Technical Research Report

Unlocking Cost and Efficiency Gains with Next-Generation Single-Socket Servers

This report compares the power usage and performance advantages of consolidating from a dual-socket server to a single-socket server using Dell™ PowerEdge™ R740xd and Dell PowerEdge R570 servers as the testing ground.

Executive Summary

Businesses face ongoing pressure to improve performance while reducing operating costs and simplifying IT infrastructure. Consolidating workloads from dual-socket to single-socket servers can help achieve these goals by shifting workloads onto fewer, more efficient servers to boost performance and reduce energy consumption. This paper compares the benefits and business value of shifting from the dual-socket Dell PowerEdge R740xd server to the single-socket PowerEdge R570 server, both running on Intel® Xeon® processors. In this report, Prowess Consulting examines how this modern system architecture can offer greater raw performance and better performance per watt.

To ensure accuracy, we ran multiple configurations, found where the dual-socket PowerEdge R740xd server reached peak performance, and then made our comparisons with the PowerEdge R570 server in the same configuration. The hardware used in this study represents the type of hardware a business might adopt during a typical server refresh. We focused on efficiency and performance to demonstrate how consolidation can impact real-world workloads by comparing generational performance in the same configuration. To that end, we conducted a series of controlled tests to quantify the benefits of this consolidation strategy, including:

- Speech-to-text-to–sentiment analysis
- Apache HTTP Server™ benchmarking
- NGINX®/Redis® caching workloads

Our initial findings from these tests indicate that consolidating workloads onto a modern single-socket PowerEdge R570 server can deliver meaningful efficiency and performance benefits in critical areas like energy consumption, performance, and infrastructure footprint. For organizations planning a refresh cycle, this consolidation approach provides a clear and practical path to improving efficiency and scalability without compromising real-world performance.

Highlights

Prowess Consulting found that in the same configuration, compared to the dual-socket Dell PowerEdge R740xd server, the single-socket PowerEdge R570 server delivered up to:

2.36x

greater raw performance in sentiment analysis

2.32x

greater performance per watt in media caching

1.82x

greater raw performance in web serving

Determining Server Use Cases and Value for Organizations

In this study, Prowess Consulting evaluated the benefits and potential drawbacks of transitioning from dual-socket servers to single-socket servers using PowerEdge servers and Intel Xeon processors as our testbed. Our testing focused on workloads that would help us quantify the potential benefits of consolidation for compute density, raw performance, and energy consumption. These metrics speak to the critical use cases in which these servers are often deployed. In this study, our use cases include AI-powered speech-to-text and sentiment analysis workloads, web serving and caching, and application acceleration. These common scenarios reflect real-world demands on small and medium-sized businesses (SMB) and enterprise IT environments, where responsiveness, throughput, resource efficiency, and raw power directly impact user experience, operational effectiveness, and cost management.

Why Businesses Might Consider Consolidation

The PowerEdge R740xd server, powered by two Intel Xeon Gold 5217 processors in our testing, provides multi-socket scalability, but it has relatively low per-socket core counts. However, the PowerEdge R570 server uses a single Intel Xeon 6787P processor with a significantly higher core density, delivering more compute power per socket. The PowerEdge R570 server also comes with native Integrated Dell™ Remote Access Controller 10 (iDRAC10), rather than iDRAC9, which is included with the PowerEdge R740xd server. iDRAC10 provides enhanced security features, richer system telemetry, and increased automation, which can all contribute to lower administrative overhead in large or distributed environments. From a business perspective, these architectural improvements translate into tangible operational benefits such as better performance per watt, more raw performance, and stronger throughput.

The architectural improvements we observed also position the PowerEdge R570 server as a more capable consolidation platform, allowing organizations to support mixed workloads on fewer systems without sacrificing reliability or responsiveness. By enabling higher infrastructure density and more efficient resource utilization, the single-socket design from Dell Technologies can help businesses simplify deployment by reducing operational overhead and creating more room for future growth within existing power and space constraints.

Our testing covered three workloads, each representing common everyday business tasks, which makes our findings directly relevant to IT decision-makers (ITDMs) evaluating infrastructure efficiency, performance, and cost-effectiveness:

