AN EVALUATION OF THE OPEN COMPUTE MODULAR HARDWARE SPECIFICATION

SUMMARY

The Open Compute Project's Modular Hardware System (OCP-MHS or simply MHS) sub-project under the Server Project Group has one mission: interoperability among key elements of datacenter, edge, and enterprise infrastructure. It achieves this by creating common standards for the physical, signaling, and protocol interfaces of server components, making it easier for datacenter architects to build and integrate datacenter infrastructure.

Of the three MHS projects, the datacenter MHS (DC-MHS) is particularly interesting because it significantly impacts major server vendors servicing both the hyperscale and enterprise server market segments. This project focuses on delivering a modular server hardware specification that enables hardware vendors to more easily and quickly source and manufacture the server infrastructure that powers the datacenter.

The progression of DC-MHS is noteworthy, given the accelerated pace of semiconductor and hardware innovation in response to the AI explosion. This Moor Insights & Strategy (MI&S) research brief will explore three key areas:

1. The history of OCP and how the DC-MHS might be the realization of a vision laid out more than 10 years ago
2. How the DC-MHS delivers real value to enterprise IT organizations by embracing a cloud operating model at a time when access to the latest hardware technology is critical to keep pace in the marketplace
3. Dell's and Intel's roles in the DC-MHS and Dell's recent release of DC-MHS-compliant servers

OCP—A SHORT HISTORY OF INNOVATION

Historically, there has been a chasm between enterprise and hyperscale datacenters. Enterprise IT organizations have long placed a premium on the quality of design and manufacturing delivered by original equipment manufacturers (OEMs) such as Dell and HPE. While these servers came at a slightly higher price, their quality, performance, manageability, and security capabilities were unmatched.

Hyperscalers, by contrast, have traditionally consumed infrastructure from original design manufacturers (ODMs) to reduce capital expenses. Though these servers tend to lag in terms of quality, manageability, resilience, and security, hyperscale datacenter operators have been able to mitigate these challenges by using homegrown toolchains and to match performance through the sheer quantity of ODM servers deployed.

Further, the hyperscalers' approach to deploying finely tuned focused workloads and monolithic environments has allowed them to influence server designs to meet their specific needs (i.e., semi-customized platforms that are environment-specific).

There are considerably more ODM vendors than OEM vendors. ODMs also lacked any real shared standards to address the challenges of deploying, securing, and managing infrastructure, which was exacerbated by the need for bespoke infrastructure for each hyperscaler. As a result, Meta (then Facebook) founded OCP in 2011 to deliver a common set of standards around server design for the hyperscale datacenter. It was the exclusive domain of hyperscalers, with a clear focus on compute.

Over time, OCP evolved to consider the needs of the entire hyperscale datacenter—from networking and storage to manageability and security. Yet it is still exclusively the domain of hyperscale datacenters; enterprise IT organizations have continued to shy away from OCP as OEMs have not fully embraced the movement. Despite this lack of enterprise adoption, one can see the (enterprise) appeal of servers that are designed and manufactured with the speed and variety of the ODM market but with the quality, security, and management toolchain of the OEM community.

OCP-MHS—Like LEGO Bricks For Servers

One reason OEMs had been hesitant to embrace OCP was the perceived lack of value for them; there was a prevailing belief that adopting OCP standards would undercut OEM value-adds for security, management, and performance. In other words, why would an OEM embrace OCP if it rendered servers as mere commodities?

