intel.

IT@Intel:
Increasing EDA Performance and Throughput with the Intel® Xeon® Processor Scalable Family

Intel IT testing of the Intel® Xeon® Gold 6400 processor Series yields the best per-core performance (up to 1.30x) compared to the Intel Xeon Gold 6300 processor Series. The Intel Xeon Platinum 8462Y+ processor delivers the best throughput (up to 2.40x) with high-core-count CPUs compared to the Intel Xeon Gold 6246R processor.

Authors

Shesha Krishnapura
Intel Fellow and IT CTO

Murty Ayyalasomayajula
Senior Staff Engineer

Shaji Kootaal Achuthan
Senior Staff Engineer

Vipul Lal
Senior Principal Engineer

Archana Somasekhara
Systems Engineer

Ty Tang
Senior Principal Engineer

Table of Contents

Executive Overview

Intel IT operates 54 data center modules at 15 data center sites. These sites have a total capacity of 104 megawatts, housing more than 390,000 servers that underpin the computing needs of more than 120,000 employees. Intel IT has four main segments of operation: Design, Office, Manufacturing and Enterprise. This paper focuses on only the Design segment.

Intel's silicon Design engineers need significant increases in computing capacity to deliver each new generation of silicon chips. To meet those requirements, Intel IT conducts ongoing throughput performance tests using real-world Intel silicon Design workloads. These tests measure Electronic Design Automation (EDA) workload throughput and help us analyze the performance improvements—and in turn, the business benefits offered by newer generations of Intel® processors.

We recently tested two-socket servers based on the Intel® Xeon® Platinum 8400 and Gold 6400 processor Series, operating single- and multi-threaded EDA applications running Intel silicon Design workloads for more than 113 hours (about 4 and a half days). Select results include the following:[1]

- Higher frequency for per-core performance. For critical-path EDA workloads, selecting a high-frequency CPU like the Intel Xeon Gold 6444Y processor (32 cores per server) can deliver up to 1.14x higher per-core performance compared to lower-frequency, higher-core-count CPUs in the same generation of processors.

- Higher core counts for throughput. For volume validation runs, selecting a higher-core-count CPU at optimal frequency like the Intel Xeon Platinum 8462Y+ processor (64 cores per server) can deliver up to 1.75x higher Register Transfer Level (RTL) Simulation throughput per server when compared to a lower-core-count CPU (32 cores per server) in the same generation of processors. The Intel Xeon Platinum 8462Y+ processor (64 cores per server) completed workloads up to 2.17x faster than a previous-generation Intel Xeon Gold 6346 processor-based server, which has only 32 cores. Compared to a 2nd Gen Intel Xeon Gold 6246R processor (32 cores per server), the server with the newer processor outperformed the older processor by up to 2.40x in throughput.

Based on our performance assessment and our refresh cycle, we are deploying servers based on the 4th Gen Intel Xeon Scalable processor family in our data centers. By doing so, we have significantly increased EDA throughput performance to improve the overall EDA design cycles and optimize time to market of Intel® chips.

Intel IT Contributors

Matt Ammann, Principal Engineer for providing server network operations support

Juan J. Sanchez and Chandra Sudireddy, IT Operations for providing multiple server configurations and operational support

Background

Silicon chip Design engineers at Intel face ongoing challenges: integrating more features into ever-shrinking silicon chips, bringing products to market faster and keeping Design engineering and manufacturing costs low. Design engineers run more than 273 million compute-intensive batch jobs every week. Each job takes from a few seconds to several days to complete.
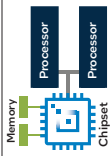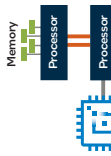
As design complexity increases, so do the requirements for compute capacity, so refreshing servers and workstations with higher-performing systems is cost-effective and offers a competitive advantage by enabling faster chip design. Refreshing older servers also enables us to realize data center cost savings. By taking advantage of the performance and power-efficiency improvements in new server generations, we can increase computing capacity within the same data center footprint, helping to avoid expensive data center construction and reduce operational costs due to reduced power consumption.