- **Speech-to-text-to–sentiment analysis:** Workloads like speech transcription and sentiment scoring are increasingly common for customer support and operational intelligence. Crucial metrics include files processed per second and energy efficiency. In our research, we found that the single-socket PowerEdge R570 server offers a strong hardware configuration for faster processing, in addition to better performance per watt, thereby enhancing productivity. This translates into faster decision-making for customer experience teams while enabling lower operational costs.
- **Apache HTTP Server:** Web-serving workloads directly impact end-user experiences and operational continuity. Measured throughput, request latency, and concurrency scaling are key indicators of how effectively a server can handle real-world traffic. In this study, we found that the PowerEdge R570 server demonstrated higher requests per second with lower latency than the PowerEdge R740xd server, while maintaining CPU headroom under sustained load. This enables the consolidation of workloads without sacrificing responsiveness, ultimately reducing infrastructure costs and simplifying maintenance.
- **NGINX and Redis caching:** For e-commerce platforms and media-heavy services, caching and content delivery efficiency are critical. Metrics such as request latency, requests per second (RPS), and power utilization highlight how well a server can support high-demand, real-time workloads. We found that the PowerEdge R570 server's architectural improvements allow for better performance per watt, more RPS, and lower latency under sustained high-concurrency traffic. This ensures smoother customer experiences through reduced latency for end users while maintaining large-scale media services.

Initial Conclusions and Potential Benefits

Our evaluation of 17th-generation PowerEdge R570 servers found that single-socket consolidation can provide clear performance advantages over older dual-socket 14th-generation PowerEdge R740xd servers. Testing across real-world workloads showed that the PowerEdge R570 server consistently delivers higher throughput with greater raw performance and better performance per watt in a single-socket configuration.

The power-efficiency improvements are equally notable, with the PowerEdge R570 server offering significantly better performance per watt as workloads scale. On average, within the same test configuration and workload, the PowerEdge R570 server delivered much better results, showing that, for the energy consumed, the PowerEdge R570 server offers greater performance, making it the better choice for businesses. Upgrading to the 17th-generation PowerEdge R570 server ensures organizations gain sustained efficiency and measurable real-world impact from their infrastructure investments.

Test Plan and Results

We compared the performance, efficiency, and consolidation capabilities of the PowerEdge R740xd and PowerEdge R570 servers across the three representative workloads described earlier. For our testing, we configured both servers with the latest firmware and with the required tools to perform the tests. We performed no custom operating system (OS) tuning beyond standard best practices, and we captured metrics including workload throughput and latency, in addition to server power consumption and resource utilization, providing a comprehensive view of each system's real-world performance and operational efficiency.

For a complete breakdown of the system configurations, refer to the **Appendix**.

Tested Workloads

Modern businesses increasingly rely on data-intensive and latency-sensitive applications to drive operations, customer experience, and revenue. To evaluate how effectively servers can handle these demands, we selected three representative workloads to stress different aspects of server performance: CPU and memory utilization, throughput, and, most importantly, raw performance, all of which provide a comprehensive view of real-world application behavior.

By measuring the three chosen workloads, we can quantify how server choice impacts operational efficiency, responsiveness, and energy consumption. For SMBs, these performance differences translate directly into system reliability and cost-effectiveness. Understanding how modern single-socket servers with Intel Xeon processors handle these scenarios can help IT teams make informed decisions about consolidation during infrastructure upgrades and application deployment strategies.

Note: To ensure a fair comparison between the configurations, our conclusions are based on performance data using the exact same configuration for each server. This consistency ensures results aren't based on the higher available core counts of the PowerEdge R570 server but on performance data from identical configurations. We included additional information on performance numbers for configurations below and above this standardized point to provide further insight into our testing results, but these numbers are not factored into our conclusions.

Speech-to-Text-to–Sentiment Analysis

The speech-to-text-to–sentiment analysis workload simulates how SMBs can use customer support interactions to improve service quality and track team effectiveness. For this workload, recorded audio from support calls or customer interactions is processed to generate transcriptions and sentiment scores, which can be used for training, quality assurance (QA), and identifying opportunities for operational improvement.

This workload reflects the high-power processing common in both SMB and enterprise deployments. It combines high CPU and memory utilization with significant power-draw demands, making it an effective benchmark to determine how servers powered by Intel Xeon processors handle parallel high-volume, data-intensive workloads.

Our testing captured detailed performance metrics, including power draw, concurrent processes, throughput, and performance per watt to evaluate how efficiently each server configuration handled the workloads. This enables us to determine raw performance numbers and average energy efficiency in each respective test case.

Better Performance in Sentiment Analysis

Our results indicate a clear performance and efficiency advantage for the single-socket PowerEdge R570 server. At the matched operating point, where the prior-generation system reached peak practical throughput, the PowerEdge R570 server sustained higher overall workload throughput and better performance per watt.

Notably, the PowerEdge R570 server supported greater concurrency, sustaining an average of 15.71 simultaneous processes compared to 6.63 on the PowerEdge R740xd server in the 32-core workload, offering up to 2.36x higher throughput before reaching performance limits (see Figure 1). Table 1 compares the resulting data at 8, 16, and 32 concurrent core workloads, with the 32 concurrent workloads highlighted as our key standout result.