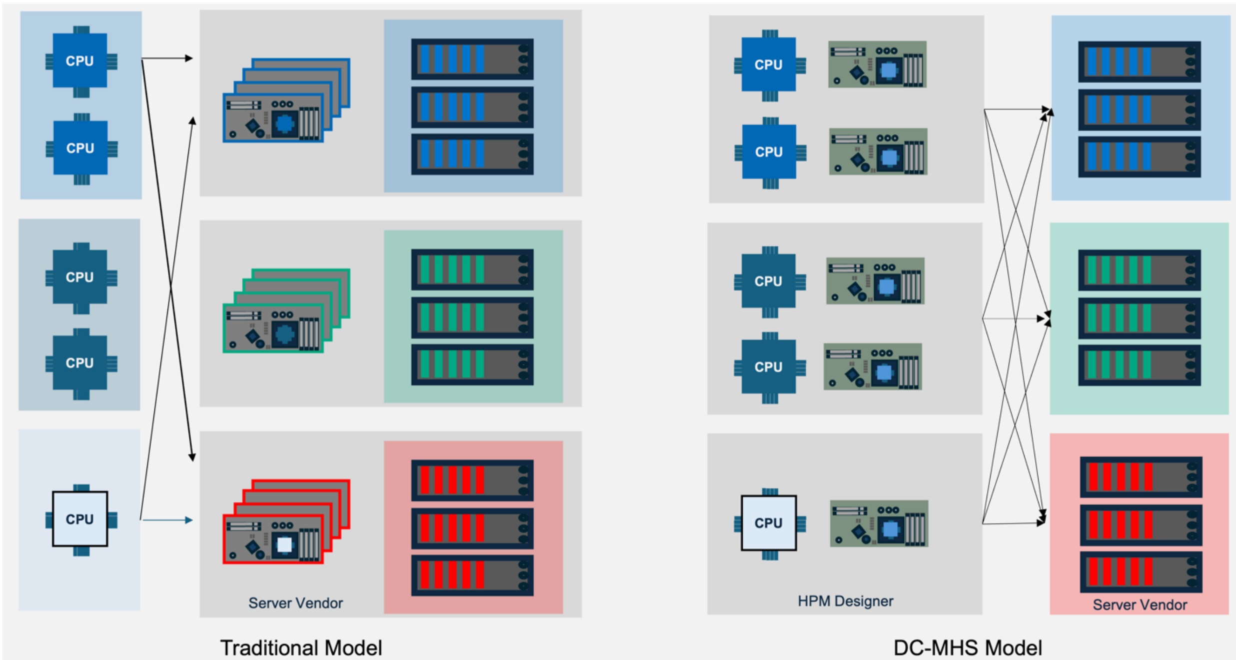
Countering this belief, MHS is the enabler that allows OEMs to increase the rate of innovation through standardization of foundational elements of a server's design—specifically, the physical and logical interfaces that connect components. The thinking is that if the industry can coalesce around a simpler manufacturing process, OEMs might increase the pace of innovation, turning out differentiated platforms faster. This is why the OCP Server Project Group formed the MHS sub-project, resulting in the DC-MHS revision 1.0 specification approved in 2022.

To better understand how DC-MHS R1 drives efficiency in the infrastructure development lifecycle, it is worth exploring the legacy approach to developing a server platform. When a CPU vendor such as Intel designs a new CPU, such as its latest Intel Xeon 6 processor with E-cores, it builds reference platforms and works with hardware partners such as Dell to build specific server models.

As part of this process, Intel will invest non-recurring engineering (NRE) budget to build each Dell platform. The portfolio that results can grow significantly for a company like Dell that develops infrastructure optimized for different workloads and environments. For example, as part of adopting Intel Xeon 6 processors with E-cores and P-cores (efficiency cores and performance cores), Dell launched the PowerEdge R670 and R770 server platforms to support traditional virtualization as well as AI/HPC-like workloads that span the edge, the enterprise datacenter, and the cloud.

When Intel wants to activate other server vendors, the exercise is repeated, which implies more reference platforms, NRE investments, and platform designs that require excessive time and money. But what if Intel could create a single host processor module (HPM) per CPU that could be used by all its server vendors? An HPM that could be leveraged to build out a range of server platforms and form factors to meet its full range of deployment and workload requirements? This is what DC-MHS R1 aims to achieve.

FIGURE 1: OPTIMIZING SERVER DESIGN WITH DC-MHS



DC-MHS streamlines the server design process, enabling different CPU designs to work without customization across a variety of servers from different vendors.
Source: Moor Insights & Strategy

The DC-MHS R1 specification includes several workstreams that account for the entire platform, from form factors and physical mountings to the power supply and system connectivity. The completeness of the specification, covering all of the physical connectivity aspects, becomes apparent.

A good way to think of MHS is like Lego bricks, the plastic building toys. MHS looks at each module like a Lego brick. Those bricks may vary by manufacturer, but their physical and electrical interfaces are consistent. This common set of standards enables manufacturers and their suppliers to drive efficiency and speed.

