

intel® XEON®

Elevate Performance
with 5th Gen Intel®
Xeon® Processors
Featuring Intel®
Accelerator Engines

E-Guide



Table of contents

What are built-in accelerators and why should you use them?	3
The real-world benefits of Intel® Accelerator Engines	4
Which Intel Accelerator Engines are right for your business?	5
Intel® AI Engines	6
Intel® Security Engines	7
Intel® HPC Engines	8
Intel® Network Engines	9
Intel® Analytics Engines	10
Intel® Storage Engines	11
Conclusion	12



What are built-in accelerators and why should you use them?

What if, instead of building customized workload solutions every time you need to establish new capabilities, you could rely on the technology already built into your CPU? With Intel Xeon Scalable processors, you can.

These trusted CPUs you already depend on to run all your workloads are supporting growing data center needs with Intel Accelerator Engines, which increase performance and performance per watt across some of the most demanding workloads.

In this guide, we'll explore how these purpose-built, integrated accelerators support workloads spanning AI, security, HPC, networking, analytics and storage.



The real-world benefits of Intel Accelerator Engines

Whether you're using Intel Xeon Scalable processors for your workloads on premises, in the cloud or at the edge, Intel Accelerator Engines can help your business reach new heights. These built-in accelerators are designed to provide a range of benefits, including **increased application performance, reduced costs and improved power efficiency**:

Performance	Intel Accelerator Engines are designed to help deliver higher performance on targeted workloads.
Cost savings	They allow you to improve performance without having to purchase additional specialized hardware.
Power savings	Additionally, Intel Accelerator Engines can help improve power efficiency by offloading common tasks from the CPU cores, boosting overall application performance while minimizing power usage to meet your sustainability goals.



Which Intel Accelerator Engines are right for your business?

With the most built-in accelerators of any CPU on the market, 4th Gen Intel® Xeon® Scalable processors are designed to improve performance and help protect data across today's fastest-growing workloads. To help you decide which Intel Accelerator Engines can best support your business, let's look at our top offerings across six key categories: AI, security, HPC, networking, analytics and storage.

Intel® AI Engines

Intel® Advanced Matrix Extensions (Intel® AMX)
Intel® Advanced Vector Extensions 512 (Intel® AVX-512)

Intel® Security Engines

Intel® Software Guard Extensions (Intel® SGX)
Intel® Trust Domain Extensions (Intel® TDX)
Intel® Crypto Acceleration

Intel® HPC Engines

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

Intel® Network Engines

Intel® QuickAssist Technology (Intel® QAT)
Intel® Dynamic Load Balancer (Intel® DLB)

Intel® Analytics Engines

Intel® In-Memory Analytics Accelerator (Intel® IAA)
Intel® Data Streaming Accelerator (Intel® DSA)
Intel® QuickAssist Technology (Intel® QAT)

Intel® Storage Engines

Intel® Data Streaming Accelerator (Intel® DSA)
Intel® QuickAssist Technology (Intel® QAT)
Intel® Volume Management Device (Intel® VMD)

Intel AI Engines

Intel Advanced Matrix Extensions (Intel AMX)

Intel AMX is Intel's solution for deep-learning training and inference on 5th Gen Intel Xeon Scalable processors. Ideal for workloads spanning the AI pipeline, such as natural language processing, recommendation systems for image recognition, GenAI and LLMs, this feature extends the built-in AI acceleration capabilities of previous Intel Xeon Scalable processors.

Up to

15X real-time object detection
inference performance
(SSD-ResNet34)

and

15X higher
performance/watt

on 5th Gen Intel Xeon with AMX INT8 compared to
3rd Gen Intel Xeon processors¹

Intel Advanced Vector Extensions 512 (Intel AVX-512)

A continuing feature of Intel Xeon Scalable processors, [Intel AVX-512](#) is a general-purpose performance-enhancing accelerator with a wide range of uses. When it comes to AI, Intel AVX-512 can accelerate classical machine learning training and inference workloads. 5th Gen Intel Xeon Scalable processors with Intel AVX-512 are also designed to speed up data preprocessing — and provide performance, sustainability and total cost of ownership (TCO) benefits.