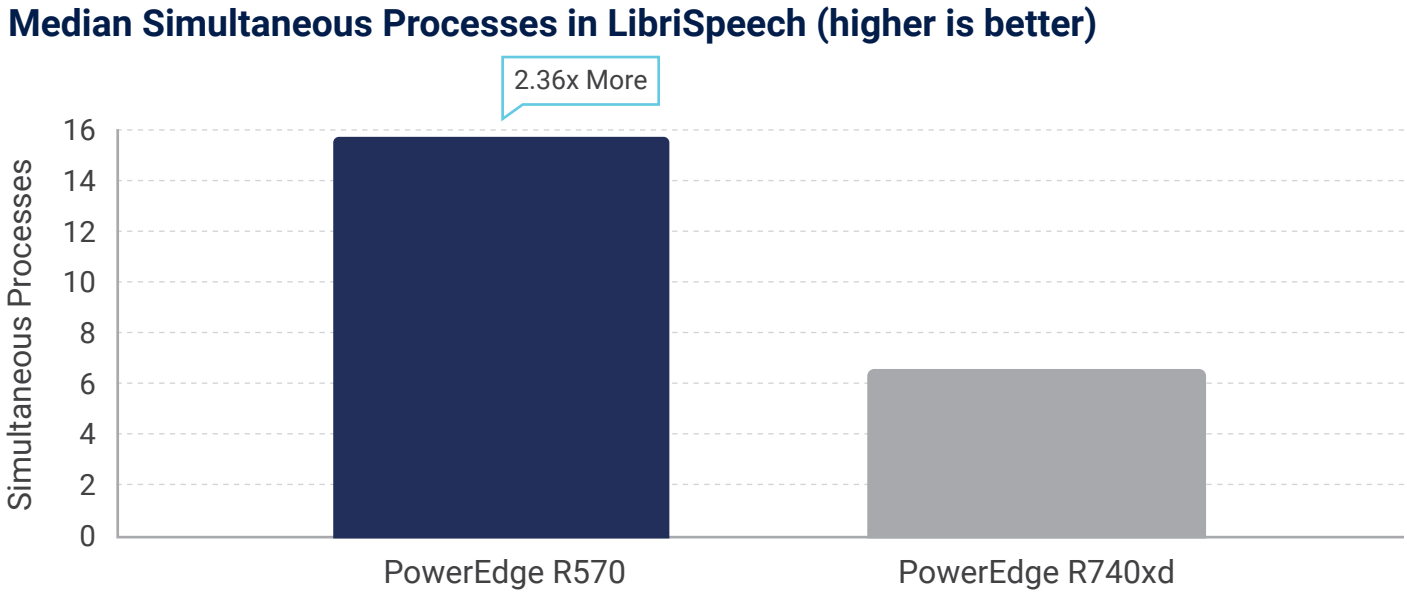

Figure 1 | Bar chart comparing simultaneous processes between the dual-socket Dell PowerEdge R740xd server and the single-socket PowerEdge R570 server in a sentiment analysis workload at the matched 32-core comparison point

Table 1 | The greater performance of the single-socket Dell PowerEdge R570 server versus the dual-socket PowerEdge R740xd server

Core Count	Dell PowerEdge R740xd Median	Dell PowerEdge R570 Median	Raw Performance Advantage of Dell PowerEdge R570 vs. PowerEdge R740xd
8	3.33	5.47	1.64
16	5.77	9.37	1.62
32	6.63	15.71	2.36

From a performance and consolidation angle, this increased processing capacity allows more customer interactions to be analyzed in parallel, leading to better operational responsiveness and more efficient use of power. These gains also directly translate into improved infrastructure efficiency.

Consolidating to a single-socket PowerEdge R570 server configuration can enable organizations to process more sentiment analysis workloads on fewer CPU sockets, creating better performance per watt and better power efficiency. For example, the PowerEdge R570 server achieved 2.36x higher performance than the two-socket PowerEdge R740xd server but drew only 1.73x the power (449.74 W versus 259.73 W), netting 1.34x higher performance per watt.[1]

Overall, these results confirm that, compared to the dual-socket PowerEdge R740xd server, the single-socket PowerEdge R570 server provides:

- Improved performance per watt
- Greater workload concurrency and processing capacity
- More efficient infrastructure utilization for sentiment analysis workloads

Apache HTTP Server (Apache® Bench and Apache JMeter™)

The Apache HTTP Server workload, tested using Apache Bench and Apache JMeter, simulates how a business would evaluate the performance and scalability of a web server under realistic traffic conditions. Apache Bench provides a lightweight command-line tool for generating controlled HTTP request loads, while Apache JMeter enables more complex, multi-threaded test scenarios that mimic real user behavior. Together, these tools measure how efficiently a server handles concurrent requests and maintains low latency while sustaining high throughput.

