Solution Brief

Network-Optimized 5th Gen Intel® Xeon® Scalable Processors Content Delivery Network (CDN)

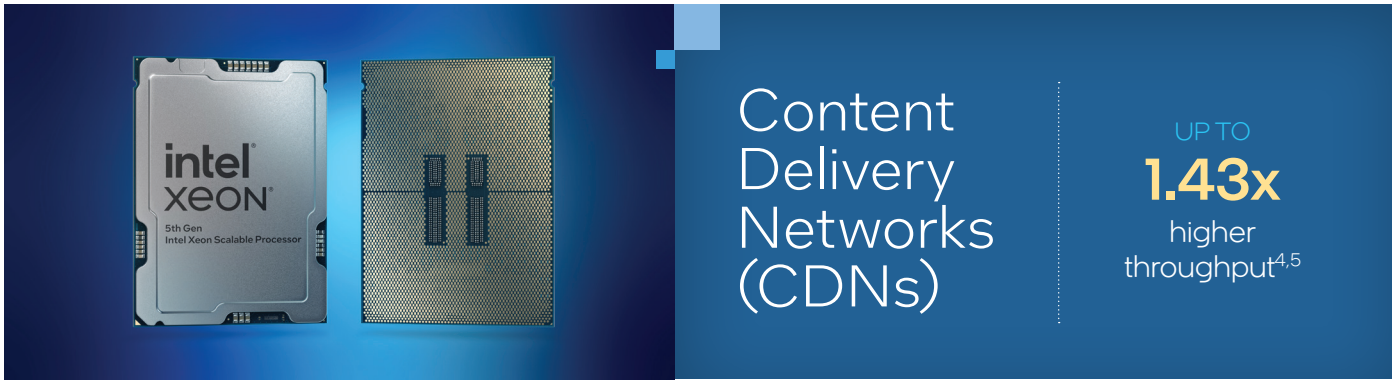
Platform and Software Innovation Improves CDN Performance and Power Efficiency

Network-optimized 5th Gen Intel® Xeon® Scalable processors help Content Delivery Networks (CDNs) break through existing limits on the number of subscribers they can support per server while also reducing energy requirements. Platform improvements coupled with optimized software enable efficient memory access and TLS encryption for CDN workloads. This lays the foundation for high-throughput CDNs that can cost-effectively scale for surging capacity requirements.

The expanding demand for high-bandwidth media content creates new urgency in longstanding challenges for operators to expand CDN capacity while improving cost efficiency. In service of that aim, they must innovate toward improving metrics such as sessions per node and energy requirements per stream or per Gbps. A leaner, more efficient infrastructure is instrumental to meet growth requirements while maintaining relatively steady CapEx and OpEx. Size of this market segment is projected to grow at a CAGR of 16.3% to reach a value of $36.51 billion by 2028.[1]

As CDNs move physically closer to the end users they serve, operators are adding low-latency compute infrastructure for the evolution of edge compute beyond core CDN workloads. Server platforms that can meet those low-latency requirements are vital to this evolution. 5th Gen Intel® Xeon® Scalable processors provide higher per-core performance, faster memory and improved performance per watt vs. predecessors, offering significant value and efficiency advantages. Network-optimized SKUs (with "N" appended to the SKU number) are designed specifically for distributed data plane and signaling workloads. To provide a balanced platform, 5th Gen Intel Xeon Scalable processors incorporate advances across execution, memory and I/O:

- **High-throughput, high-efficiency execution resources**. Improved per-core performance and up to 64 cores per socket, with the industry's most built-in accelerators and energy savings with Optimized Power Mode.
- **Enhanced memory subsystem**. Up to 16% increased DDR5 memory speed and up to 3x total last level cache[2,3] compared to its predecessor, to hold more user plane data close to the processor for enhanced throughput.
- **Expanded I/O speed and capacity**. Up to 80 lanes of PCIe Gen 5.0 per socket, with Intel Ultra Path Interconnect (Intel UPI) 2.0 speeds up to 20 GT/s and support for Compute Express Link (CXL) Types 1, 2 and 3.



With these increased hardware capabilities, software stack selection becomes more important to extract the best performance and efficiency from the hardware. Intel works with an ecosystem of software partners to help CDN operators improve performance from Intel processor-based platforms. With well-optimized software, Intel-based platforms can deliver excellent performance for CDN workloads without requiring the added cost of hardware offloads. Network-optimized 5th Gen Intel Xeon Scalable processors have been tested with CDN workloads from market leaders Broadpeak and Varnish and up to 1.43x higher CDN live-linear throughput[4,5] was observed in two separate benchmark efforts at Intel and at Broadpeak, compared to 3rd Gen. For complex workloads such as CDN applications that involve significant data movement to deliver bits over the network and significant compute to TLS encrypt those bits, tools including Intel® VTune™ Profiler and Intel Performance Counter Monitor can help identify areas of optimization for the entire workload, rather than simply replacing one algorithm with another.