Intel Security Engines

Intel Software Guard Extensions (Intel SGX)

Intel SGX provides application isolation and is designed to enhance data protection in use. This continuing feature of Intel Xeon Scalable processors provides the foundation for confidential computing solutions across edge and multi-cloud infrastructures.

Intel SGX is the most researched, updated and deployed confidential computing technology in data centers on the market at the moment. For customers that require 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.

Intel Trust Domain Extensions (Intel TDX)

Intel TDX offers isolation and confidentiality at the virtual machine (VM) level. 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. Intel TDX offers a simpler migration path for existing applications to adopt confidential computing and move to a trusted execution environment (TEE).

5th Gen Intel Xeon Scalable processor with Intel TDX offers enhanced VM security with small, ~3% performance overhead on key workloads.²

Intel Crypto Acceleration

Intel Crypto Acceleration 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.



Intel HPC Engines

Intel Advanced Vector Extensions 512 (Intel AVX-512)

With ultrawide 512-bit vector operations capabilities, Intel AVX-512 is especially suited to handle the most demanding computational tasks commonly encountered in HPC applications. It's used by organizations across educational, financial, enterprise, engineering and medical industries. By enabling users to run complex workloads on existing hardware, Intel AVX-512 accelerates performance for tasks like financial analytics, 3D modeling and scientific simulations.



Intel Advanced Matrix Extensions (Intel AMX)

Intel AMX is also designed to deliver performance gains across popular HPC workloads. This built-in accelerator transforms large matrix math calculations into a single operation and uses a two-dimensional register file to store large chunks of data.

Intel Data Streaming Accelerator (Intel DSA)

Intel DSA is a feature designed to optimize and speed up streaming data movement and transformation operations common in networking, data-processing-intensive applications and high-performance storage. Intel DSA accelerates HPC workloads by offloading the most common data movement tasks that cause CPU overhead in data-center-scale deployments.

Up to

1.31X higher LAMMPS performance

on 5th Gen Intel Xeon Scalable platform vs. prior gen³

Intel Network Engines

Intel QuickAssist Technology (Intel QAT)

Intel QAT boosts performance to meet the demands of today's networking workloads, helping systems serve more clients. It can deliver significant workload acceleration for cryptography, including symmetric and asymmetric encryption and decryption. It also allows users to accelerate data-compression workloads on existing hardware and increases power management efficiency.

Intel QAT using RSA4K can increase client density on an open-source NGINX web server compared to software running on CPU cores without acceleration.

Intel Dynamic Load Balancer (Intel DLB)

Intel DLB is a feature that enables the efficient distribution of network processing across multiple CPU cores. It also restores the order of networking data packets processed simultaneously on CPU cores.

With Intel DLB, customers can gain higher performance on packet forwarding compared to software queue management on cores without acceleration. Additionally, applications can achieve higher packet processing performance than the previous generation.

Up to

1.73X higher NGINX TLS handshake performance

with 5th Gen Intel Xeon vs. 3rd Gen Intel Xeon processors⁴



Intel Analytics Engines

Intel In-Memory Analytics Accelerator (Intel IAA)

Intel IAA is designed to accelerate database and analytics performance while being more energy efficient. By increasing query throughput and decreasing the memory footprint for in-memory databases and advanced analytics workloads, this feature can provide faster data movement and improve CPU core utilization by reducing dependency on CPU cores.

Intel IAA is ideal for in-memory databases, open-source databases and data stores like RocksDB, Redis, Cassandra, ClickHouse and MySQL. With Intel IAA, customers can gain higher throughput for data decompression when running the open-source RocksDB database engine compared to using software compression on CPU cores without acceleration. RocksDB achieves up to 3.7x higher performance with 5th Gen Intel Xeon Platinum 8592+ using integrated Intel IAA versus 3rd Gen Intel Xeon Platinum 8380 using Zstd.⁵ Customers can also experience Clickhouse gains up to 2.49x in performance using integrated Intel IAA versus 3rd Gen Intel Xeon processors using Zstd.⁶

Intel Data Streaming Accelerator (Intel DSA)

Intel DSA is designed to optimize and speed up streaming data movement and transformation operations common in data-processing-intensive applications, driving better business outcomes. By offloading tasks like data movement, data copying and error checking, Intel DSA enables the CPU to focus on business-critical database functions or other analytics workloads. This reduces query latencies and increases throughput, delivering faster data processing.