This workload simulates an SMB evaluating the deployment of an Apache web server in a remote office or branch environment. Using servers powered by Intel Xeon processors, the test captures real-world web-serving scenarios. We measured metrics such as request latency, concurrent workloads, and RPS to compare performance between the dual- and single-socket configurations. This workload reflects the sustained, user-facing demand common in both SMB and enterprise web environments. Because it monitors high concurrency, it's an effective benchmark for evaluating how servers maintain responsiveness and stability under real-world web-serving conditions.

Lower Latency, More RPS

At the matched operating point, where the prior-generation system peaked, our results show clear performance improvements on the PowerEdge R570 server compared to the PowerEdge R740xd server. In this testing configuration with Apache Bench, the PowerEdge R570 server handled more RPS with lower average latency and with zero errors from both servers. With an equivalent configuration, the PowerEdge R570 server managed 16% more RPS (see Figure 2) with 13% lower average latency (see Figure 3) when running a workload with 10 parallel processes.

Responses per Second in Apache Bench (higher is better)

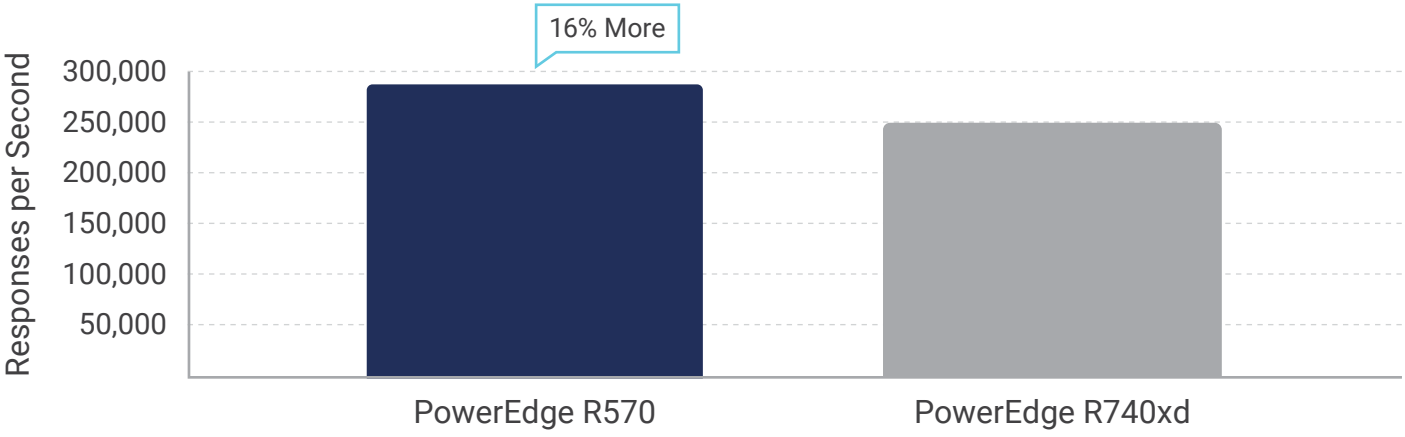

Figure 2 | RPS between the dual-socket Dell PowerEdge R740xd server and the PowerEdge R570 server in Apache Bench at the matched 32-core comparison point

Median Latency in Apache Bench (lower is better)