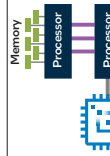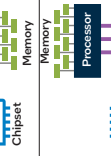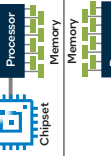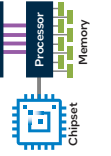Intel IT conducts ongoing performance tests, based on the latest Intel silicon Design data, to analyze the potential performance and data center benefits of introducing servers based on new processors into our Electronic Design Automation (EDA) computing environment.

Evolution of the Intel® Xeon® Processor and the EDA Workflow

The architectural enhancements shown in Table 1 illustrate how the Intel® Xeon® processor has evolved over nearly two decades. We have found that refreshing data center servers to use the latest processor technology substantially improves EDA throughput.

While our assessments focus on EDA applications, throughput improvements may also be achieved with other applications used in high-performance computing environments where simulation and verification are large parts of the workflow, including:

- Computational fluid dynamics and simulation in the aeronautical and automobile industries
- Synthesis and simulation applications in the life sciences industry
- Simulation in the oil and gas industries

Table 1. Comparison of Two-socket Servers Based on Intel® Xeon® Processors Over Time

	2004-2005	2006-2008	2009-2011	2012	2013	2014	2016	2017-2020	2021-2022	2023
Design										
Intel® Chipset	E7520	5400	5520	C600		C610		C620	C620A	C741
Process Technology	90nm	65nm and 45nm	45nm and 32nm	32nm	22nm			14nm	10nm	Intel® 7
Cores per Socket	1	2 or 4	4 or 6	8	10	14	22	28	40	60
Interconnect Speed	6.4 GT/s	21-25 GT/s	25.6 GT/s	32 GT/s		38.4 GT/s		41.6 GT/s	44.8 GT/s	64 GT/s
Intel® QPI	—	—	Yes	Yes				—	—	—
Intel® UPI	—	—	—	—				Yes	Yes	Yes
DIMMs	Up to 8	Up to 16	Up to 18	Up to 24					Up to 32	Up to 32
Memory Type	DDR2	FB-DIMM/DDR2	DDR3	DDR3	DDR3	DDR4	DDR4	DDR4	DDR4	DDR5
Memory Speed (MHz)	400	667 or 800	800/1066/1333	1333/1600	1333/1600/1866	1600/1866/2133	2400	2666	3200	4800
Memory Bandwidth	Up to 6.4 GB/s	Between 21-25 GB/s	Up to 32 GB/s	Up to 51.2 GB/s	Up to 59.7 GB/s	Up to 68 GB/s	Up to 76.8 GB/s	Up to 128 GB/s	Up to 204.76 GB/s	Up to 307.14 GB/s
Max Memory per Server (2004-08) or CPU (2009-23)	16 GB	64 GB or 128 GB	144 GB or 288 GB	384 GB or 768 GB		768 GB or 1.5 TB		Up to 768 GB; 1 TB or 4.5 TB	Up to 6 TB[a]	Up to 4 TB

[a] 6 TB with Intel® Optane™ memory+DDR memory, 4 TB with DDR memory.

As shown in Figure 1, EDA includes several phases, including front-end logic design, followed by back-end physical design and then by tape-in/tape-out. This paper discusses selected tools in the front-end and back-end design phases.
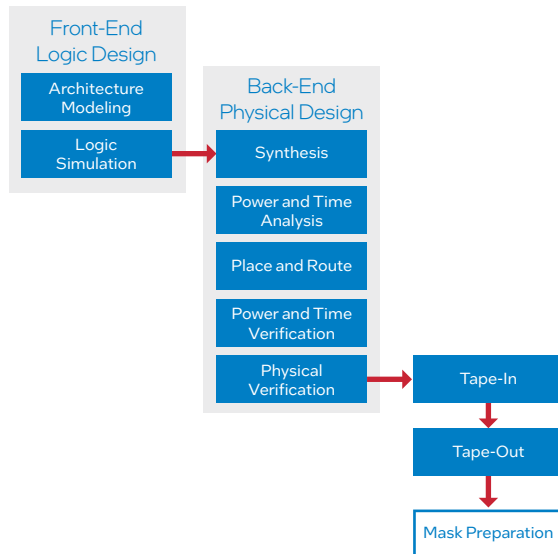
Figure 1. The EDA phases of silicon design.

Front-end Logic Design

In the silicon design process, front-end logic design includes architecture specification and functional verification design. Front-end EDA workloads are single-threaded or lightly multi-threaded and run in a highly distributed compute environment.