5th Gen Intel Xeon delivers up to 2.85x higher IOPs and up to 65% latency reduction for large packet sequential read with 5th Gen Intel Xeon Platinum 8592+ using integrated Intel DSA versus 3rd Gen Intel Xeon.⁷

Intel QuickAssist Technology (Intel QAT)

Intel QAT is an established technology previously implemented in networking systems on a chip, in Intel Xeon Scalable processor chipsets, and as a standalone accelerator device and PCIe adapter card. Intel QAT accelerates backups across many databases. When enabled for Microsoft SQL, the performance was even faster and more efficient. Customers can achieve up to 2.56x higher MS SQL+Backup performance with 5th Gen Intel Xeon Platinum 8592+ using integrated Intel QAT versus 3rd Gen Intel Xeon Platinum 8380 using default compression.⁸

The benefits of Intel QAT go beyond immediate performance gains. Enterprises can, for example, reduce core count, leading to more streamlined processor utilization. Intel QAT also empowers businesses to realize new efficiencies and scalability, particularly for large-scale operations such as cloud environments, data centers, extensive data lakes and various storage tiers.

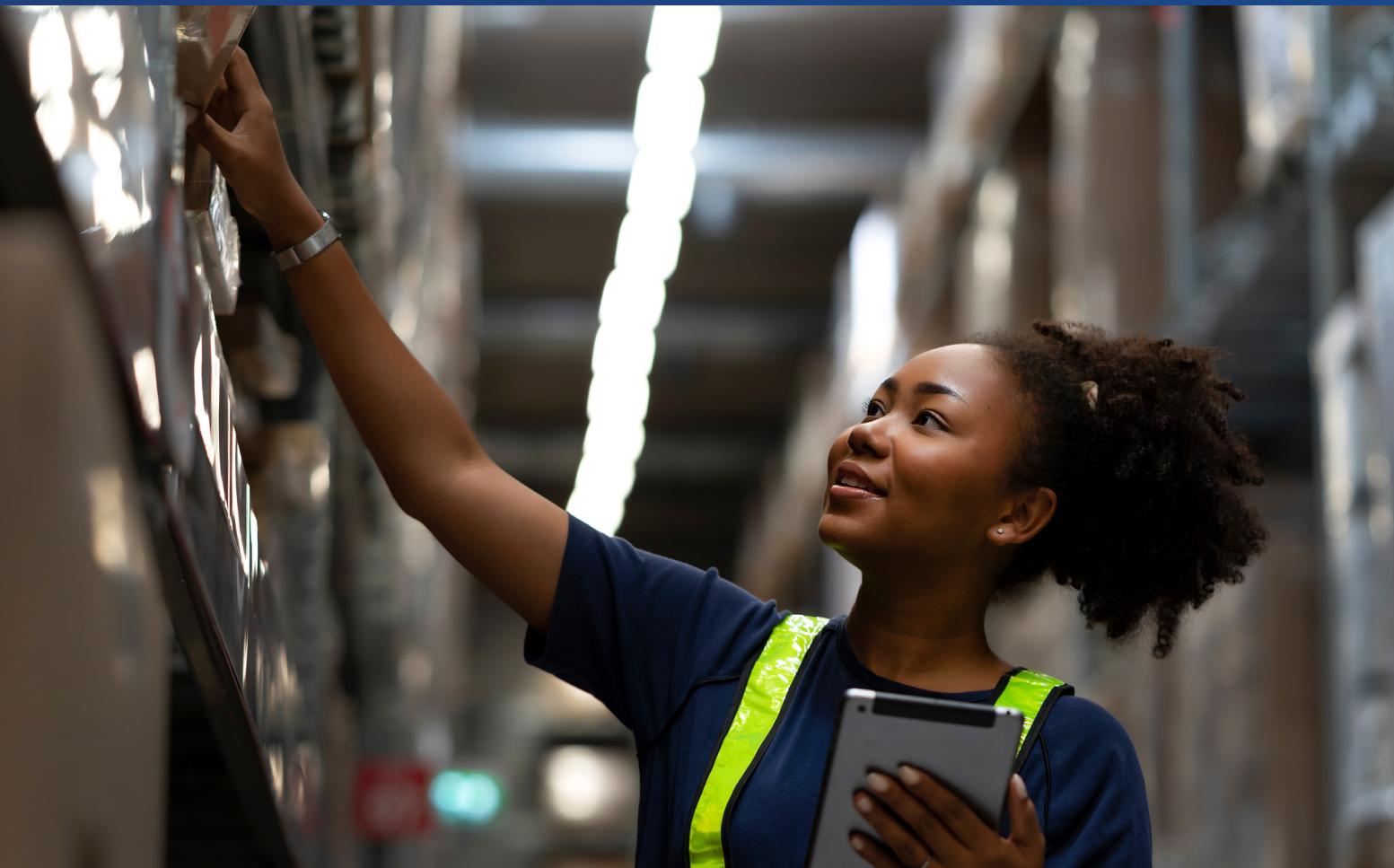
Intel Storage Engines

Intel Data Streaming Accelerator (Intel DSA)

The Intel Data Streaming Accelerator is Intel's next-generation direct memory access (DMA) engine. It speeds transfers between volatile and persistent memory and supports virtualized memory and I/Os.

Intel DSA works on the CPU — between DRAM, caches and processor cores — and extends across I/Os to attached memory and storage, as well as networked resources.

Customers can achieve up to 2.45x higher IOPs and up to 59% latency reduction for small packet random read with 5th Gen Intel Xeon Platinum 8592+ using integrated Intel DSA versus 3rd Gen Intel Xeon Platinum 8380.⁹



Intel QuickAssist Technology (Intel QAT)

Intel QAT increases performance of storage workloads and applications by accelerating in-line encryption and data compression/decompression. For instance, using Intel QAT as an offload engine provides a significant throughput improvement on compression compared to running the same compression algorithm on CPU cores without acceleration.

Customers can achieve up to 2.17x higher NGINX TLS handshake performance per core with 5th Gen Intel Xeon Platinum 8592+ with integrated QAT versus OOB software.¹⁰

Intel Volume Management Device (Intel VMD)