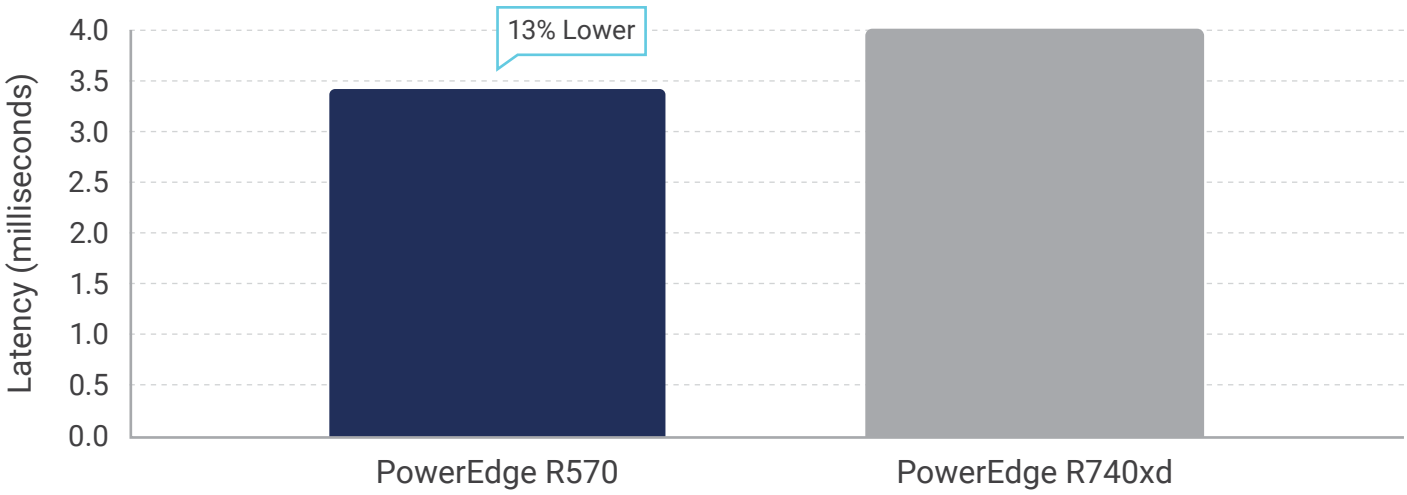

Figure 3 | Median latency numbers between the dual-socket Dell PowerEdge R740xd server and the PowerEdge R570 server in Apache Bench at the matched 32-core comparison point

Because the PowerEdge R570 server's higher throughput and lower latency allow more web requests to be processed simultaneously, the server offers better results for businesses interested in consolidating to single-socket servers. The PowerEdge R570 server demonstrated improved responsiveness under realistic traffic conditions with more efficient use of system resources when comparing equal workloads, as compared to the PowerEdge R740xd server. These gains directly translate into better infrastructure efficiency to handle higher web-serving loads with fewer servers and with better power efficiency.

In the equivalent configuration for Apache JMeter (16 worker nodes x 64 threads each), the PowerEdge R570 server delivered 82% more RPS (see Figure 4).

Responses per Second in Apache JMeter (higher is better)

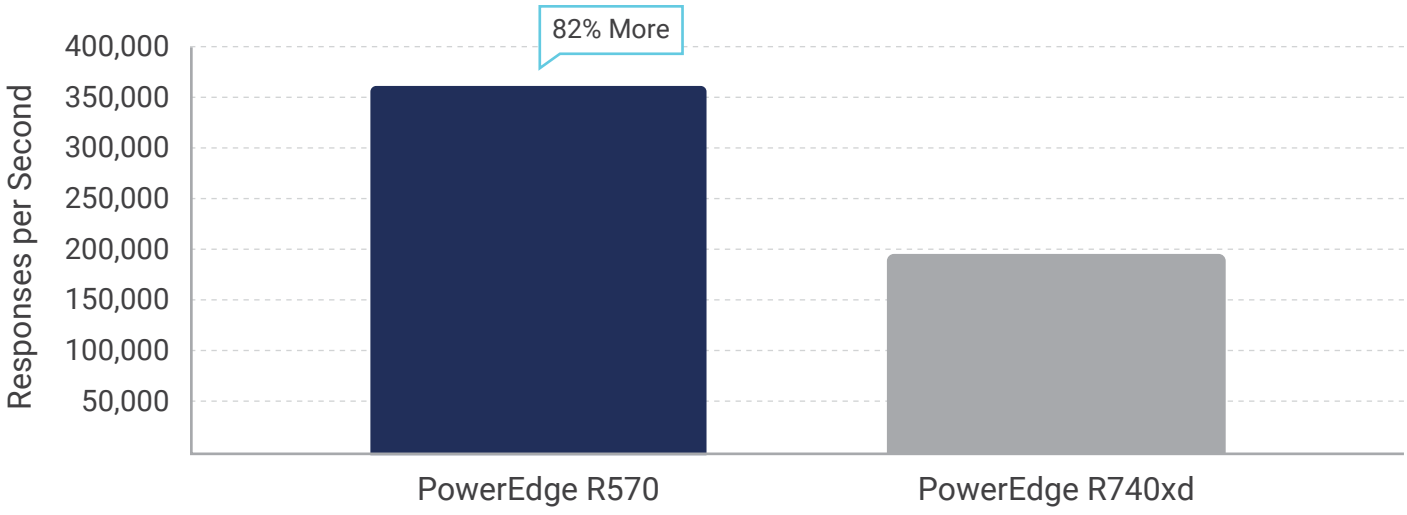

Figure 4 | RPS between the dual-socket Dell PowerEdge R740xd server and the PowerEdge R570 server in Apache JMeter at the matched 32-core comparison

Because the PowerEdge R570 server provides higher performance with a lower power-draw than the PowerEdge R740xd server for the same amount of work completed (314.05 W versus 331.14 W), it achieved 91% better performance per watt (see Figure 5).[2]

