Product Brief

Network-Optimized 4th Gen Intel® Xeon® Scalable Processors

Drive High Throughput and Low Latency with Next-Generation Network-Optimized Processors

Network-optimized 4th Gen Intel® Xeon® Scalable processors drive improvements in performance, energy efficiency and return on investment (ROI). The processor architecture includes the industry's largest array of built-in hardware accelerators to enhance the full range of network applications. These include 5G core and virtualized Radio Access Networks (vRANs), secure access service edge (SASE) deployments, content delivery networks (CDNs), IPSec Transport Layer Security (TLS) and virtual Cable Modem Termination System (vCMTS) workloads.

Communication service providers (CoSPs) and other network operators face continued pressure to reduce power consumption, to maximize cost efficiency and environmental sustainability. A key strategy for meeting these requirements has been to implement increasingly sophisticated traffic processing, based on network functions distributed from edge to core to cloud.

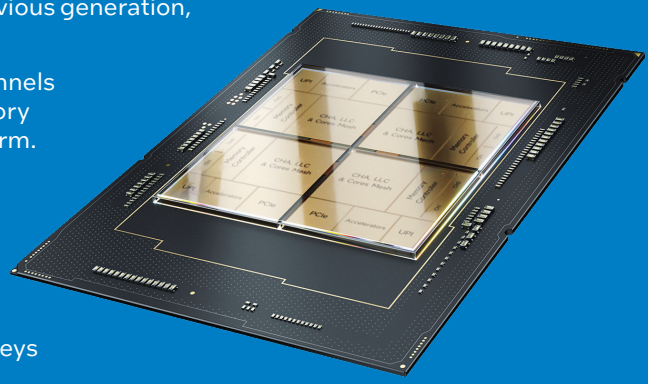
4th Generation Intel® Xeon® Scalable processors advance that vision, with industry-leading per-core performance and energy efficiency, with the industry's widest array of built-in accelerators, for functions from security to AI. Instead of building custom systems using expensive external accelerators, network operators have a fast path to deploying new services on a standardized, scalable platform that meets the performance needs of a wide variety of deployments — both today and in the future.

Network-optimized 4th Generation Intel Xeon Scalable processors (N-SKUs) are designed for a range of broadly deployed network and 5G workloads. This platform supports network functions and solutions with low latency, high throughput, deterministic performance and extended supply life. They are oriented toward fast packet and signal processing from a balanced platform that incorporates innovations and advances across execution, memory and I/O. The N-SKUs provide multiple operating modes for networking, server and compute-optimized workloads, enabling CoSPs to deploy a single server type with flexibility across use cases, simplifying acquisition and logistics.

vRAN	5G Core UPF	CDN	VCMTS	SASE
UP TO	UP TO	UP TO	UP TO	UP TO
2X	1.88X	1.46X	2X	1.71X
Higher throughput capacity with new Intel AVX instuctions[1]	Higher throughput with user plane functions in 5G wireless core network[2]	More live linear connections[3]	Better throughput[4]	Higher throughput for next gen firewall[5]

Purpose Built for Network Workloads

- **Advanced execution resources** in a range of core counts and feature sets, with improved per-core performance compared to the previous generation, enhanced by the most built-in accelerators in the industry.
- **Enhanced memory subsystem** with up to eight DDR5 channels operating at up to 4800 MT/s, a 1.5x improvement in memory bandwidth and speed compared to the predecessor platform.
- **Faster, higher capacity I/O** based on up to 80 lanes of PCIe 5.0 per socket, compared to 64 lanes of PCIe 4.0 per socket in the prior generation.
- **Data protected while in use** by confidential computing based on Intel® Software Guard Extensions (Intel® SGX), which isolates secrets such as passwords and encryption keys in private memory enclaves.



The multiple operating modes, enabled by Intel® Speed Select Technology (Intel® SST), make it possible for operators to better tailor execution to specific sets of requirements. Network Mode is performance-optimized for high throughput and low latency, using high processor frequencies for data plane, control plane and signal processing workloads. Server Mode reduces the core frequency while still operating in the same power envelope, for general-purpose processing. IT Mode is intended for IT workloads with lower latency requirements. For example, during off-peak hours, systems in IT Mode could perform back-office functions such as Operations and Business Support Systems (OSS/BSS).

Having multiple operating modes available enhances software reuse across the environment, with reduced platform validation requirements for network operators.

