

SINGLE SOCKET SERVERS ARE DISRUPTING THE SERVER MARKET

HPE PROLIANT DL325 GEN10 BASED ON THE SINGLE SOCKET AMD EPYC PROCESSOR DELIVERS 2P PERFORMANCE WITH 1P ECONOMICS

INTRODUCTION

Dual socket servers – servers that ship with two “sockets” to support two central processing units (CPUs) – are the mainstay of nearly every datacenter and server room. But walk through any server room or datacenter and slide three servers on their rails, pull off the cover and you will probably find at least one of those servers will be populated with only one CPU. And if you go to a console and run the performance monitor (PerfMon) utility against any server, you’ll more than likely see CPU utilization in the 25% - 35% range.

Advances in CPU microarchitectures and design have led to server platforms that can reach incredible performance levels. In many cases, compute power and resources are overmatched for the workloads supported. Even in large enterprise IT organizations, CPUs sit highly underutilized. This begs the question – why continue to deploy two socket servers if they are underutilized in your server room?

Moor Insights & Strategy sees an inflection point in the marketplace. We believe companies like Hewlett Packard Enterprise are responding to the trends in the server market by designing high-performing single socket servers that can address the needs of IT organizations of all sizes. This paper will examine the case for single socket servers and the HPE ProLiant DL325 Gen10 server.¹

HOW WE GOT HERE

When Advanced Micro Devices (AMD) announced the first x86 native dual-core processor in 2005, it signaled a shift in the industry. That new microarchitecture, combined with advances in process technology, allowed for silicon and server vendors to pack more compute power and resources into server platforms. Today’s servers pack up to 32 cores and 64 threads in a single CPU.

¹ For brevity and readability, hereafter referred to in this paper as the DL325.

As multicore and manufacturing process technology advanced, a second technology shift impacted server rooms and datacenters – virtualization. The benefit to companies was (and is) obvious – reduced physical server footprint and simplified management.

The benefits of virtualization are only partly realized by IT organizations. Application and workload resiliency is increased through live migration technologies. But the complexity of managing virtualized environments has offset the benefit of managing fewer physical servers. And while virtualization should increase datacenter operational efficiency, IT organizations are generally not seeing significant increases in server utilization rates.

Why do IT organizations continue to deploy highly underutilized, underpopulated two socket servers? In a simple word, resources. Single socket servers have traditionally shipped with lower performing CPUs married to lower memory and resource addressability (e.g. add-in cards and storage). Mid-market IT organizations looking to deploy virtualization in their organizations have traditionally been forced into “buying up” to achieve necessary performance levels.

MI&S sees the launch of servers such as the DL325 as an inflection point in the server industry. For the first time, enterprise class processor and memory performance, I/O expandability and security are available in a single socket package. The DL325 may be the ideal virtualization server for mid-sized and small enterprise IT organizations. Additionally, this server is well suited for workloads dependent on memory or I/O, such as software-defined storage (SDS).

TWO SOCKET PERFORMANCE, SINGLE SOCKET ECONOMICS

As previously described, in order for an IT organization to achieve enterprise-level virtualization performance² from a server, it would traditionally have to pay a two-socket premium. In other words, it would need to purchase a server that is designed and manufactured for a second processor in order to get the processor performance as well as memory and I/O expandability, regardless of how many CPUs populate that server.

As an enterprise class single socket rack mount server, the DL325 is tailored for very dense virtualized environments. This server boasts up to 32 multi-threaded cores, 8 memory channels, 2TB DDR4 RAM, 10 NVMe drives and 3 PCIe 3.0 slots. This configuration is richer than the average server populating today’s datacenter.

