT On, Turbo On, NUMA configuration SNC4, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, WRF v3.9.1.1 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -xCORE-AVX512 -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low -w -ftz -align array64byte -fno-alias -fimf-use-svml=true -inline-max-size=12000 -inline-max-total-size=30000 -vec-threshold0 -qno-opt-dynamic-align ".
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, WRF v3.9.1.1 build with Intel® Fortran Compiler Classic and Intel® MPI from 2022.3 Intel® oneAPI HPC Toolkit with compiler flags "-ip -O3 -xCORE-AVX512 -fp-model fast=2 -no-prec-div -no-prec-sqrt -fimf-precision=low -w -ftz -align array64byte -fno-alias -fimf-use-svml=true -inline-max-size=12000 -inline-max-total-size=30000 -vec-threshold0 -qno-opt-dynamic-align "

Resources and Configuration (3 of 7 for HPC)

YASK (Anelsatic Wave Propagation, ISO3DFD, Standard Staggered Grid)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, YASK v3.05.07
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode revision=0xaa0000a0, CentOS Stream 8, Linux version 4.18.0-365.el8.x86_64, YASK v3.05.07
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, YASK v3.05.07

FSI Kernels (Binomial Options, Black Scholes, Monte Carlo)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Binomial Options v1.1, Black Scholes v1.4, Monte Carlo v1.2
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Binomial Options v1.1, Black Scholes v1.4, Monte Carlo v1.2
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Binomial Options v1.1, Black Scholes v1.4, Monte Carlo v1.2

DeePMD (Multi-Instance Training)

- Intel® Xeon® 8380: Test by Intel as of 10/20/2022. 1-node, 2x Intel® Xeon® 8380 processor, Total Memory 256 GB, kernel 4.18.0-372.26.1.el8_6.crt1.x86_64, compiler gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-10), https://github.com/deepmodeling/deepmd-kit, Tensorflow 2.9, Horovod 0.24.0, oneCCL-2021.5.2, Python 3.9
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® 8480+, Total Memory 512 GB, kernel 4.18.0-365.el8_3x86_64, compiler gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-10), https://github.com/deepmodeling/deepmd-kit, Tensorflow 2.9, Horovod 0.24.0, oneCCL-2021.5.2, Python 3.9
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/12/2022. 1-node, 2x Intel® Xeon® Max 9480, Total Memory 128 GB (HBM2e at 3200 MHz), kernel 5.19.0-rc6.0712.intel_next.1.x86_64+server, compiler gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-13), https://github.com/deepmodeling/deepmd-kit, Tensorflow 2.9, Horovod 0.24.0, oneCCL-2021.5.2, Python 3.9

GROMACS (benchMEM, benchPEP, benchPEP-h, benchRIB, hecbiosim-3m, hecbiosim-465k, hecbiosim-61k, ion_channel_pme_large, lignocellulose_rf_large, rnase_cubic, stmv, water1.5M_pme_large, water1.5M_rf_large)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Converge GROMACS v2021.4_SP
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, GROMACS v2021.4_SP
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, GROMACS v2021.4_SP

Resources and Configuration (4 of 7 for HPC)

LAMMPS (Atomic Fluid, Copper, DPD, Liquid_crystal, Polyethylene, Protein, Stillinger-Weber, Tersoff, Water)

- Intel® Xeon® 8380: Test by Intel as of 10/11/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LAMMPS v2021-09-29 cmkl:2022.1.0, icc:2021.6.0, impi:2021.6.0, tbb:2021.6.0; threads/core:; Turbo:on; BuildKnobs:-O3 -ip -xCORE-AVX512 -g -debug inline-debug-info -qopt-zmm-usage=high;
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 9/29/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LAMMPS v2021-09-29 cmkl:2022.1.0, icc:2021.6.0, impi:2021.6.0, tbb:2021.6.0; threads/core:; Turbo:off; BuildKnobs:-O3 -ip -xCORE-AVX512 -g -debug inline-debug-info -qopt-zmm-usage=high;
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/29/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, LAMMPS v2021-09-29 cmkl:2022.1.0, icc:2021.6.0, impi:2021.6.0, tbb:2021.6.0; threads/core:; Turbo:off; BuildKnobs:-O3 -ip -xCORE-AVX512 -g -debug inline-debug-info -qopt-zmm-usage=high;