Back-end Physical Design

Synthesis converts a Register Transfer Level (RTL) description to a structural gate-level netlist, which instantiates standard cells, macros and areas that compose the circuit and its connections. The synthesized netlist is verified for

functionality and timing to ensure it operates as intended before the Place-and-Route stage translates the gate-level netlist into a physical design. Static-Timing-Analysis as well as other post-layout static and dynamic analyses are then performed to check all possible paths for timing violations, voltage drop analysis and more, and to deliver accurate signoff information for timing, signal integrity and power analysis for the design. Finally, the Physical-Verification stage ensures the physical design meets manufacturing constraints imposed by process technology; the verification includes Design Rule Check, Layout versus Schematic and Electrical Rule Check. Back-end EDA workloads are generally multi-threaded and consume large memory and terabytes of data.

Tape-in/Tape-out

During tape-in, Intel chip design teams create multi-gigabyte hierarchical layout databases that specify the design to be manufactured. During tape-out, these layout databases are processed using EDA tools, which apply extremely compute-intensive resolution enhancement techniques (RET) to update layout data for mask manufacturability and verify the data for compliance to mask manufacturing rules.

Test Methodology

We performed various tests on two-socket servers. Some tests compared several different CPUs in the 4th Gen Intel® Xeon® Platinum 8400 and Gold 6400 processor Series (see Table 2). Other tests compared the 4th Gen processors to a baseline of 3rd Gen Intel Xeon Gold 6300 and 2nd Gen Intel Xeon Gold 6200 processor Series. We conducted front-end and back-end tests using industry-leading EDA applications to run single- and multi-threaded Intel silicon Design workloads.

Our goal was to assess performance and throughput improvements by measuring the time needed to complete a specific number of Design workloads. To maximize throughput, we configured each application to utilize all available cores, resulting in one job or process per core wherever possible.

Table 2. Specification of the Systems Used for Testing

	2nd Gen Intel® Xeon® Scalable Processor Family		3rd Gen Intel Xeon Scalable Processor Family		4th Gen Intel Xeon Scalable Processor Family		
Processor	Intel Xeon 6246R	Intel Xeon 6240R	Intel Xeon 6346	Intel Xeon 6342	Intel Xeon 6444Y	Intel Xeon 6442Y	Intel Xeon 8462Y+
Cores per Socket	16	24	16	24	16	24	32
Frequency	3.4 GHz	2.4 GHz	3.1 GHz	2.8 GHz	3.6 GHz	2.6 GHz	2.8 GHz
Cache per CPU	35.75 MB	35.75 MB	36 MB	36 MB	45 MB	60 MB	60 MB
Bus Speed	10.4 GT/s	10.4 GT/s	11.2 GT/s	11.2 GT/s	16 GT/s	16 GT/s	16 GT/s
RAM	768 GB	768 GB	1 TB	1 TB	1 TB	1 TB	1 TB
Memory Type	DDR4-2933 MHz	DDR4-2933 MHz	DDR4-3200 MHz	DDR4-3200 MHz	DDR5-4800 MHz	DDR5-4800 MHz	DDR5-4800 MHz

Testing by Intel IT as of January through March 2023.

Results: Faster Servers Process More EDA Jobs in Less Time

Test results are shown in Figures 2 through 6; system specifications and runtimes are provided in Table 2 (on previous page) and Table 3 (below).

Benefits of the 4th Gen Intel Xeon Scalable Processor Family

We conducted tests to understand throughput and the core-to-core speed increase that we obtained using 4th Gen Intel Xeon Scalable processors instead of 3rd Gen Intel Xeon Scalable processors. For this test, we selected three different SKUs from each generation:

- A 16-core high-frequency SKU with a total of 32 cores in a two-socket system—Intel Xeon Gold 6346 processor versus Intel Xeon Gold 6444Y processor.
- A 24-core scalable performance SKU with a total of 48 cores in a two-socket system—Intel Xeon Gold 6342 processor versus Intel Xeon Gold 6442Y processor.
- A 32-core higher scalable performance SKU with a total of 64 cores in a two-socket system—Intel Xeon Platinum 8462Y+ processor.

As shown in Figure 2, comparing the performance of same-core-count CPUs across generations results in performance increases up to 1.30x.