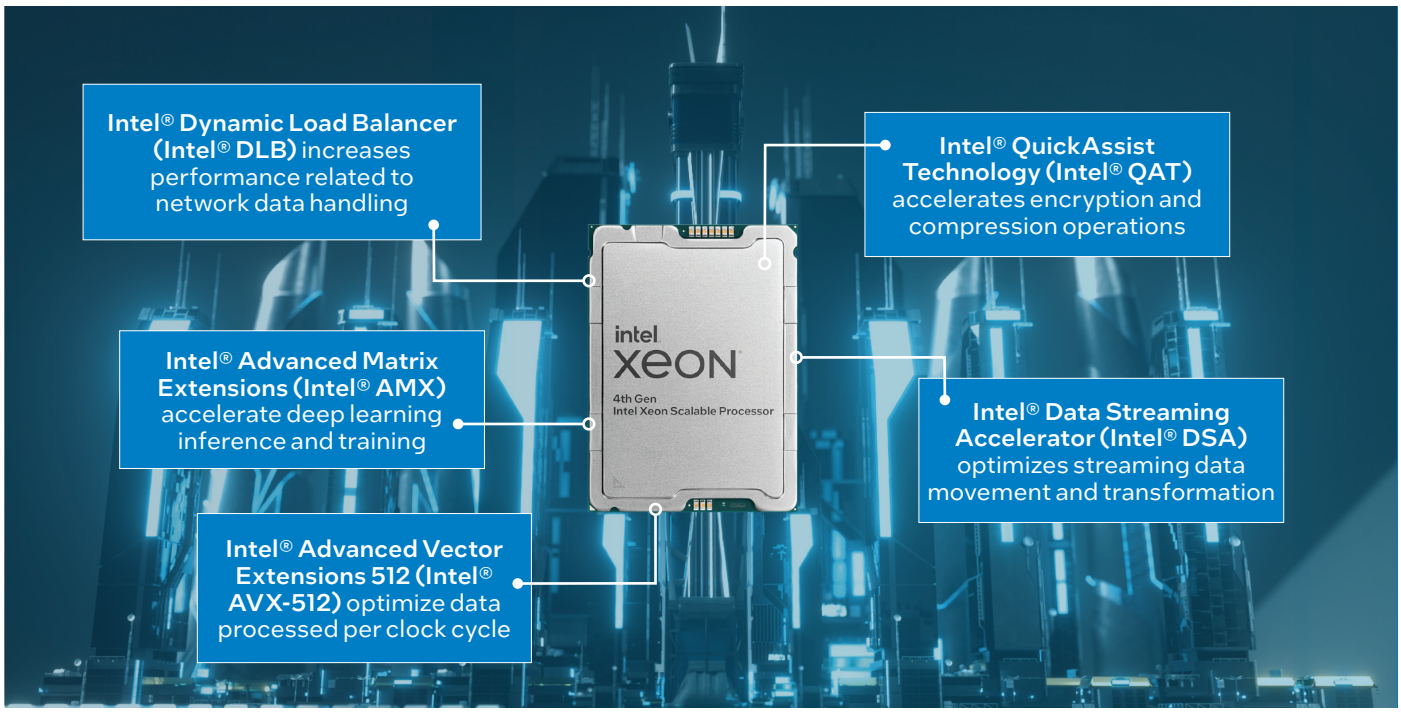
Network Mode	Server Mode	IT Mode
High-Performance Network Workloads	General-Purpose Server (Compute Processing	IT Workloads and Back-Office Systems

Network-optimized 4th Generation Intel Xeon Scalable processors ship in Medium Core Count (MCC) and Extreme Core Count (XCC) configurations. The MCC processors range from 24-32 cores per socket at 165-205 watts, while the XCC processors provide 52 cores per socket at 300 watts.

Accelerators built into both MCC and XCC processors speed up tasks such as data movement and compression for faster networking, boost query throughput for more responsive analytics, and offload scheduling and queue management to dynamically balance loads across multiple cores. To ensure usability of accelerators for developers, Intel supports the ecosystem with open source and proprietary software enablement, from applications to low-level elements such as drivers, libraries and APIs.

Built-in hardware accelerators for distributed network workloads

4th Gen Intel Xeon Scalable processors usher in a new paradigm based on an unprecedented array of hardware accelerators built directly into the CPU. This advance lets network operators tailor execution on general-purpose systems more than ever before, to optimize performance, power efficiency and security. That ability delivers more efficient CPU utilization, greater sustainability, lower electricity consumption and higher ROI.