Quantum Espresso (AUSURF112, Water_EXX)

- Intel® Xeon® 8380: Test by Intel as of 9/30/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Quantum Espresso 7.0, AUSURF112, Water_EXX
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), ucode revision= 0x90000c0, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Quantum Espresso 7.0, AUSURF112, Water_EXX
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Quantum Espresso 7.0, AUSURF112, Water_EXX

VASP (CuC, Si, PdO4, PdO4_k221)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, VASP6.3.2
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, VASP6.3.2
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, VASP6.3.2

Altair AcuSolve (HQ Model)

- Intel® Xeon® 8380: Test by Intel as of 09/28/2022. 1-node, 2x Intel® Xeon® 8380, HT ON, Turbo ON, Quad, Total Memory 256 GB, BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode 0xd000270, Rocky Linux 8.6, kernel version 4.18.0-372.19.1.el8_6.crt1.x86_64, Altair AcuSolve 2021R2
- Intel® Xeon® 6346: Test by Intel as of 10/08/2022. 4-nodes connected via HDR-200, 2x Intel® Xeon® 6346, 16 cores, HT ON, Turbo ON, Quad, Total Memory 256 GB, BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode 0xd000270, Rocky Linux 8.6, kernel version 4.18.0-372.19.1.el8_6.crt1.x86_64, Altair AcuSolve 2021R2
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 09/28/2022. 1-node, 2x Intel® Xeon® 8480+, HT ON, Turbo ON, SNC4, Total Memory 512 GB, BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode 0xaa0000a0, CentOS Stream 8, kernel version 4.18.0-365.el8.x86_64, Altair AcuSove 2021R2
- Intel® Xeon® CPU Max Series: Test by Intel as of 10/03/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode 2c000020, CentOS Stream 8, kernel version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Altair AcuSolve 2021R2

Resources and Configuration (5 of 7 for HPC)

Altair RADIOSS (Neon1M @ 80 ms)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Altair RADIOSS 2022.2, Intel MPI 2021.7
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Altair RADIOSS 2022.2, Intel MPI 2021.7
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Altair RADIOSS 2022.2, Intel MPI 2021.7

Ansys Fluent (pump_2m, sedan_4m, rotor_3m, aircraft_wing_14m, combustor_12m, exhaust_system_33m)

- Intel® Xeon® 8380: Test by Intel as of 08/24/2022. 1-node, 2x Intel® Xeon® 8380, HT ON, Turbo ON, Quad, Total Memory 256 GB, BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode 0xd000270, Rocky Linux 8.6, kernel version 4.18.0-372.19.1.el8_6.crt1.x86_64, Ansys Fluent 2022R1
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 09/02/2022. 1-node, 2x Intel® Xeon® 8480+, HT ON, Turbo ON, SNC4, Total Memory 512 GB, BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode 0xaa0000a0, CentOS Stream 8, kernel version 4.18.0-365.el8.x86_64, Ansys Fluent 2022R1
- Intel® Xeon® CPU Max Series: Test by Intel as of 08/31/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo ON, SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode 2c000020, CentOS Stream 8, kernel version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Ansys Fluent 2022R1

Ansys LS-DYNA (ODB-10M)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s DDR4), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LS-DYNA R11
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of ww41'22. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, LS-DYNA R11
- Intel® Xeon® CPU Max Series: Test by Intel as of ww36'22. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, LS-DYNA R11

Ansys Mechanical (V22iter-1, V22iter-2, V22iter-3, V22iter-4, V22direct-1, V22direct-2, V22direct-3)