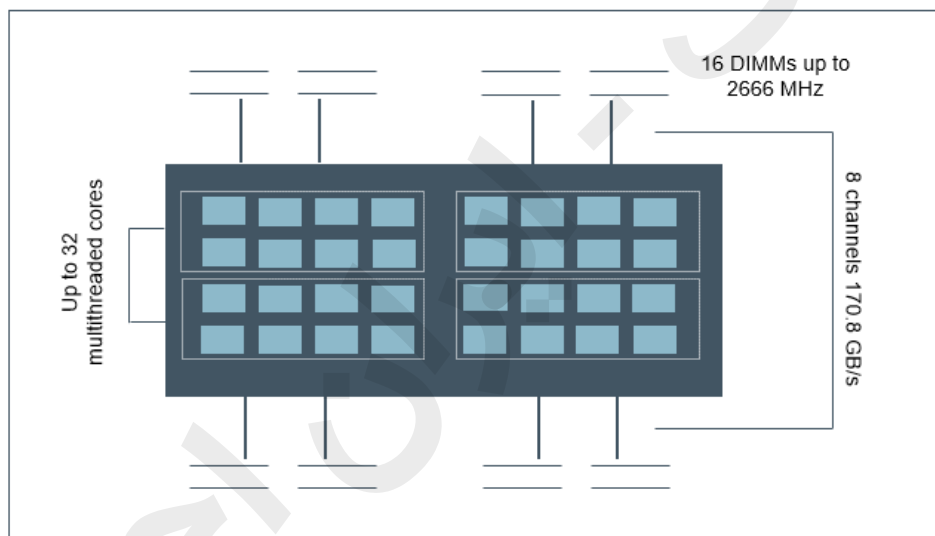
² MI&S sees virtualization performance as a combination of VM density *and* acceptable performance levels

As a result, IT organizations could enjoy performance similar to the majority of two-socket servers for certain workloads without paying for unused motherboard components.

AMD EPYC – THE FOUNDATION OF THE HPE PROLIANT DL325 GEN10

How is the DL325 able to achieve such levels of virtualization performance? It starts with the AMD EPYC System on a Chip (SoC). Its balanced architecture allows for more virtual machines to run more efficiently.

FIGURE 1: AMD EPYC SOC IS AT THE HEART OF THE HPE PROLIANT DL325 GEN10



Source: Moor Insights & Strategy

One of the key features of the EPYC SoC is its memory addressability. Over the past decade, the needs of virtualization have required more memory bandwidth and capacity. Specifically, the lack of sufficient memory channels inhibited full utilization of compute resources. Because of EPYC's eight memory channels, more bandwidth is available as virtual machines utilize compute resources. As a result, IT organizations should be able to achieve higher server utilization without sacrificing performance.

FIGURE 2: AMD EPYC SINGLE SOCKET PRODUCT LINEUP

AMD EPYC Model	Core Count	Memory Channels	Memory Bandwidth	I/O Bandwidth
7551P	32	8	2TB	128
7401P	24	8	2TB	128
7351P	16	8	2TB	128

Source: Moor Insights & Strategy

Note the fidelity of capabilities across the EPYC product offerings, as outlined in Figure 2. AMD has enabled rich functionality across all products. Memory addressability, capacity and expandability are fully enabled.

While we show the EPYC “single socket” lineup, it is important to note that IT organizations can outfit the DL325 server with AMD’s more (or less) performant CPUs for workload suitability. For example, an IT organization may want to populate their server with the fastest 32 core processor to support an analytics workload. This is easily achievable through HPE’s configuration tools.

HPE PROLIANT DEMOCRATIZES ENTERPRISE CLASS PERFORMANCE

MI&S looks at “enterprise class performance” as more than just a benchmark. While raw performance is important, IT organizations should consider the other features of a server that contribute to overall workload performance, manageability, reliability, availability and security (RAS).

The DL325 server seems well-positioned in every aspect.

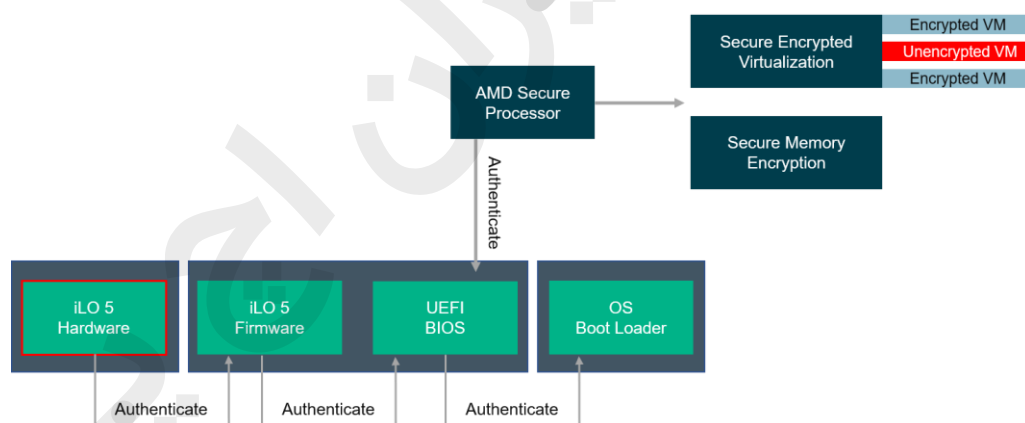
- Raw Performance – In published results, the DL325 platform sets the bar for single socket virtualization performance by scoring 1500@84 VMs in the SPECvirt_sc2013 benchmark.³ What does this mean to the mid-market IT organization? More virtual machines per physical server that perform faster.
- Workload optimization - HPE designed several features in the ProLiant portfolio to enable optimized workload performance.