Because these accelerators are built into the processor, their increased throughput and reduced latency are available to network operators without requiring external hardware. This architecture provides a significant energy efficiency advantage by eliminating the need to power additional devices or pass data over the PCIe bus between the cores and accelerator hardware.

Additionally, general-purpose hardware resources can be applied to other work during non-peak times, such as training AI and machine learning models or running analytics. The proximity of the accelerators to the processor cores also eliminates PCIe-bus transit time, improving latency for sensitive network workloads.

Intel invests substantially in software enablement that streamlines the implementation of built-in accelerators and other Intel architecture features and capabilities for developers. Industry-wide collaborations with solution providers of all types and sizes brings highly enabled solutions to market. Intel contributes code to the open source community that optimizes popular projects for the latest Intel architecture.

Intel libraries and oneAPI software development tools implement hardware optimizations to code that can execute across Intel hardware platforms, including CPUs, GPUs, FPGAs and other accelerators. Software-enablement activities lower the barriers and cost for software providers to adopt the latest acceleration features of Intel platforms, including 4th Gen Intel Xeon Scalable processors.

Intel® Advanced Matrix Extensions (Intel® AMX): Accelerated deep learning

Machine learning is proving effective at tuning network workloads to be more efficient and effective, particularly in SASE deployments and vRAN. Intel AMX is a built-in hardware accelerator that provides a significant leap in the performance of inference and training by speeding up the tensor processing at the heart of deep learning algorithms.

The technology includes TILEs, a set of up to eight expandable 2D register tiles per core that store larger chunks of data than predecessors, as well as TMUL (Tile Matrix Multiply), a set of matrix multiplication instructions that are the first operators on TILEs. Intel AMX accelerates time to value by enabling deep learning software to complete more inference in a given period of time or to close in on insights more quickly.

Intel® Advanced Vector Extensions 512 (Intel® AVX-512): The latest x86 vector instruction set

Progressively more sophisticated vectorization has contributed to faster calculations on larger data sets over many technology generations, including in network workloads such as SASE and vRAN. Intel® AVX-512, the latest x86 vector instruction set, builds on the vector processing power of its predecessors to accelerate the completion of data-intensive workloads.

Network functions can pack 32 double-precision and 64 single-precision floating point operations per clock cycle within the 512-bit vectors, as well as eight 64-bit and 16 32-bit integers with two 512-bit fused multiply-add (FMA) units for the most demanding computational workloads. The technology doubles the width of data registers, the number of registers and the width of FMA units compared to Intel® Advanced Vector Extensions 2 (Intel AVX2).

Intel® Dynamic Load Balancer (Intel® DLB): Optimized network data handling across cores

As network traffic management becomes more sophisticated, large workloads such as packet processing must be distributed across many cores, exemplified by a CDN supporting many simultaneous video streaming connections. Intel DLB improves overall system performance by optimizing data handling across processor cores. The accelerator dynamically redistributes work as the system load varies and restores the order of networking data packets processed simultaneously on separate cores. Intel DLB optimizes core resource usage, improving traffic handling capacity per node and system ROI.

View the latest performance data at www.intel.com/PerformanceIndex

Intel® Data Streaming Accelerator (Intel® DSA): Enhanced streaming data movement

Data movement and transformation operations are critical to the performance of streaming network workloads, including for CDNs. Intel DSA drives up streaming performance by offloading the most common data movement tasks that cause overhead in large-scale deployments. By shouldering almost all data movement operations, including checksum, memory compare and checkpointing, Intel relieves the CPU cores of overhead associated with moving data in and out of memory, storage and networking subsystems. Intel DSA optimizes the handling of streaming data across the CPU, memory and caches, as well as all attached memory, storage and network devices.

Intel® QuickAssist Technology (Intel® QAT): Accelerated encryption and compression

Reducing the overhead associated with encryption and data compression can play a significant role in improving network performance, including for SASE, CDN and 5G UPF workloads. Intel QAT is now built in as a hardware accelerator in 4th Gen Intel Xeon Scalable processors that enables faster data encrypt and decrypt on the fly, as well as more efficient data compression.

This latest version of the technology accelerates cryptographic ciphers, secure hashes, public key encryption and compression/decompression performance relative to prior generations. By offloading these tasks from the processor cores, Intel QAT frees up resources for other work, increasing overall throughput. Intel QAT contributes to zero-trust security strategies that protect data at all stages in any infrastructure — at rest, in flight and in use — without loss of performance for critical workloads.

