Solution Brief

Content Delivery Networks (CDNs)
Intel® Xeon® 6 Processors with P-cores

Increased Capacity and Efficiency for CDNs Based on Single-Socket Servers

Rich 1 Socket Intel® Xeon® 6 processors offer industry-leading I/O capacity with 136 PCIe 5.0 lanes, supporting content delivery network (CDN) deployments that can access enhanced storage capacity using single-socket servers. With optimized software that can improve efficiency of memory accesses and TLS encryption, solutions based on Rich 1 Socket Intel Xeon 6 processors can offer reduced power consumption, solution complexity and TCO.

Continuing expansion of consumer demand for online media and entertainment, such as streaming video and gaming, is putting increased pressure on CDN operators to both expand content delivery capacity and improve their financial efficiency. Reliable, high-quality delivery of high-bandwidth content is a baseline requirement for providing experiences that attract and retain customers. This places a high value on low latency, driving CDNs out closer to the customers they serve.

Operators need to meet these evolving demands while maintaining favorable infrastructure costs and operating expenses, as the CDN market grows at an expected CAGR of more than 11% for the next several years.[1] Intel® platform innovations bring strategic advantages to support this growth.

Single-socket CPU architecture for CDNs

Rich 1 Socket Intel® Xeon® 6 processors enable a new class of high-density, single-socket servers based on intel architecture that can now address use cases with storage and network density requirements that would have called for a dual-socket system in prior generations. When compared to single-socket prior-generation Intel Xeon CPUs, the Rich 1 Socket Intel Xeon 6 processors offer substantially more PCIe lanes, allowing for increased native connectivity to storage, networking, or GPUs. To optimize single-socket performance and I/O density, the Rich 1 Socket architecture offers an additional 48 PCIe lanes in place of the Intel® Ultra Path Interconnect interfaces that would ordinarily handle communication between processors in a dual-socket configuration. This change provides a total of 136 PCIe 5.0 lanes for increased connectivity to storage, networking, GPUs, or other PCIe devices.

One-socket servers reduce the bill of materials to build out CDN server fleets and use less energy by powering fewer components. Software complexity is also reduced for single-socket systems, eliminating possible latency overhead associated with traversing the interconnect between processors when the additional compute and memory bandwidth of a dual-socket platform are not required. Platform advances of the Intel Xeon 6 processors help meet the computation, memory and I/O requirements for a balanced platform to handle CDN workloads:

- **High-throughput execution resources**. Improved per-core performance based on the latest Intel 3 technology, with up to 80 Performance-cores for Rich 1 Socket SKUs, as well as enhancements to instructions and built-in accelerators. (Higher core count options for regular SKUs are available.)
- **Enhanced memory subsystem.** Up to 14% increased DDR5 memory speed and 42% increased MR DDR5 memory speed provide better access to warm user plane data close to the processor cores.
- **Dramatically expanded I/O.** The processor's 136 lanes of PCIe 5.0 connectivity are key to enabling one-socket CDN servers and realizing the associated cost benefits.

As CDNs move outward, closer to end customers, these hardware resources enable them to maintain dependable operation and high data rates across thousands of parallel HTTPS request paths. Hardware acceleration for encryption and other key workloads helps ensure that excellent user experiences accompany favorable solution TCO.

CDNs on Intel® Xeon® 6 processors

UP TO

2.45x

higher throughput[2]

In a content cache capacity bound scenario, where throughput is bound by the application cache, a CDN can use the increased PCIe fanout of Rich 1 Socket Intel Xeon 6 processors to enable up to 2x the native NVMe storage capacity for each 800 Gbps node compared to the previous generation. The 136 PCIe lane count makes it possible to equip a single-socket 2U server with dual Gen5 x16 network interface cards and 24 Gen5 x4 NVMe drives, with PCIe capacity left to spare for GPUs or other uses. Alternatively, a converged edge node design could host multiple workloads on a shared system, such as AI inference for predictive content caching in addition to the core CDN workload.

Rich 1 Socket Intel® Xeon® 6 processors enable one-socket servers for CDNs

Replacing the Intel Ultra Path Interconnect interfaces with additional PCIe lanes on Rich 1 Socket Intel Xeon 6 processors enables single-socket servers that are ideal for CDN deployments.

- **Storage and networking density** equivalent to or better than previous-generation dual-socket systems.
- **136 PCIe 5.0 lanes** (88 lanes northbound, 48 lanes southbound) for storage, networking or GPUs.
- **Lower server and energy costs** from the efficiency of deploying one processor per system instead of two.

The reduced software and hardware complexity associated with the shift to single-socket systems makes it easier and more cost effective to deploy, manage and maintain CDN environments. The simpler server designs are more energy- and size-efficient, with the high I/O density supporting edge deployments with space and thermal constraints. In cases where software is licensed by the core, consolidating workloads onto single-socket servers may provide additional cost benefits.

Workload optimizations for the CDN solution ecosystem

Intel has a longstanding commitment to the CDN ecosystem that includes open source and commercial software enablement for features and capabilities of Intel hardware platforms. The optimizations outlined in this section help organizations meet rising performance goals while increasing efficiency to reduce operating costs.