Relative Performance per Watt in Apache JMeter (higher is better)

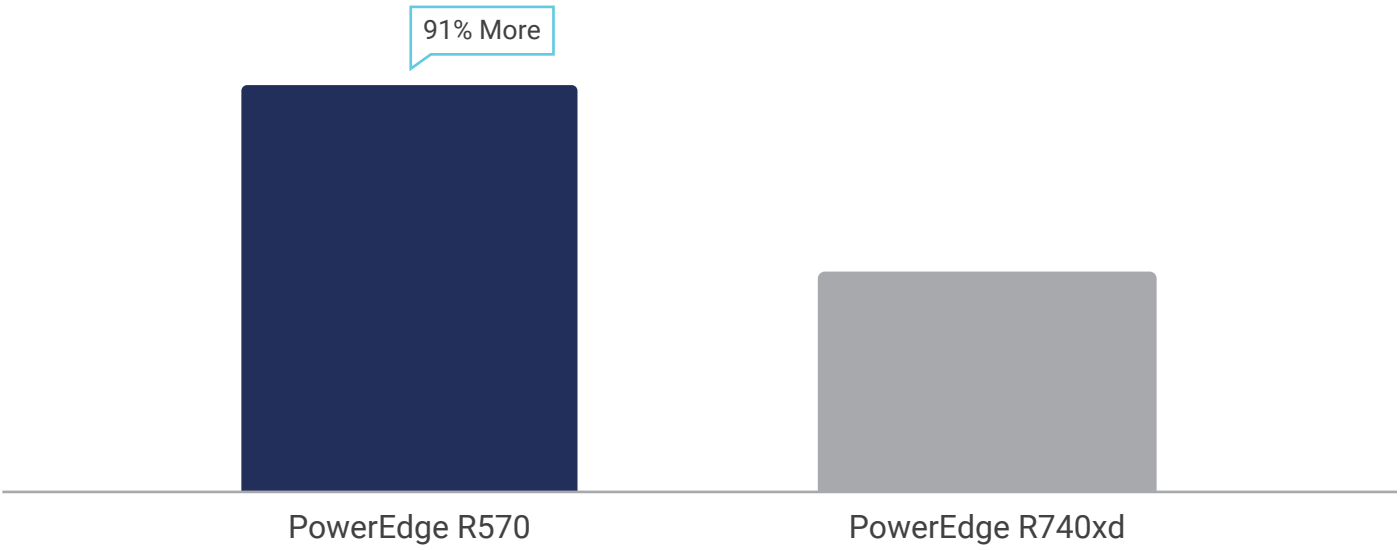

Figure 5 | Performance per watt between the dual-socket Dell PowerEdge R740xd server and the PowerEdge R570 server in Apache JMeter at the matched 32-core comparison

Overall, these results confirm that, compared to the dual-socket PowerEdge R740xd server, the single-socket PowerEdge R570 server provides:

- Lower latency and improved responsiveness under concurrent requests
- Lower power-draw and better performance per watt
- Greater workload concurrency and scalability
- More efficient infrastructure utilization for web-serving workloads

OpenResty® (NGINX and LUA®), Redis, and WRK®

The OpenResty caching workload, built on NGINX with integrated Lua scripting and paired with Redis, simulates how a business can evaluate the performance and scalability of a modern web caching layer under realistic application traffic. Combined with WRK, a high-performance HTTP benchmarking tool, OpenResty enables executing Lua logic directly within the NGINX event loop to accelerate request handling, while Redis provides ultra-low-latency in-memory storage for frequently accessed objects such as media files or session data. Together, these tools measure how efficiently a server can deliver cached content, process edge-side logic, handle concurrent requests, and sustain high throughput during peak demand on web servers.

Our testing simulated an SMB deploying an e-commerce caching tier to accelerate product page recommendations and improve load times. Using servers powered by Intel Xeon processors, we captured real-world behaviors such as cache hit rates, response times, Lua script execution, and Redis retrieval performance under load. Measuring CPU and memory utilization alongside these application-level metrics provides insight into how efficiently a modern single-socket server can consolidate caching and application acceleration compared to multiple legacy systems.

Results

In the media caching workloads, the PowerEdge R570 server outperformed the PowerEdge R740xd server in the equivalent configuration. For the OpenResty results, the single-socket PowerEdge R570 server had 2.65x more RPS (see Figure 6).