- Intel® Xeon® 8380: Test by Intel as of 08/24/2022. 1-node, 2x Intel® Xeon® 8380, HT ON, Turbo ON, Quad, Total Memory 256 GB, BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode 0xd000270, Rocky Linux 8.6, kernel version 4.18.0-372.19.1.el8_6.crt1.x86_64, Ansys Mechanical 2022 R2
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 09/02/2022. 1-node, 2x Intel® Xeon® 8480+, HT ON, Turbo ON, SNC4, Total Memory 512 GB DDR5 4800 MT/s, BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode 0xaa0000a0, CentOS Stream 8, kernel version 4.18.0-365.el8.x86_64, Ansys Mechanical 2022 R2
- Intel® Xeon® CPU Max Series: Test by Intel as of 08/31/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo ON, SNC4, Total Memory 512 GB DDR5 4800 MT/s, 128 GB HBM in cache mode (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode 2c000020, CentOS Stream 8, kernel version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Ansys Mechanical 2022 R2

Resources and Configuration (6 of 7 for HPC)

CONVE.RGE (SI8_engine_PFI_SAGE_transient_RAN)

- Intel® Xeon® 8380: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, NUMA configuration SNC2, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C620.86B.01.01.0006.2207150335, ucode revision=0xd000375, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Converge CFD 3.0.17
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/7/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, SNC4, Total Memory 512 GB (16x32GB 4800MT/s, DDR5), BIOS Version SE5C7411.86B.8713.D03.2209091345, ucode revision=0x2b000070, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, Converge CFD 3.0.17
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT ON, Turbo ON, NUMA configuration SNC4, Total Memory 128 GB (HBM2e at 3200 MHz), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, Converge CFD 3.0.17

OpenFOAM (Geomean of Motorbike 20M, Motorbike 42M)

- Intel® Xeon® 8380: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode revision=0xd000270, Rocky Linux 8.6, Linux version 4.18.0-372.19.1.el8_6.crt1.x86_64, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode revision=0xaa0000a0, CentOS Stream 8, Linux version 4.18.0-365.el8.x86_64, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, OpenFOAM 8, Motorbike 20M @ 250 iterations, Motorbike 42M @ 250 iterations

ParSeNet (SplineNet)

- Intel® Xeon® 8380: Test by Intel as of 10/18/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode revision=0xd000270, Rocky Linux 8.6, Linux version 4.18.0-372.19.1.el8_6.crt1.x86_64, ParSeNet (SplineNet), PyTorch 1.11.0, Torch-CCL 1.2.0, IPEX 1.10.0, MKL (20220804), oneDNN (v2.6.0)
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 10/18/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16x32GB 4800MT/s, Dual-Rank), BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode revision=0xaa0000a0, CentOS Stream 8, Linux version 4.18.0-365.el8.x86_64, ParSeNet (SplineNet), PyTorch 1.11.0, Torch-CCL 1.2.0, IPEX 1.10.0, MKL (20220804), oneDNN (v2.6.0)
- Intel® Xeon® CPU Max Series: Test by Intel as of 09/12/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, ParSeNet (SplineNet), PyTorch 1.11.0, Torch-CCL 1.2.0, IPEX 1.10.0, MKL (20220804), oneDNN (v2.6.0)

Siemens Simcenter Star-CCM+ (civil, HIMach10AoA10Sou, kcs_with_physics, lemans_poly_17m.amg, reactor, TurboCharger7M)