CDN Software Optimizations Deliver Performance and Value on Intel® Xeon® 6 Processors

- Accelerated AI Inference to Guide Operations
- Improved Memory and Cache Data Handling
- Optimized Copy-and-Forward Operations
- Reduced Performance Overhead from Encryption

Accelerated AI inference to guide CDN operations

AI is an increasingly important component in CDN operational software, for tasks such as learning network behaviors to prevent traffic congestion and loading content as efficiently as possible on the right servers at the right time. AI inference is also an important part of modern cybersecurity workloads for capabilities such as anomaly detection. Intel Xeon 6 processors with P-cores provide hardware acceleration for AI workloads with Intel® Deep Learning Boost (Intel® DL Boost) to enable high-performance inference on the CPU. Acceleration with Intel DL Boost is also built into popular AI frameworks such as PyTorch and TensorFlow.

Intel DL Boost includes some Intel® Advanced Vector Instructions 512 (Intel® AVX-512) instructions as well as Intel Advanced Matrix Extensions (Intel® AMX). These technologies vectorize matrix multiplication, a core AI task, by enabling more data to be processed per clock cycle, increasing throughput for higher performance. The Intel Xeon 6 processors with P-cores introduce support for FP16 datatypes, which enables faster inference on some GPU-trained models. Deploying AI on Intel CPUs provides cost-effective resources for inference without add-on hardware, while the Rich 1 Socket Intel Xeon 6 processors also provide plenty of PCIe connectivity to add a GPU or other accelerator if needed.

Improved memory and cache data handling

Efficient data movement on CDN servers is critical to delivering performance and value from the hardware. Memory accesses commonly create contention over bandwidth and add latency to the delivery of content over the network to end customers. TLS bulk encryption adds latency to read data from memory and write it back, and additional delay accrues when data must be read from storage rather than memory. Intel works with CDN solution providers to optimize their software with techniques such as cache tuning to avoid unnecessary memory accesses, helping take full advantage of the memory subsystem advances built into Intel Xeon 6 processors.

Optimized copy-and-forward operations

The CPU of a CDN server constantly responds to content requests from the network interface with copy-and-forward operations. A typical flow is for the processor to fetch content from storage, write it to memory, encrypt it, copy it back to memory and then pass it to the outgoing network interface. Optimizing these copy-and-forward operations for better memory and CPU utilization can increase the capacity of the CDN node and reduce energy consumption, helping improve solution TCO. These measures also enable higher throughput on the CPU, helping reduce solution complexity and cost by avoiding the need for offloads to external hardware.

Reduced performance overhead from encryption

The need to encrypt content at scale is a core requirement for delivering high-bandwidth content over HTTPS, and performing TLS operations with high performance is a prerequisite for delivering great viewer experiences. Intel works with ecosystem partners to reduce the performance overhead associated with pervasive encryption. Intel Xeon 6 processors incorporate Intel® Crypto Acceleration, which draws on technologies including Intel AVX-512 and Intel® AES New Instructions to accelerate bulk cryptography and public key encryption. Updated crypto libraries also include an OpenSSL engine that is certified according to the Federal Information Processing Standard (FIPS). These crypto optimizations help drive throughput and add to the capacity of CDN servers, increasing their value.

We have worked with Intel closely optimizing content delivery performance and energy efficiency on many generations of Intel® Xeon® platforms. The Rich 1 Socket Intel Xeon 6 platform builds on the performance and efficiency gains by increasing the number of PCIe lanes allowing for greater fanout of NVMe drives versus previous single-socket generations. This delivers greater storage and networking density in a single-socket server and enables operators to reduce system and operational costs while improving energy efficiency compared to dual-socket configurations.

- Espen Braasad, Chief Architect, Varnish

CDN Software and the Intel® Industry Solution Builders Program

The Intel Industry Solution Builders program is an unparalleled ecosystem of solutions for communication service providers and enterprises that includes industry-leading CDN software optimized, tested and verified on Rich 1 Socket Intel Xeon 6 processors. In addition to enhancing the performance and energy efficiency of CDN workloads, this work benefits edge compute and content delivery providers with accelerated time to production and reduced risk.

Conclusion

Rich 1 Socket Intel Xeon 6 processors support single-socket CDN server deployments with up to 136 PCIe 5.0 lanes, enabling massive storage and networking density in a single-socket server. This advance helps reduce both system and operating cost, with a smaller bill of materials for CPUs and memory as well as reduced energy consumption when compared to a similarly configured dual-socket system. The expanded I/O resources are balanced with the CPU's increased performance per core, performance per watt and enhanced memory subsystem compared to predecessors. Intel enablement for CDN solutions on this hardware helps operators drive new capabilities, performance and energy efficiency for a lasting competitive advantage.

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

Learn More
www.intel.com/xeon
networkbuilders.intel.com

[1] Mordor Intelligence. "Content Delivery Network Market Size & Share Analysis - Growth Trends & Forecasts Analysis (2024 - 2029)." https://www.mordorintelligence.com/industry-reports/content-delivery-market.

[2] See [7N24] at intel.com/processorclaims: Intel® Xeon® 6. Results may vary.

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.

0225/DK/MESH/PDF 356898-001US