Responses per Second in OpenResty (higher is better)

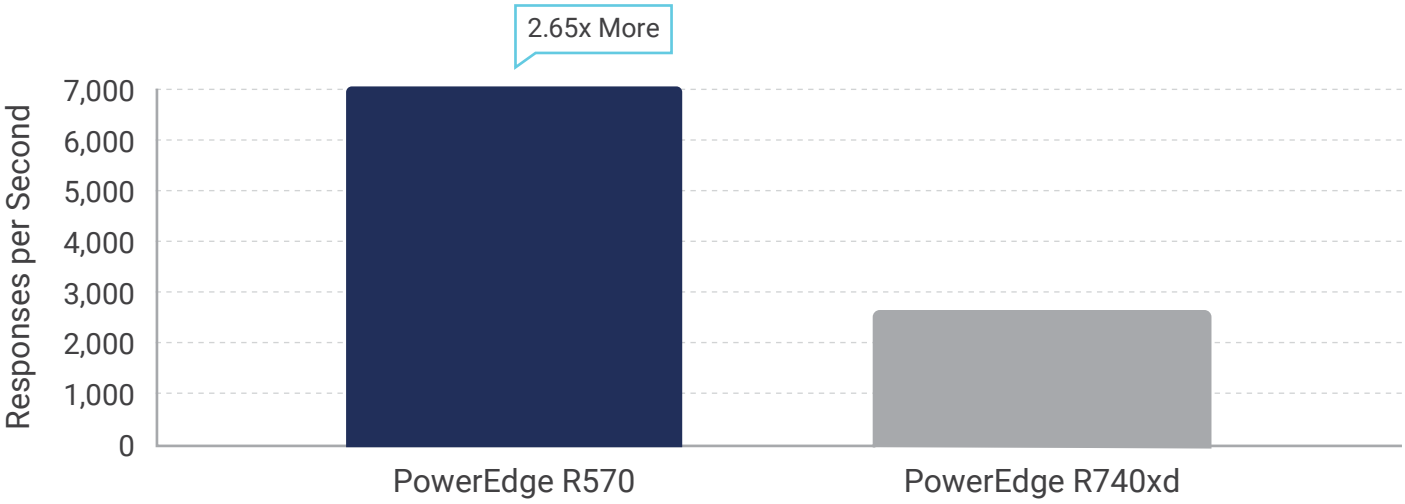

Figure 6 | Bar chart comparing RPS between the Dell PowerEdge R740xd server and the PowerEdge R570 server in OpenResty at the matched 32-core comparison

Because the PowerEdge R570 server could provide that much more performance than the dual-socket PowerEdge R740xd server for only an 11% greater power-draw (447.03 W versus 384.00 W), it achieved 2.32x greater performance per watt (see Figure 7).[3] These numbers continue to demonstrate the effectiveness of the PowerEdge R570 server, with 59% lower latency at the matched operating point.

Relative Performance per Watt in OpenResty (higher is better)

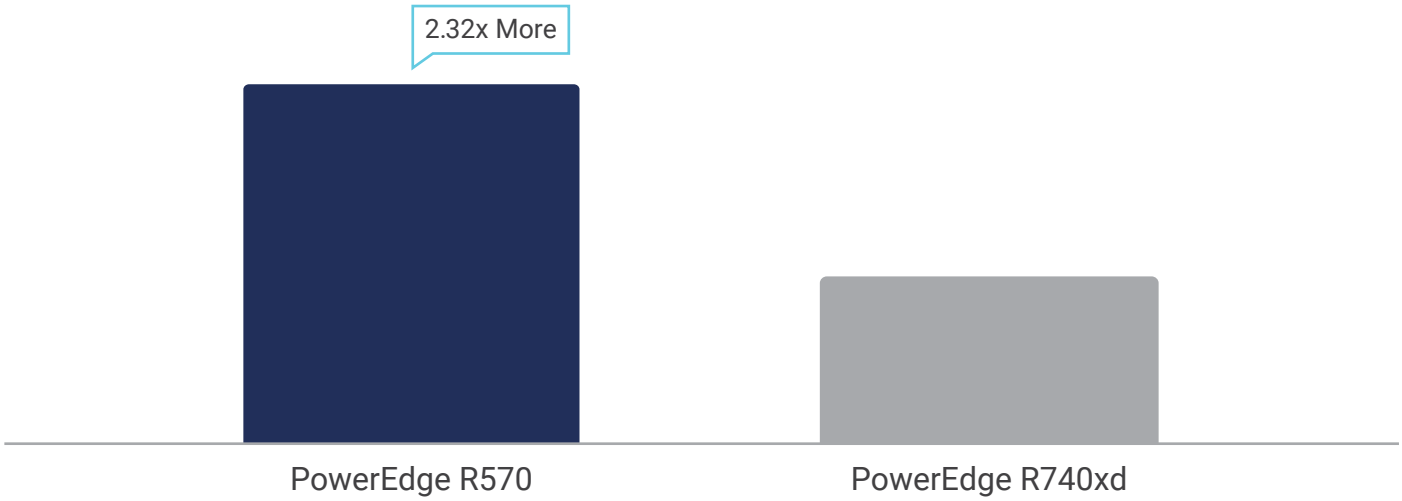

Figure 7 | Bar chart comparing performance per watt between the Dell PowerEdge R740xd server and the PowerEdge R570 server in OpenResty at the matched 32-core comparison