- Intel® Xeon® 8380: Test by Intel as of 25-Oct-22. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode revision=0xd000270, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, StarCCM+ 17.04.007, reactor 9m @ 20 iterations, lemans_poly_17m @ 20 iterations, civil 20m @ 20 iterations, TurboCharger7M @ 20 iterations, HIMach10AoA10Sou 6.4m @ 20 iterations, kcs_with_physics 3m @ 20 iterations
- 4th Gen Intel® Xeon® Scalable Processor: Test by Intel as of 14-Sep-22. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 1024 GB (16x64GB 4800MT/s, Dual-Rank), BIOS Version EGSDCRB1.86B.0083.D22.2206290535, ucode revision=0xaa000090, CentOS Stream 8, Linux version 4.18.0-394.el8.x86_64, StarCCM+ 17.04.007, reactor 9m @ 20 iterations, lemans_poly_17m @ 20 iterations, civil 20m @ 20 iterations, TurboCharger7M @ 20 iterations, HIMach10AoA10Sou 6.4m @ 20 iterations, kcs_with_physics 3m @ 20 iterations
- Intel® Xeon® CPU Max Series: Test by Intel as of 14-Sep-22. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, StarCCM+ 17.04.007, reactor 9m @ 20 iterations, lemans_poly_17m @ 20 iterations, civil 20m @ 20 iterations, TurboCharger7M @ 20 iterations, HIMach10AoA10Sou 6.4m @ 20 iterations, kcs_with_physics 3m @ 20 iterations

Resources and Configuration (7 of 7 for HPC)

CosmoFlow (training on 8192 image batches)

- Intel® Xeon® 8380: Test by Intel as of 06/07/2022. 1-node, 2x Intel® Xeon® 8380, 40 cores, HT On, Turbo On, Total Memory 512 GB (16 slots/ 32 GB/ 3200 MHz, DDR4), BIOS SE5C6200.86B.0022.D64.2105220049, ucode 0xd0002b1, OS Red Hat Enterprise Linux 8.5 (Ootpa), kernel 4.18.0-348.7.1.el8_5.x86_64, https://github.com/mlcommons/hpc/tree/main/cosmoflow, AVX-512, FP32, Tensorflow 2.9.0, horovod 0.23.0, keras 2.6.0, oneCCL-2021.4, oneAPI MPI 2021.4.0, ppn=8, LBS=16, ~25GB data, 16 epochs, Python 3.8
- 4th Gen Intel® Xeon® Scalable Processor (AMX BF16): Test by Intel as of 10/18/2022. 1-node, 2x Intel® Xeon® 8480+, HT On, Turbo On, Total Memory 512 GB (16 slots/ 32 GB/ 4800 MHz, DDR5), BIOS EGSDCRB1.86B.0083.D22.2206290535, ucode 0xaa0000a0, CentOS Stream 8, kernel 4.18.0-365.el8.x86_64, https://github.com/mlcommons/hpc/tree/main/cosmoflow, AMX, BF16, Tensorflow 2.9.1, horovod 0.24.3, keras 2.9.0.dev2022021708, oneCCL 2021.5, ppn=8, LBS=16, ~25GB data, 16 epochs, Python 3.8
- Intel® Xeon® Processor Max Series HBM (AMX BF16): Test by Intel as of 10/18/2022. 1 -node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, Total Memory 128 HBM and 512 GB DDR (16 slots/ 32 GB/ 4800 MHz), BIOS SE5C7411.86B.8424.D03.2208100444, ucode 0x2c000020, CentOS Stream 8, kernel 5.19.0-rc6.0712.intel_next.1.x86_64+server, https://github.com/mlcommons/hpc/tree/main/cosmoflow, AMX, BF16, TensorFlow 2.9.1, horovod 0.24.0, keras 2.9.0.dev2022021708, oneCCL 2021.5, ppn=8, LBS=16, ~25GB data, 16 epochs, Python 3.9