Efficient software to reduce memory bandwidth amplification

The CDN workload involves a lot of data movement to deliver content over the network to end users. It is important to avoid unnecessary memory accesses, as memory bandwidth is a finite resource and every memory access has non-zero latency. The networking overhead for a CDN node to deliver 1 Tbps of content may be as much as 125 GB/s. Similarly, that same node will need to read the data from memory (125 GB/s) and then write it back to memory (125 GB/s) for TLS bulk encryption. If the data is read from storage rather than memory, there can be another 125 GB/s of memory bandwidth. Software can be optimized to avoid unnecessary memory accesses and to improve the utilization of the processor caches, helping to keep the memory bandwidth requirement under control.

Extracting value from optimized solutions based on Intel Xeon Scalable processors

HTTPS CDN delivery involves large numbers of copy-and-forward data operations. Requests come in from a network interface to the CPU, which fetches content from storage, copies it to DRAM, encrypts it, copies it back to DRAM and passes it to an outgoing NIC port. In aggregate, optimizing these operations can improve the memory and CPU efficiency of a software stack. With that improved efficiency, the solution can consume less energy for the same number of users, support more users with the same power consumption, or those resources could be allocated to performing an additional edge workload. Additionally, processor- and software-based solutions do not require the use of additional dedicated hardware offloads, which can increase solution cost and complexity.

Reduced performance overhead from encryption operations

Delivering high-bandwidth content over HTTPS requires TLS encryption at scale, with deterministic high throughput to safeguard quality of experience. Intel has worked with ISVs to improve the efficiency of their TLS operations to reduce CPU and memory requirements without requiring the use of added-cost hardware offloads. These resources can then be used for other important edge functions, adding to the efficiency of CDN platforms. The latest generation of Intel Xeon Scalable processors includes Intel Crypto Acceleration, which builds upon Intel AVX-512 and Intel AES-NI to improve bulk cryptography as well as public-key encryption, both of which are critical to scalability of CDN nodes delivering content over HTTPS.

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

Conclusion

5th Gen Intel Xeon Scalable processors offer a significant step forward in the cost-effective delivery of CDN services, with higher densities of simultaneous subscriber connections per CDN node and higher performance per watt. The balanced platform provides core count, memory bandwidth and I/O expandability improvements as the foundation for optimized software stacks. The solutions improve performance and efficiency of critical aspects of the CDN workload while freeing up compute resources and reducing energy usage.

Learn More

www.intel.com/5thgenxeon-network

networkbuilders.intel.com

[1] Mordor Intelligence, "Content Delivery Network Market Size & Share Analysis - Growth Trends & Forecasts (2023 - 2028)." https://www.mordorintelligence.com/industry-reports/content-delivery-market.

[2] On select SKUs.

[3] Results have been estimated or simulated comparing theoretical performance of DDR5-4800 to DDR5-5600. Performance varies by use, configuration and other factors.

[4] Performance varies by use, configuration and other factors. Actual results may vary.

Up to 1.43x higher CDN live-linear throughput with Broadpeak BkS450 on Dual Socket Xeon 6538N vs. 3rd Generation.

2S 6538N: Test by Broadpeak as of 11/20/2023. 1 node, Lenovo SR650 V3 with 2x Intel® Xeon Gold 6538N Processor, 32 cores, Hyperthreading on, Turbo on, total memory 512GB (16 slots, 32GB, 4800 MT/s), 1x Broadcom N1200G SN37B02874, 1x Mellanox ConnectX-6 Dx OF6FXM, 3x Mellanox ConnectX-7 MCX755106AC-HEAT, BIOS ESE122O-3.10 released 10/23/2023 (ucode 0x21000170), Arch Linux, kernel 6.6.1-arch1-1, OpenSSL 3.1.4 24 Oct 2023 (library: OpenSSL 3.1.4 24 Oct 2023), Broadpeak BkS450 version 6f8bc8, gcc-12 (Ubuntu 12.1.0-2ubuntu1~22.04) 12.1.0, Test Tool version 6d70f0 (keep-alive, 8192 total connections). Throughput measured with 100% Transport Layer Security (TLS) traffic with 100% target cache hit ratio.

2S 5318Y: Test by Broadpeak as of 11/20/2023. 1 node, Lenovo SR650 V2 with 2x Intel® Xeon Gold 5318Y Processor, 24 cores, Hyperthreading on, Turbo on, total memory 512GB (16 slots, 32GB, 3200 MT/s @ 2933 MT/s), 1x Mellanox ConnectX-6 Dx MCX623436AC-CDAB, 3x Mellanox ConnectX-6 Dx SN37A28327, BIOS AFE118M-1.32 released 06/29/2022 (ucode 0xd0003b9), Arch Linux, kernel 6.6.1-arch1-1, OpenSSL 3.1.4 24 Oct 2023 (library: OpenSSL 3.1.4 24 Oct 2023), Broadpeak BkS450 version 6f8bc8, gcc-12 (Ubuntu 12.1.0-2ubuntu1~22.04) 12.1.0, Test Tool version 6d70f0 (keep-alive, 8192 total connections). Throughput measured with 100% Transport Layer Security (TLS) traffic with 100% target cache hit ratio.

[5] Performance varies by use, configuration and other factors. For configurations, see [N010] at intel.com/processorclaims: 5th 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.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a nonexclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

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

1223/LH/MESH/353956-001US