Optimizing Platform Selection within the 4th Gen Intel Xeon Scalable Processor Family

For volume validation runs, overall cluster throughput is desirable; but for critical-path runs, the highest per-core performance is needed. Both types of workloads can be supported using a variety of Intel Xeon Scalable processor family SKUs. We compared selected 16-core, 24-core and 32-core offerings from the 4th Gen Intel Xeon Scalable processor family. This provides a choice for critical path versus volume validation runs.

- A 16-core high-frequency SKU: Intel Xeon Gold 6444Y processor.
- A 24-core scalable performance SKU: Intel Xeon Gold 6442Y processor.
- A 32-core higher scalable performance SKU: Intel Xeon Platinum 8462Y+ processor.

Relative Two-Socket System Throughput for EDA Workloads
3rd Gen Intel® Xeon® Processors vs. 4th Gen Intel® Xeon® Processors
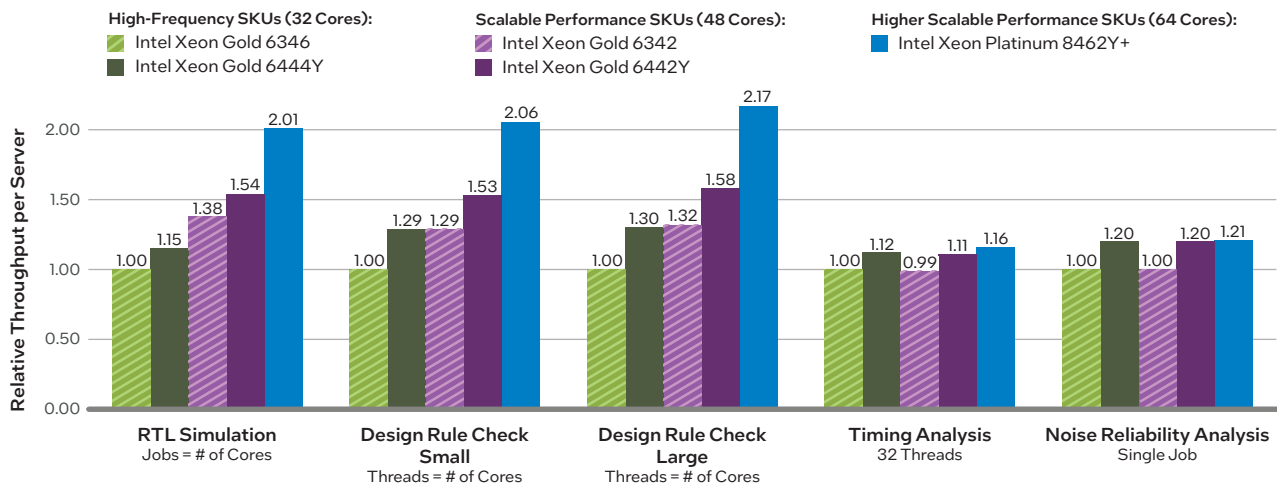Higher Is Better

Figure 2. 4th Generation Intel® Xeon® Scalable processor vs. 3rd Generation Intel Xeon Scalable processor: More cores provide more system throughput for EDA workloads. Note: Same application binary used across all of the platforms.[1]

Table 3. Workload Run Times

| Workload Cores per Server	2nd Gen Intel® Xeon® Scalable Processor Family		3rd Gen Intel Xeon Scalable Processor Family		4th Gen Intel Xeon Scalable Processor Family		
	Intel Xeon 6246R 32 cores	Intel Xeon 6240R 48 cores	Intel Xeon 6346 32 cores	Intel Xeon 6342 48 cores	Intel Xeon 6444Y 32 cores	Intel Xeon 6442Y 48 cores	Intel Xeon 8462Y+ 64 cores
RTL Simulation	1:10:19	1:00:02	1:07:19	0:48:45	0:58:31	0:43:37	0:33:29
Design Rule Check - Small	1:16:06	1:09:05	1:08:51	0:53:24	0:53:19	0:44:57	0:33:30
Design Rule Check - Large	3:50:27	3:24:29	3:28:27	2:37:33	2:40:44	2:12:18	1:36:09
Static Timing Analysis	3:25:30	3:42:33	2:58:04	2:59:40	2:39:39	2:40:51	2:33:03
Power Noise Reliability	7:36:40	7:50:06	6:12:57	6:11:43	5:10:20	5:11:05	5:07:23

NOTE: Benchmarks were conducted with all the cores loaded; static timing analysis workload is limited to a maximum of 32 threads; power noise reliability is a single job. Results are dependent on tool type, version, and data set.