LAMMPS (Liquid Crystal) – Intel® Data Center GPU Max Series

- Intel® Xeon® 8380: Test by Intel as of 10/28/2022. 1-node, 2x Intel® Xeon® 8380 CPU, HT On, Turbo On, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version SE5C6200.86B.0020.P23.2103261309, ucode revision=0xd000270, Rocky Linux 8.6, Linux version 4.18.0-372.19.1.el8_6.crt1.x86_64
- Intel® Xeon® CPU Max Series HBM: Test by Intel as of 10/28/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, Total Memory 128 GB HBM2e, BIOS EGSDCRB1.DWR.0085.D12.2207281916, ucode 0xac000040, SUSE Linux Enterprise Server 15 SP3, Kernel 5.3.18, oneAPI 2022.3.0
- Intel® Data Center GPU Max Series with DDR Host: Test by Intel as of 10/28/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, Total Memory 1024 GB DDR5-4800 + 128 GB HBM2e, Memory Mode: Flat, HBM2e not used, 6x Intel® Data Center GPU Max Series, BIOS EGSDCRB1.DWR.0085.D12.2207281916, ucode 0xac000040, Agama pvc-prq-54, SUSE Linux Enterprise Server 15 SP3, Kernel 5.3.18, oneAPI 2022.3.0
- Intel® Data Center GPU Max Series with HBM Host: Test by Intel as of 10/28/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, Total Memory 128 GB HBM2e, 6x Intel® Data Center GPU Max Series, BIOS EGSDCRB1.DWR.0085.D12.2207281916, ucode 0xac000040, Agama pvc-prq-54, SUSE Linux Enterprise Server 15 SP3, Kernel 5.3.18, oneAPI 2022.3.0

HPCG Performance per Watt Comparison

- AMD EPYC 7773X: Test by Intel as of 10/7/2022. 1-node, 2x AMD EPYC , HT On, Turbo On, cTDP – 280, Total Memory 256 GB (16x16GB 3200MT/s, Dual-Rank), BIOS Version M10 rev5.22, ucode revision=0xa001224, Rocky Linux 8.6, Linux version 4.18.0-372.26.1.el8_6.crt1.x86_64, HPCG from MKL_v2022.1.0. Power calculated using CPU TDP and 7W per DIMM for 100 nodes for equal performance
- Intel® Xeon® CPU Max Series: Test by Intel as of 9/2/2022. 1-node, 2x Intel® Xeon® Max 9480, HT On, Turbo On, SNC4, Total Memory 128 GB (8x16GB HBM2 3200MT/s), BIOS Version SE5C7411.86B.8424.D03.2208100444, ucode revision=0x2c000020, CentOS Stream 8, Linux version 5.19.0-rc6.0712.intel_next.1.x86_64+server, HPCG from MKL_v2022.1.0. Power calculated using CPU TDP and 7W per DIMM for 31 nodes for equal performance

Numenta BERT-Large

- AMD Milan: Tested by Numenta as of 11/28/2022. 1-node, 2x AMD EPYC 7R13 on AWS m6a.48xlarge, 768 GB DDR4-3200, Ubuntu 20.04 Kernel 5.15, OpenVINO 2022.3, BERT-Large, Sequence Length 512, Batch Size 1
- Intel® Xeon® 8480+: Tested by Numenta as of 11/28/2022. 1-node, 2x Intel® Xeon® 8480+, 512 GB DDR5-4800, Ubuntu 22.04 Kernel 5.17, OpenVINO 2022.3, Numenta-Optimized BERT-Large, Sequence Length 512, Batch Size 1
- Intel® Xeon® Max 9468: Tested by Numenta as of 11/30/2022. 1-node, 2x Intel® Xeon® Max 9468, 128 GB HBM2e 3200 MT/s, Ubuntu 22.04 Kernel 5.15, OpenVINO 2022.3, Numenta-Optimized BERT-Large, Sequence Length 512, Batch Size 1

Resources and Configurations

Intel ® Quick Assist Technology

- **Up to 84% fewer cores to achieve same connections/s on NGINIX with built-in QAT vs. out-of-the-box software** - QAT Configuration HW/SW on 8490H: 1-node, pre-production platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel QuickAssist Accelerator (Intel QAT), Number of QAT device utilized=4(1 socket active), 1024GB (16x64 GB) total DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, QAT engine v0.6.14, QAT v20.l.0.9.1, NGINX 1.20.1, OpenSSL 1.1.1l, IPP crypto v2021_5, IPSec v1.1 , TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel September 2022.

  QAT Configuration SW on 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, QAT engine v0.6.14, NGINX 1.20.1, OpenSSL 1.1.1l, IPP crypto v2021_5, IPSec v1.1 , TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel October 2022.

