

5 ways Intel® Xeon® Scalable processors with built-in accelerators can support your fastest-growing workloads

Workloads are evolving — and so is compute architecture. Traditionally, adding more cores to your CPU or choosing a higher-frequency CPU would improve workload performance and efficiency, but these techniques alone can no longer guarantee the same benefits they achieved in the past. Modern workloads place increased demands on compute, network and storage resources. In response, a growing trend exists to deploy power-efficient accelerators to offload specialized functions and reserve compute cores for general-purpose tasks. Offloading specialized tasks to AI, security, HPC, networking, analytics and storage accelerators can result in faster time to results and power savings.

As a result, Intel has integrated the broadest set of built-in accelerators in 4th Gen Intel® Xeon® Scalable processors to boost performance, reduce latency and increase power efficiency. **Below are five ways Intel Xeon Scalable processors with Intel® Accelerator Engines can help your business solve today's most rigorous workload challenges across cloud, networking and enterprise deployments.**

01

Run workloads faster on existing hardware

Instead of customizing systems for new applications, which adds complexity, enterprises can achieve the performance needed to meet a variety of deployments with a scalable platform. 4th Gen Intel Xeon Scalable processors have the most built-in accelerators of any CPU on the market to deliver performance and power efficiency advantages across the fastest-growing workloads.

While certain workload requirements warrant additional specialized hardware, in many cases these purpose-built accelerators will enable you to run your workloads effectively as is on your CPU. For instance, Intel® AI Engines on Intel Xeon Scalable processors are designed to run complex AI workloads — such as deep-learning training and inference — on the same CPU as other tasks. The most popular AI frameworks used by data scientists, such as TensorFlow and PyTorch, are already optimized for Intel Xeon Scalable processors.

Additionally, in the case of storage workloads, Intel® Volume Management Device (Intel® VMD) enables direct control and management of NVMe SSDs from the PCIe bus without the need for power-consuming RAID hardware adapters like host bus adapters (HBAs).



02

Improve performance across various workloads

For AI workloads, the new Intel® Advanced Matrix Extensions (Intel® AMX) significantly accelerates deep-learning capabilities on your CPU. Customers using 4th Gen Intel Xeon Scalable processors with Intel AMX (BF16) will gain up to 10x higher PyTorch performance for both real-time inference and training workloads compared to the previous generation (FP32).¹

Other new Intel Accelerator Engines speed up data movement and compression for faster networking, boost query throughput for more responsive data analytics, and offload scheduling and queue management to dynamically balance loads across multiple cores. For example, with Intel® In-Memory Advanced Analytics Accelerator (Intel® IAA), customers using 4th Gen Intel Xeon Platinum 8490H with Intel IAA will gain up to 3x higher RocksDB performance compared to the previous generation.²



Increase power efficiency

With built-in accelerators and software optimizations, Intel Xeon Scalable processors can help deliver leading performance per watt on targeted real-world workloads. This results in more efficient CPU utilization, lower electricity consumption and higher ROI, helping businesses achieve their sustainability goals.

4th Gen Intel Xeon Scalable processors have built-in accelerators to speed up encryption, data compression and data movement, promoting optimal CPU use and power resources. Intel® QuickAssist Technology (Intel® QAT), featured in our 4th Gen Xeon Scalable processors, boosts networking and storage performance by moving key functions off CPU cores to this newly built-in accelerator engine. Customers using 4th Gen Intel Xeon Platinum 8490H with Intel QAT for networking workloads will require up to 95% fewer CPU cores and gain 2x higher data compression throughput compared to the previous generation.³ Additionally, Intel® Data Streaming Accelerator (Intel® DSA) helps accelerate storage I/O per second and reduce latency. Customers using 4th Gen Intel Xeon Platinum 8490H with Intel DSA will gain up to 1.6x higher IOPS and up to 37% latency reduction for large packet sequential read compared to the previous generation.⁴

03

04

Enhance protection of your most sensitive data

Intel Xeon Scalable processors enable confidential computing solutions that enhance protection for your data on prem, at the edge and in the cloud. Intel® Software Guard Extensions (Intel® SGX) is the most researched, updated and deployed confidential computing technology in data centers on the market today.

Intel SGX helps protect sensitive data and application code while actively in use. This helps defend against breaches, leaks or attacks that could halt business operations, compromise critical data or break compliance. For customers that need the least amount of code to access confidential data, Intel SGX provides the smallest trust boundary of any confidential computing technology in the data center today.⁵

Intel® Trust Domain Extensions (Intel® TDX) is a new capability available through select cloud providers in 2023 that offers increased confidentiality at the virtual machine (VM) level, enhancing privacy and control over your data. Within an Intel TDX confidential VM, the guest OS and VM applications are isolated from access by the cloud host, hypervisor and other VMs on the platform.

Additionally, Intel® Crypto Acceleration — an instruction set in the Intel Xeon Scalable processor core architecture — uses single instruction, multiple data (SIMD) techniques to process more encryption operations in every clock cycle. This can help increase the total throughput of applications that require strong data encryption, with minimal impact on performance and user experience.



Gain more processing capacity for your most data-intensive workloads

The need for greater computing performance in business, science and academia has never been higher. Intel has helped organizations design system architectures that execute the most demanding of workloads. Whether these organizations are trying to crack the biggest challenges in medicine, economics or engineering, Intel® HPC Engines on Intel Xeon Scalable processors boost workload performance to conquer highly advanced computational tasks. Customers using 4th Gen Intel Xeon Scalable processors will gain up to 1.56x higher average HPC performance compared to the previous generation.⁶

For data-intensive workloads such as modeling and simulation, Intel Xeon Scalable processors not only allow code to take advantage of Intel® Advanced Vector Extensions 512 (Intel® AVX-512), but also offer high system memory capacity and bandwidth. This helps accelerate complex workloads on existing hardware.

05



Designed and optimized for specific use cases across market segments, Intel Accelerator Engines on Intel Xeon Scalable processors enable high-performance, hardware-based acceleration across today's fastest-growing workloads.

Current and future Intel Xeon Scalable processors will continue to boost performance and results with these powerful features, giving customers the tools they need to be faster, more sustainable and more successful than before.

Learn more about Intel Xeon Scalable processors by visiting

<https://www.intel.com/xeonscalable>.

¹See [A16] and [A17] at [intel.com/processorclaims](https://www.intel.com/processorclaims): 4th Gen Intel Xeon Scalable processors. Results may vary.

²See [D1] at [intel.com/processorclaims](https://www.intel.com/processorclaims): 4th Gen Intel Xeon Scalable processors. Results may vary.

³See [N16] at [intel.com/processorclaims](https://www.intel.com/processorclaims): 4th Gen Intel Xeon Scalable processors. Results may vary.

⁴See [N18] at [intel.com/processorclaims](https://www.intel.com/processorclaims): 4th Gen Intel Xeon Scalable processors. Results may vary.

⁵See <https://www.intel.com/content/www/us/en/architecture-and-technology/software-guard-extensions-enhanced-data-protection.html>

⁶See [H1] at [intel.com/processorclaims](https://www.intel.com/processorclaims): 4th Gen Intel® Xeon® Scalable processors. Results may vary.

Notices & 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.

For workloads and configurations, visit 4th Gen Xeon Scalable processors at www.intel.com/processorclaims. Results may vary.

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

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

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

Availability of accelerators varies depending on SKU. Visit the [Intel Product Specifications page](#) for additional product details.