Ecosystem support through Intel® Network Builders

The Intel Network Builders Program helps partners innovate and adapt to evolving business, technology and end-user needs, effectively and cost-efficiently. The program provides members with a variety of technical enablement options such as hands-on support from subject matter experts, access to virtual testing and optimization labs, training, tools and other resources.

Hardware enablement for evolving networking requirements

With high throughput and low latency provided by improvements across the balanced platform — including an industry-leading set of built-in hardware accelerators — network-optimized 4th Gen Intel Xeon Scalable processors help equip network operators for expanding requirements. The platform delivers value across key networking use cases, including the following:

- **SASE implementations** benefit from network, cryptography and AI processing capabilities for SASE edge POP deployments and SASE hosted on public cloud.
- **High-density CDNs** support more sessions per node with enhanced performance that includes accelerated load balancing, media streaming and cryptography acceleration, for lower TCO.
- **5G core (UPF) deployments** optimize system performance with efficient distribution of network processing across multiple CPU cores as the system load varies, improving system efficiency. Up to 30% throughput improvement over previous-generation 3rd Gen Intel Xeon Scalable processors.[2]
- **5G vRAN implementations** provide Distributed Unit and Centralized Unit optimizations for Massive MIMO pipelines with high energy efficiency to help reduce operating costs and improve sustainability. Up to 2X the throughput of previous-generation 3rd Gen Intel Xeon Scalable processors, in the same power envelope.[6]

The performance and energy-efficiency gains provided by network-optimized 4th Gen Intel Xeon Scalable processors support ongoing business benefits. High capacity per node helps reduce equipment CapEx, while high performance per watt reduces OpEx and helps companies meet sustainability goals. The platform provides the foundation for networking innovation in an increasingly data-intensive future.

Network-Optimized 4th Gen Intel® Xeon® Scalable Processors

SKU[§]	Processor Cores	Base Frequency (GHz)	All-Core Turbo (GHz)	Max Turbo (GHz)	Cache (MB)	TDP (Watts)	Maximum Scalability	DDR5 Memory Speed	Intel® UPI Links Enabled	Default Intel® DSA Devices	Default Intel® QAT Devices	Default Intel® DLB Devices	Default Intel® IAA Devices	Intel® SGX Enclave Capacity (GB, per Processor)	Long-Life Availability[§]	Intel® On Demand Capable
Intel® Xeon® Platinum 8470N processor	52	1.7	2.7	3.6	97.5	300	2S	4800	3	4	4	4	0	128	Yes	Yes
Intel® Xeon® Platinum 8471N processor	52	1.8	2.8	3.6	97.5	300	1S	4800	0	4	4	4	0	128	Yes	Yes
Intel® Xeon® Gold 6438N processor	32	2.0	2.7	3.6	60	205	2S	4800	3	1	2	2	0	128	Yes	Yes
Intel® Xeon® Gold 6428N processor	32	1.8	2.5	3.8	60	185	2S	4000	3	1	2	2	0	128	Yes	Yes
Intel® Xeon® Gold 6421N processor	32	1.8	2.6	3.6	60	185	1S	4400	0	1	0	0	0	128	Yes	Yes
Intel® Xeon® Gold 5418N processor	24	1.8	2.6	3.8	45	165	2S	4000	3	1	2	2	0	128	Yes	Yes
Intel® Xeon® Gold 5411N processor	24	1.9	2.8	3.9	45	165	1S	4400	0	1	2	2	0	128	Yes	Yes

[§] Intel does not commit or guarantee product availability or software support by way of road map guidance. Intel reserves the right to change road maps or discontinue products, software, and software support services through standard EOL/PDN processes. Contact your Intel account rep for additional information.

For more information about Intel® On Demand, visit intel.com/ondemand.

Learn More

www.intel.com/4thgenxeon-network | networkbuilders.intel.com

[1] See [N10] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.
[2] See [N8] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.
[3] See [N60] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.
[4] See [N4] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.
[5] See [N24] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.
[6] See [N10] at intel.com/processorclaims: 4th Gen Intel Xeon Scalable processors. Results may vary.

Availability of accelerators varies depending on SKU. Visit the Intel Product Specifications page for additional product details.

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

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for configuration details. No product or component can be absolutely secure.

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

Your costs and results may vary.

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

0323/FS/MESH/350502-001US