  OOB Configurations:

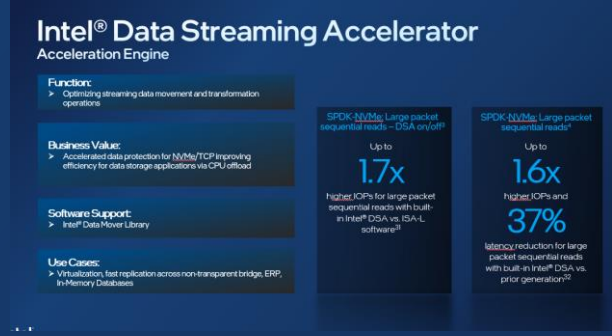
  8490H: 1-node, pre-production platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores), with 1024GB (16x64 GB) total DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, NGINX 1.20.1, OpenSSL 1.1.1l, TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel September 2022.

  8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 1x100GbE, NGINX 1.20.1, OpenSSL 1.1.1l, TLS 1.3 AES_128_GCM_SHA256, ECDHE-X25519-RSA2K, tested by Intel October 2022.

- **Up to 95% fewer cores and 2x higher level 1 compression throughput leveraging integrated QAT vs. prior generation** - 8490H: 1-node, pre-production platform with 2x 4th Gen Intel® Xeon Scalable Processor (60 core) with integrated Intel QuickAssist Accelerator (Intel QAT), QAT device utilized=8(2 sockets active), with Total 1024GB (16x64 GB) DDR5 memory, microcode 0xf000380, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, QAT v20.l.0.9.1 , QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel September 2022.

  8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo Off, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01,QAT v1.7.l.4.16, QATzip v1.0.9 , ISA-L v2.3.0, tested by Intel October 2022.

Resources and Configurations

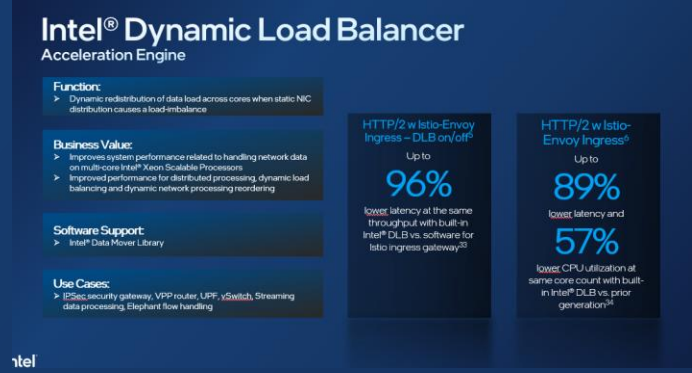


Intel® Data Streaming Accelerator

- **Up to 1.7xhigher IOPs for large packet sequential reads with built-in Intel® DSA vs. ISA-L software** - 8490H: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0xf000380, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel September 2022. 8380: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel October 2022.

- **Up to 1.6x higher IOPs and 37% latency reduction for large packet sequential reads with built-in Intel® DSA vs. prior generation** - 8490H: 1-node, 2x pre-production 4th Gen Intel Xeon Scalable processor (60 core) with integrated Intel Data Streaming Accelerator (Intel DSA), DSA device utilized=1(1 active socket), on pre-production Intel platform and software with 1024GB DDR5 memory (16x64 GB), microcode 0xf000380, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel® SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel® Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel September 2022. 8380: 1-node, 2x production 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, DDR4 memory total 1024GB (16x64 GB), microcode 0xd000375, HT On, Turbo On, SNC Off, Ubuntu 22.04.1 LTS, 5.15.0-47-generic, 1x 1.92TB Intel SSDSC2KG01, 4x 1.92TB Samsung PM1733, 1x Intel Ethernet Network Adapter E810-2CQDA2, 2x100GbE, FIO v3.30, SPDK 22.05, tested by Intel October 2022.