³ Published by the Standard Performance Evaluation Corporation (SPEC) as of 6/5/2018. See <http://spec.org/> and <http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00047709enw> for more information.

One of the more interesting features, Workload Matching, allows for IT managers to select workload profiles in system BIOS, which will automatically configure the DL325 for optimal performance. These features allow IT managers to accelerate server deployments with less manual set-up and maintenance.

- Manageability – HPE OneView accelerates infrastructure deployment with software-defined templates and increases IT efficiency by reducing maintenance time and support costs. Remote systems management is also key to IT operational efficiency. The DL325 server platform ships with integrated lights out 5 (iLO 5) management standard, with licenses available to unlock additional feature.
- Security - Security is one of the more important features in the DL325 server. MI&S has written extensively on the HPE Silicon Root of Trust capabilities. When integrating this technology with AMD's Security Processor, organizations benefit from the greatest levels of security for the datacenter.

FIGURE 3: HPE SILICON ROOT OF TRUST – AMD SECURITY HANDSHAKE



Source: Moor Insights & Strategy

HPE Silicon Root of Trust begins as iLO 5 is powered on and an immutable fingerprint in silicon is used to verify the iLO 5 firmware before the server can boot. At each point during the boot process – from firmware loads to BIOS to OS – firmware authenticity is verified. If any discrepancy is noted, administrators can automate the reprovisioning of the server to a last known good image; or the server can be taken off-line for further examination.

WHERE THE HPE PROLIANT DL325 GEN10 SHINES

There are countless use cases where organizations could benefit from deploying the DL325. The ratio of cores – to memory – to I/O makes it an attractive fit for mid-market IT organizations and service providers architecting scale-out compute environments. Some of the more interesting use cases are detailed below.

MID-MARKET VIRTUALIZED INFRASTRUCTURE

There are several ways to measure VM performance in an IT organization - VM density, VM “raw” performance, price per VM, etc. Regardless, the DL325 is in a leadership position. Its 32 multi-threaded cores and up to 2TBs of memory allow for maximum VM density. In just released benchmarks on SPEC.org, the DL325 set the bar for single socket virtualization platforms with a score of 1500@84 VMs.

I/O INTENSIVE WORKLOADS

The richness of PCIe Gen3 lanes translates into richness of storage and richness of expandability. The DL325 ships with three PCIe slots that can be used for accelerators such as Field Programmable Grid Arrays (FPGAs) and Application Specific Integrated Circuits (ASICs), enabling greater performance for workloads like data analytics. It is important to note, the DL325 does not support GPU accelerators at the time of publication (6/5/2018).

SOFTWARE-DEFINED STORAGE

The 128 PCIe lanes in the AMD EPYC architecture can translate into a powerful SDS platform from HPE. HPE took those 128 PCIe lanes and developed a server that supports up to 10 NVMe drives, embedded 4x1GB NICs and 3 PCIe 3.0 slots. Tying this storage and networking with up to 2TB of DDR4 RAM and 32 multi-threaded cores creates a platform that capitalizes on performance and capacity.

TCO SAVINGS FOR THE MID-MARKET

Total cost of ownership (TCO) is usually a meaningless term for any IT organization not deploying thousands of servers. While a detailed TCO analysis that fully considers all elements of the TCO equation was not conducted, we believe mid-market organizations deploying the DL325 could realize significant savings in capital expenditures (CapEx) and Operational Expenditures (OpEx). These savings can be derived from server bill of materials (BOM) costs and software licensing (depending on licensing models).