Our findings, shown in Figures 3 and 4, indicate the following best practices:

- Selecting a higher-frequency CPU (3.6 GHz) can deliver up to 1.14x higher RTL Simulation per-core performance for critical-path EDA runs.

- Selecting a higher-core-count CPU (32 cores) can deliver up to 1.75x higher RTL Simulation throughput per server when compared to a lower-core-count SKU, which is ideal for volume validation runs. The choice of optimal SKU is based on the end-user workload use model. EDA jobs rarely scale beyond 32 threads, so a 32-core processor with optimal frequency is sufficient to accommodate high thread counts and high-throughput, single-threaded EDA jobs.

Relative Performance Per Core For RTL Simulation
Higher Is Better (2 Sockets)
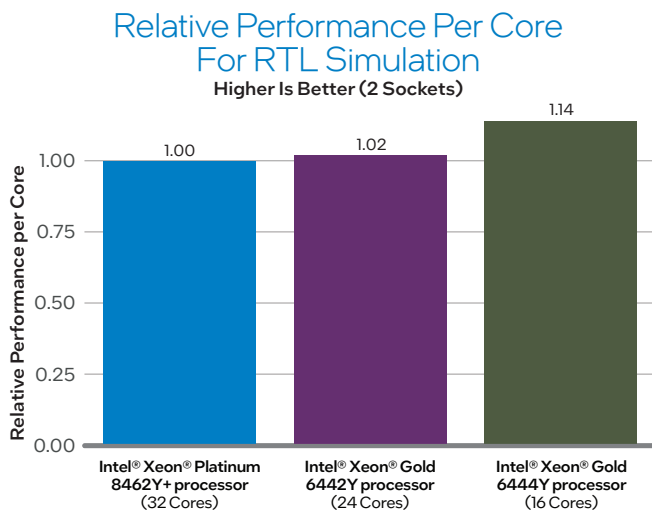
Figure 3. 4th Generation Intel® Xeon® Scalable processors: A higher frequency results in better per-core performance for critical-path EDA workloads. Note: Same application binary used across all of the platforms.[1]

Relative System Throughput Per Server For RTL Simulation
Higher Is Better (2 Sockets)
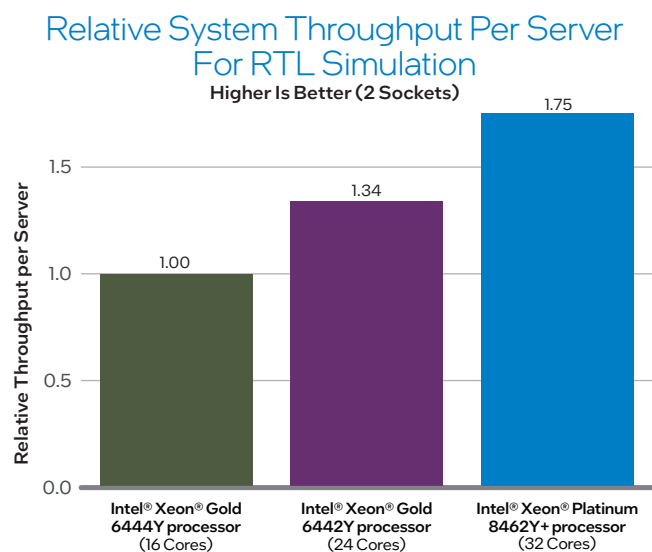
Figure 4. 4th Generation Intel® Xeon® Scalable processor: A higher core count results in better per-system throughput for volume validation runs. Note: Same application binary used across all of the platforms.[1]

Intel's Latest Processors for Data Center Workloads

4th Gen Intel® Xeon® Scalable processors are packed with performance- and security-enhancing features, including the following:

- Intel® 7 process technology.

- Enhanced per-core performance, with up to 60 cores in a standard socket.

- Enhanced memory performance with support for up to 4800 MT/s DIMMs—50% more than the previous generation.

- Up to 4 TB memory capacity with up to eight channels.