Resources and Configurations



Intel® Dynamic Load Balancer

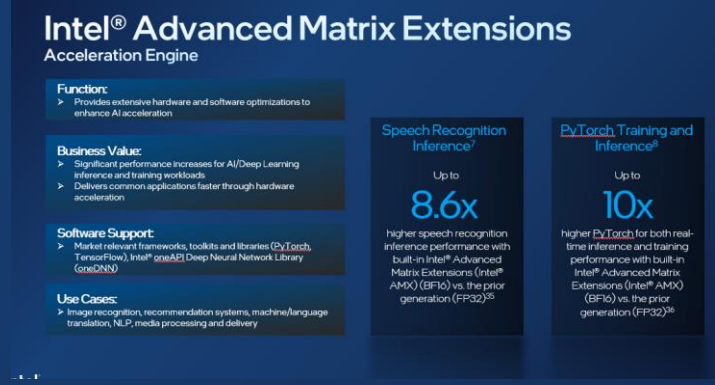
- **Up to 96%lower latency at the same throughput with built-in Intel® DLB vs. software for Istio ingress gateway** - 8480+: 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ with Intel DLB on Intel ArcherCity with GB (32 slots/ 32GB/ DDR5 4800) total memory, ucode 0x2b0000a1, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-40-generic, 1x 54.9G INTEL SSDPEK1A058GA, 4x Ethernet Controller E810-C for QSFP, 2x Ethernet Controller XXV710 for 25GbE SFP28, 1x Ethernet Controller I225-LM, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 10/27/2022.

  8360Y: 1-node, 2x Intel(R) Xeon(R) Platinum 8360Y on Intel M50CYP2SBSTD with GB (32 slots/ 16GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-50-generic, 2x 1.8T INTEL SSDPE2KX020T8, 1x Ethernet Controller E810-C for QSFP, 2x BCM57416 NetXtreme-E Dual-Media 10G RDMA Ethernet Controller, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 11/3/2022.

- **Up to 89% lower latency and 57% lower CPU utilization at same core count with built-in Intel® DLB vs. prior generation -** 8480+: 1-node, pre-production platform with 2x Intel(R) Xeon(R) Platinum 8480+ with Intel DLB on Intel ArcherCity with GB (32 slots/ 32GB/ DDR5 4800) total memory, ucode 0x2b0000a1, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-40-generic, 1x 54.9G INTEL SSDPEK1A058GA, 4x Ethernet Controller E810-C for QSFP, 2x Ethernet Controller XXV710 for 25GbE SFP28, 1x Ethernet Controller I225-LM, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 10/27/2022.

  8360Y: 1-node, 2x Intel(R) Xeon(R) Platinum 8360Y on Intel M50CYP2SBSTD with GB (32 slots/ 16GB/ DDR4 3200) total memory, ucode 0xd000375, HT on, Turbo off, Ubuntu 22.04 LTS, 5.15.0-50-generic, 2x 1.8T INTEL SSDPE2KX020T8, 1x Ethernet Controller E810-C for QSFP, 2x BCM57416 NetXtreme-E Dual-Media 10G RDMA Ethernet Controller, Nighthawk, gcc version 11.2.0, Docker 20.10.17, Kubernetes v1.22.3, Calico 3.21.4, Istio 1.13.4. DLB SW v 7.8, qatlib is 22.07.1, Nighthawk-worker run on 40 threads, 15-25 POD's with nighthawk-server and envoy sidecar proxy, 100Gb back-to-back connections between device, aRFS enabled – NIC interrupts pinned to the core with running applications, test by Intel on 11/3/2022.