Intel VMD is a feature that enables direct control and management of NVMe SSDs from the PCIe bus without the need for additional hardware adapters. It allows for a smoother and lower-cost transition to fast NVMe storage while limiting the downtime of critical infrastructure. Plus, Intel VMD provides benefits like bootable RAID, robust surprise hot plug and blink status LED. It increases serviceability and enables you to deploy next-generation storage with confidence too.

Intel® VMD also enables Intel® Virtual RAID on CPU (Intel® VROC), a virtual RAID solution that eliminates the complexity, cost and power consumption of traditional hardware RAID host bus adapter (HBA) cards placed between the drives and the CPU. Intel® VROC requires additional software.

Learn more about what Intel Accelerator Engines on Intel Xeon Scalable processors can accomplish by visiting the [Intel Products page](#).

Conclusion

There are numerous workloads for which Intel Accelerator Engines will deliver greater business value to customers. Whether you're looking to increase performance, support sustainability initiatives or help ensure the protection of your most sensitive data, Intel Xeon processors with Intel Accelerator Engines offer a wide range of solutions without the need for additional hardware.

With the latest Intel Xeon Scalable processors, Intel redefines what customers can expect from a CPU — and delivers leading performance with the most built-in accelerators. Moreover, accelerators can be upgraded via [Intel On Demand](#), Intel's API-enabled business offerings that facilitate flexible consumption and configurations.

[**Learn more about Intel Xeon Scalable processors.**](#)

[**Learn more about Intel Developer Zone, the official hub for developing on Intel hardware and software.**](#)

[**Learn more about enabling critical workloads for 5th Gen Intel Xeon Processors and Intel Accelerator Engines.**](#)



¹See [A21] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

²See [S1] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

³See [H14] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁴See [N15] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁵See [D1] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁶See [D2] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁷See [N16] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁸See [D5] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

⁹See [N16] at [intel.com/processorclaims](#): 5th Gen Intel Xeon Scalable processors. Results may vary.

¹⁰See [N15] at [intel.com/processorclaims](#): 5th 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 5th 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.