- Built-in Deep Learning acceleration with Intel® Advanced Matrix Extensions.

- Intel® Data Streaming Accelerator (Intel® DSA) drives high performance for storage, networking and data-intensive workloads.

- Faster inter-node connections with multiple Intel® Ultra Path Interconnect links at 16 GT/s.

EDA Per-Core Performance Across Three Generations of Intel Xeon Processors

We tested three generations of Intel Xeon processors to compare the per-core performance for critical-path EDA workloads:

- 16-core Intel Xeon Gold 6246R processor (2nd Gen)
- 16-core Intel Xeon Gold 6346 processor (3rd Gen)
- 16-core Intel Xeon Gold 6444Y processor (4th Gen)

Based on the test results, for performance-critical workloads, a 16-core 4th Gen Intel Xeon Gold 6444Y processor-based server can reduce license consumption time and provide a performance boost of 1.20x to 1.47x across all the tested workloads (see Figure 5).

EDA Throughput Across Three Generations of Intel Xeon Processors

We tested three generations of Intel Xeon processors to compare the throughput:

- 24-core Intel Xeon Gold 6240R processor (2nd Gen)
- 24-core Intel Xeon Gold 6342 processor (3rd Gen)
- 24-core Intel Xeon Gold 6442Y processor (4th Gen)
- 32-core Intel Xeon Platinum 8462Y+ processor (4th Gen)

Based on the test results, a 4th Gen Intel Xeon Platinum 8462Y+ processor-based server provides up to a 2.13x increase in throughput per server, compared to a 2nd Gen Intel Xeon Gold 6240R processor-based server (see Figure 6). In other words, 17 older Intel Xeon Gold 6240R processor-based servers can be replaced with only eight servers based on the latest generation of Intel Xeon Scalable processors.

Relative Two-Socket Performance Per Core for EDA Performance
Across 3 Generations of Intel® Xeon® Processors
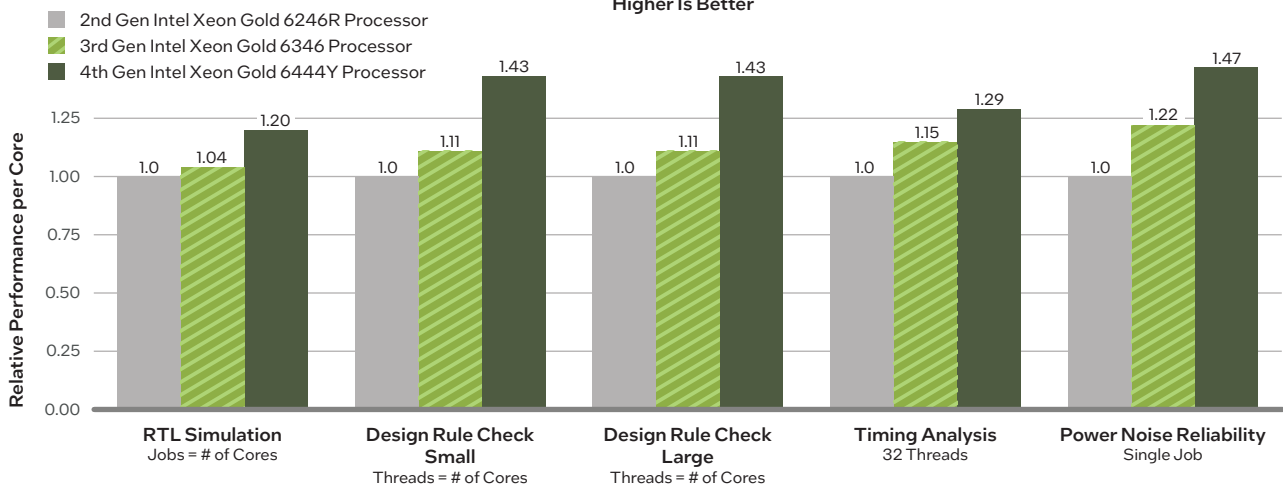##### Higher Is Better

Figure 5. Per-core EDA performance of select Intel® Xeon® processors across three generations. Note: Same application binary used across all of the platforms.[1]