Resources and Configurations

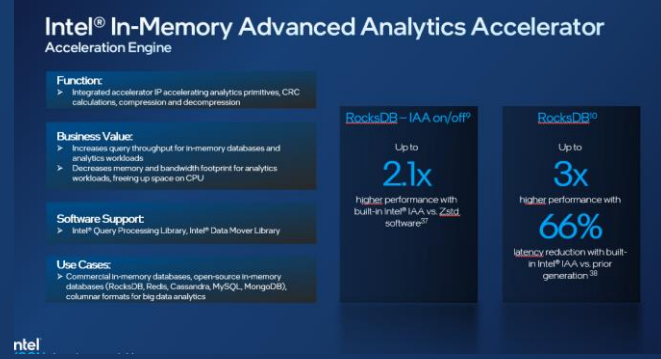


Intel® Advanced Matrix Extensions

- **Up to 8.6x higher speech recognition inference performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)** - 8480+: 1-node, pre-production platform with 2x Intel Xeon Platinum 8480+ on Archer City with 1024 GB (16 slots/ 64GB/ DDR5-4800) total memory, ucode 0x2b0000a1, HT on, Turbo on, CentOS Stream 8, 5.15.0, 1x INTEL SSDSC2KW256G8 (PT)/Samsung SSD 860 EVO 1TB (TF), Resnext101 32x16d, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, amx bf16=1,64, amx int8=1,116, ImageNet, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.; 8380: 1-node, 2x Intel Xeon Platinum 8380 on M50CYP2SBSTD with 1024 GB (16 slots/ 64GB/ DDR4-3200) total memory, ucode 0xd000375, HT on, Turbo on, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1x INTEL SSDSC2KG960G8, Resnext101 32x16d, Inference: bs=1 [4cores/instance], bs=n [1socket/instance], bs: fp32=1,64, int8=1,116, ImageNet, Framework: https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66; Modelzoo: https://github.com/IntelAI/models/tree/spr-launch-public, PT:1.13, IPEX: 1.13, OneDNN: v2.7, test by Intel on 10/24/2022.

- **Up to 10x higher PyTorch for both real-time inference and training performance with built-in Intel® Advanced Matrix Extensions (Intel® AMX) (BF16) vs. the prior generation (FP32)-**
  - 8. 3.5-10x & 4.9x: PyTorch geomean of ResNet50, Bert-Large, DLRM, MaskRCNN, SSD-ResNet34, RNN-T. 2.3-5.5x & 3.6x: TensorFlow geomean of ResNet50, Bert-Large, SSD-ResNet34, Transformer.
  - 5.7-10x & 7x: PyTorch geomean of ResNet50, Bert-Large, MaskRCNN, SSD-ResNet34, RNN-T, Resnext101.
  - 2.5-4.8x & 3.6x: PyTorch geomean of ResNet50, Bert-Large,SSD-ResNet34, Resnext101.
  - 1.8-9.6x & 4.6x: TensorFlow geomean of ResNet50, Bert-Large, SSD-ResNet34, Transformer, 3D Unet, DIEN.
  - 2.1-4.7x & 2.9x: TensorFlow geomean of ResNet50, Bert-Large, SSD-ResNet34, Transformer, 3D Unet.

Resources and Configurations



Intel® In-Memory Advanced Analytics Accelerator

- **Up to 2.1x higher performance with built-in Intel® IAA vs. Zstd software** -8490H: 1-node, pre-production Intel platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0xf000380, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel September 2022. 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel October 2022.
- **Up to 3x higher performance with 66% latency reduction with built-in Intel® IAA vs. prior generation** - 8490H: 1-node, pre-production Intel platform with 2x 4th Gen Intel Xeon Scalable Processor (60 cores) with integrated Intel In-Memory Analytics Accelerator (Intel IAA), HT On, Turbo On, Total Memory 1024GB (16x64GB DDR5 4800), microcode 0xf000380, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, QPL v0.1.21,accel-config-v3.4.6.4, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel September 2022. 8380: 1-node, 2x 3rd Gen Intel Xeon Scalable Processors(40 cores) on Coyote Pass platform, HT On, Turbo On, SNC Off, Total Memory 1024GB (16x64GB DDR4 3200), microcode 0xd000375, 1x 1.92TB INTEL SSDSC2KG01, Ubuntu 22.04.1 LTS, 5.18.12-051812-generic, ZSTD v1.5.2, RocksDB v6.4.6 (db_bench), tested by Intel October 2022.