For general IT infrastructure workloads, the cost savings on the DL325 do not come with a performance penalty; in fact, quite the opposite is true. When measuring performance of a server platform, there are several benchmarking organizations to reference. The gold standard is [SPEC](#) (Standard Performance Evaluation Corporation). In recent submitted benchmarks from HPE, the DL325 more than held its own:

- [SPEC CPU2017](#) is a suite of tests that examines the performance of server platforms at performing different tasks, based on specific processors. SPECrate 2017 Integer and SPECrate 2017 Floating Point measure work over a period of time. These two specific tests are good proxies of CPU performance. A DL325 with a single EPYC processor drew equivalent performance to a third-party server outfitted with two price comparable competitive platforms.

FIGURE 4: HPE PROLIANT DL325 GEN10 PERFORMANCE

Server	CPU	Qty	SPECrate2017_int_base ⁴	Price/SPECrate2017_int_base ^{5,6}	SPECrate2017_fp_base ⁴	Price/SPECrate2017_fp_base ^{5,6}
HPE ProLiant DL325 Gen 10	AMD EPYC 7551P	1	126	\$85.63	123	\$87.72
Competitive Server Platform	Intel Xeon Gold 5118	2	119	\$114.03	126	\$107.69

Source: Moor Insights & Strategy

Note the performance and price performance of the two servers. The single socket DL325 is delivering equivalent integer and floating point performance of competitive platform. And it does so with a significant advantage in cost. As mentioned previously, this translates into real cost savings for IT organizations of all sizes.

- [SPEC virt_sc2013](#) is a test that measures overall virtualization performance of server platforms. As previously mentioned, the DL325 shows impressive performance numbers, scoring [1500@84 VMs](#) loaded on a single AMD EPYC 7551P CPU. This places the DL325 in the leadership position for single socket virtualization.

⁴ Published SPEC scores can be found at <https://www.spec.org/cpu2017/results/>, <http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00047709enw>, and <http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00048319enw>

⁵ DL325 pricing provided by HPE on 29 May 2018

⁶ Competitive server platform is the Dell EMC PowerEdge R440 server, outfitted with two Intel Xeon Gold 5118 CPUs. Pricing for the configuration tested in SPEC CPU2017 was obtained from Dell.com on 2 June 2018 and does not include discounts.

DOUBLE CLICKING ON SOFTWARE LICENSING

Software licensing is fluid and varies from vendor to vendor - per core, per virtual core, per socket, by term or perpetual. However, organizations deploying software applications licensed per socket can potentially reduce this operational expense by deploying the DL325. For example, VMware vSphere Enterprise is licensed “per socket”. In the traditional datacenter, this workload would reside on a fully populated two-socket server. The ability to deploy vSphere on a single socket server should reduce licensing costs by half. This could amount to thousands of dollars saved for the average mid-market IT organization.

CALL TO ACTION

The server market is in constant change, and the launch of the DL325 server marks yet another inflection point – the emergence of the single socket server with enterprise class performance.

The design of the DL325 makes it ideal for virtualization and I/O intensive workloads such as Software-Defined Storage. Mid-market IT organizations that plan to embark on new virtualization projects should strongly consider the DL325 as the cornerstone to those projects. Mid-market organizations that have already virtualized should consider migrating their virtualized workloads to the DL325 server.

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTOR

Matt Kimball, Senior Analyst at [Moor Insights & Strategy](#)

PUBLISHER

Patrick Moorhead, Founder, President, & Principal Analyst at [Moor Insights & Strategy](#)

INQUIRIES

Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

CITATIONS

This paper can be cited by accredited press and analysts but must be cited in-context, displaying author's name, author's title, and "Moor Insights & Strategy". Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

LICENSING

This document, including any supporting materials, is owned by Moor Insights & Strategy. This publication may not be reproduced, distributed, or shared in any form without Moor Insights & Strategy's prior written permission.

DISCLOSURES

This paper was commissioned by Hewlett Packard Enterprise (HPE). Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.

©2018 Moor Insights & Strategy. Company and product names are used for informational purposes only and may be trademarks of their respective owners.