Relative Two-Socket System Throughput for EDA Workloads
Across 3 Generations of Intel® Xeon® Processors
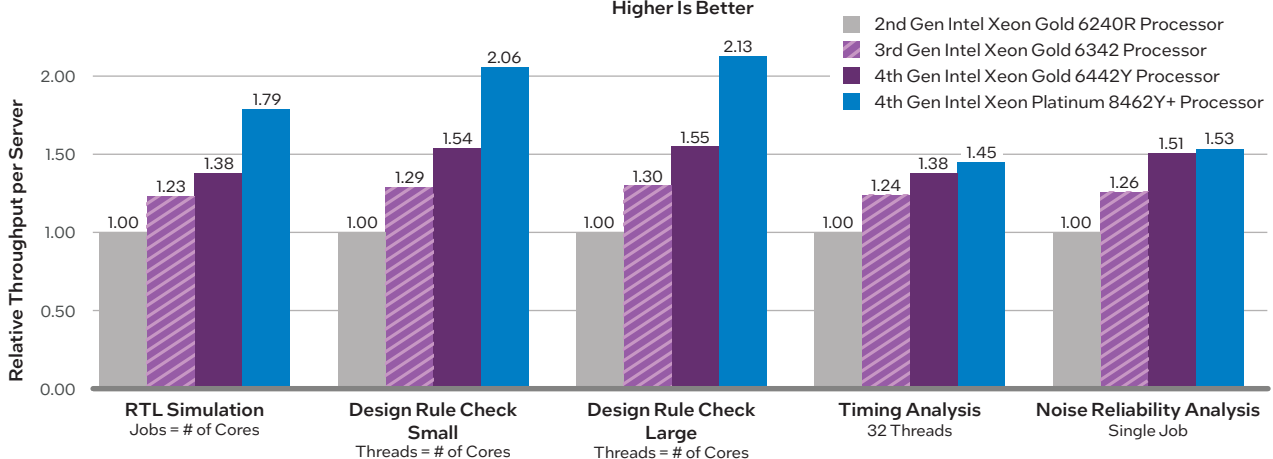##### Higher Is Better

Figure 6. EDA throughput of select Intel® Xeon® processors across three generations: Consolidate servers by up to 2.13x to reduce data center footprint, power and cooling costs, and software license costs.[1]

Conclusion

The latest Intel Xeon Scalable processor family deliver significant improvements in throughput and per-core performance for Intel silicon Design workloads across a range of EDA applications in the data center.

At the same core-count level, workloads on servers with 4th Gen Intel Xeon processors can run up to 1.3x faster than a previous-generation server based on 3rd Gen Intel Xeon processors. For performance-centric workloads, selecting a higher-frequency 4th Gen Intel Xeon processor can increase performance by up to 1.14x, compared to lower-frequency 4th Gen Intel Xeon processor. For throughput-centric workloads, an Intel Xeon 8400 processor Series SKU with high frequency and core count can deliver up to 2.13x better throughput than a 2nd Gen Intel Xeon Gold processor.

Using our tested end-to-end EDA applications based on Intel silicon Design tests, we found that we can replace up to 17 older-generation Intel Xeon Gold 6240R processor-based servers with just eight of the latest Intel Xeon Platinum 8400 processor Series-based servers. Based on our performance assessment and our refresh cycle, we have deployed servers based on the new Intel Xeon Platinum 8400 and Gold 6400 processor Series, enabling us to achieve greater throughput while realizing operational benefits such as cost avoidance of data center construction and reduced power consumption.

Our test results suggest that other technical applications with large memory requirements—such as simulation and verification applications in the auto, aeronautical, oil and gas and life sciences industries—could see similar throughput improvements, depending on their workload characteristics.

Related Content

If you liked this paper, you may also be interested in these related papers:

- Data Center Strategy Leading Intel's Business Transformation white paper
- Disaggregated Servers Drive Data Center Efficiency and Innovation white paper

For more information on Intel IT best practices, visit intel.com/IT.

IT@Intel

We connect IT professionals with their IT peers inside Intel. Our IT department solves some of today's most demanding and complex technology issues, and we want to share these lessons directly with our fellow IT professionals in an open peer-to-peer forum.

Our goal is simple: improve efficiency throughout the organization and enhance the business value of IT investments.

Follow us and join the conversation on **X** or **LinkedIn.**

Visit us today at **intel.com/IT** or contact your local Intel representative if you would like to learn more.