As the PowerEdge R570 server's higher throughput and significantly lower latency allow more media requests to be processed simultaneously, it delivers stronger results for organizations consolidating caching and content-delivery workloads onto single-socket servers. Our numbers indicate that the PowerEdge R570 server is far better suited to these types of workloads, and the improved responsiveness under realistic traffic conditions directly translates into more efficient use of system resources.

Overall, these results confirm that, compared to the dual-socket PowerEdge R740xd server, the single-socket PowerEdge R570 server provides:

- Higher usable performance for media caching
- Significantly lower latency and improved responsiveness under concurrent demand
- Superior performance per watt
- Greater workload concurrency and sustained scalability
- More efficient infrastructure utilization for media caching and content-delivery workloads

Business Relevance and Takeaways

Across all workloads, the metrics that we measured directly influence operational costs, user satisfaction, and scalability. Higher throughput, better performance per watt, and lower latency mean faster delivery of insights, content, and services, enabling businesses to remain competitive and responsive. Energy-efficient servers reduce power and cooling costs, while higher consolidation potential means better performance per workload.

Overall, we found that the single-socket PowerEdge R570 server, powered by Intel Xeon processors, provides a compelling upgrade path for businesses looking to modernize their infrastructures, consolidate their workloads, and optimize their operational efficiency. Compared to the dual-socket PowerEdge R740xd server's performance in equivalent configurations, the single-socket PowerEdge R570 server's performance gains across sentiment analysis, web serving, and media caching workloads demonstrate how architectural improvements can translate into measurable business value. For decision-makers, the choice between dual-socket and single-socket comes down to which can offer better performance, cost, and scalability while ensuring end users experience fast, reliable, and responsive services across multiple business-critical applications.

Explore

Visit the Intel website to further explore **Intel Xeon 6 processors**.

Visit the Dell Technologies site to see which **Dell PowerEdge server** is right for you.

Appendix: Configuration of Servers Tested

Model	Dell PowerEdge R570	Dell PowerEdge R740xd
CPU	1 x Intel Xeon 6787P (86 cores)	2 x Intel Xeon Gold 5217 (8 cores each)
Memory	16 x 96 GB (1,536 GB total) DDR5-5200, 2DPC	6 x 32 GB (192 GB total) DDR4-2666, 1DPC
Power	Power supply unit (PSU 1): 1,100 W PSU 2: 1,100 W	PSU 1: 750 W PSU 2: 750 W
Networking	Broadcom® 2P 100G QSFP 57508 OCP network interface controller (NIC)	Broadcom Adv. Dual 10GBASE-T Ethernet 25 Gb, 2P Intel Ethernet Network Adapter E810-XXV
Storage	8 x 1.92 TB Serial ATA (SATA)	8 x 893.75 GB SATA

Endnotes

[1]**LibriSpeech configuration: PowerEdge R570 server:** 15.71 simultaneous processes drawing 449.74 W; **PowerEdge R740xd server:** 6.63 simultaneous processes drawing 259.73 W. **Relative performance per watt:** (15.71 simultaneous processes/449.74 W)/(6.63 simultaneous processes/259.73 W) = (0.035 simultaneous processes/W)/(0.026 simultaneous processes/W) = 1.34x higher performance per watt (rounded down for analytical caution).

[2]**Apache Bench configuration:** 16 workers x 64 threads. **PowerEdge R570 server:** 356,769.76 RPS drawing 314.05 W; **PowerEdge R740xd server:** 195,990.25 RPS drawing 331.14 W. **Relative performance per watt:** (356,769.76 RPS/314.05 W)/(195,990.25 RPS/331.14 W) = (1,136.03 RPS/W)/(591.86 RPS/W) = 1.91 or 91% higher performance per watt (rounded down for analytical caution).

[3]**WRK load test configuration:** 4 threads at 200 simultaneous TCP connections. **PowerEdge R570 server:** 7,107.37 RPS drawing 447.03 W; **PowerEdge R740xd server:** 2,628.64 RPS drawing 384.00 W. **Relative performance per watt:** (7,107.37 RPS/447.03 W)/(2,628.64 RPS/384.00 W) = (15.90 RPS/W)/(6.85 RPS/W) = 2.32x higher relative performance per watt (rounded down for analytical caution).

0426/250110