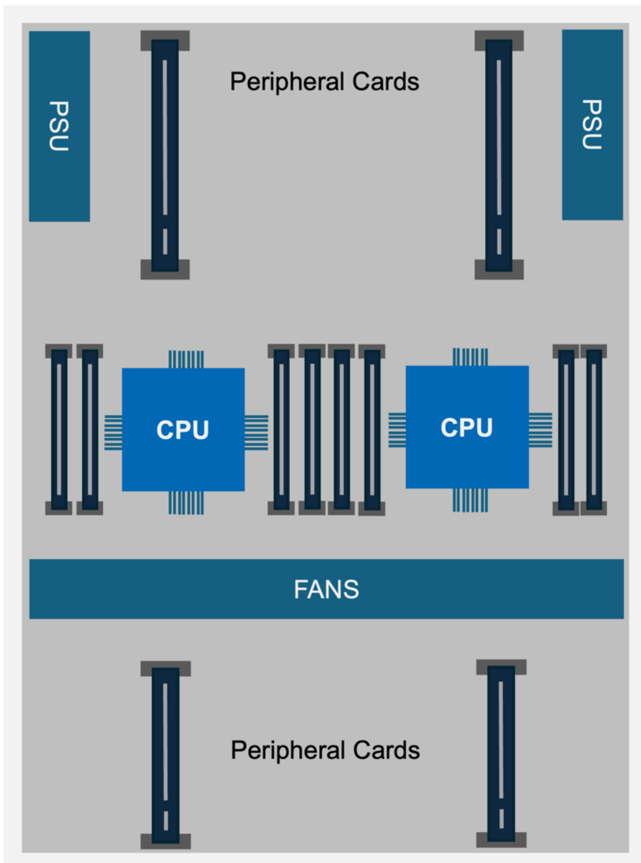
HOST PROCESSOR MODULE DETAILED VIEW

The HPM workstream details the requirements for the physical layout of a server motherboard. It enables chip makers and their partners to create a single motherboard design that can be deployed in any server chassis for any server vendor that adheres to the OCP-MHS standard. In turn, server vendors can quickly take HPMs and, along with the supporting modules previously described design server technology in a more standard way. Think of these as the blueprints for manufacturing the basic server.

This simplification of building out the fundamental elements of server technology enables server vendors such as Dell to focus on adding value to these platforms for specific deployment scenarios, workloads, and the like. Effectively, embracing the MHS standards collapses the timeline from concept to product, meaning that the latest technology is in the racks of all datacenters—large and small—much faster.

Because HPM addresses the entirety of an organization's data estate, multiple form factors support different space requirements. The traditional server that sits in a standard server rack (HPM-FLW, for "full width") supports a 1U and 2U height with up to two CPU sockets. As mentioned, this specification covers the physical and electrical interfaces as well as a base level of functionality.

FIGURE 2: THE HPM FULL-WIDTH FORM FACTOR DESIGN



Enterprise EIA-310 Specification
Source: Moor Insights & Strategy

The image shows a standard HPM implementation. However, if a vendor wanted to add value to its HPM solution (e.g., Dell's Boot Optimized Storage Solution for cost-effective hardware RAID), it would simply work with the HPM provider to integrate this functionality. This would be known as an "HPM with value-add modifications."

HPM also supports dense form factors (HPM-DNO). This specification has been created to support a variety of partial-depth and -width HPMs to support diverse types of server deployments. Whether the customer is a telco or an enterprise deploying edge infrastructure, the HPM-DNO spec enables server vendors to achieve high levels of agility and speed in offering workload- and scenario-based platforms.

The completeness of the HPM spec is significant because it enables server vendors to more quickly and easily meet the needs of all customers. This is especially true for enterprise and commercial datacenters that lack access to the cutting-edge technology enjoyed by hyperscale datacenters deploying hundreds of thousands of servers.

How MHS Benefits the Market

For most IT professionals, the servers running in racks in the enterprise datacenter are not top of mind—unless, of course, something goes wrong. How each server is designed and manufactured is even further down the list of day-to-day considerations. However, the ability to acquire and deploy a server that employs the latest CPU, accelerator, memory, and I/O innovations is top of mind. This has been even more noticeable as AI has shined a spotlight on the supply-chain issues that have caused chip and server manufacturers to struggle to meet the demands of the server market.

In the server market in particular, this dynamic has led many IT organizations to either deploy less-optimized infrastructure to support their needs or look to less-traditional server manufacturers for AI platforms. While these platforms from ODMs may be quicker to obtain, they are often considered lacking in the quality, resilience, security, and manageability that protect the enterprise.

For hyperscale and cloud datacenter operators, the dynamic is a little different. As previously mentioned, although access to more leading-edge silicon and server technologies may be required, these organizations have come to rely on ODMs as their primary server suppliers.

Regardless of the customer segment, MHS addresses this challenge by enabling server vendors to innovate and deliver new server technologies to the market faster by removing many of the mechanical and electrical design challenges associated with laying out a unique server motherboard and chassis. As an example, Dell, a server vendor that has embraced MHS, will be able to refresh a new generation of its PowerEdge portfolio in a fraction of the time it once would have. This means customers receive new server technology faster, with equal or even greater assurances of quality. Dell is also able to focus on innovation that delivers measurable value, and leverage industry-standard components where applicable.

Exploring Dell's MHS Portfolio

Dell is the first major server vendor deploying MHS-compliant platforms. The first two server platforms in its MHS portfolio, the Dell PowerEdge R670 and R770, are designed with specific components for cloud service providers (CSPs) and enterprise IT organizations tasked with simplifying the lifecycle management of servers—from deployment to provisioning to servicing.

The R670 is a 1U, two-socket server supporting a richness of memory for virtualized and containerized workloads. The R770 is a 2U, two-socket server that supports up to 2TB DDR-5 RAM and 4x16 Gen5 PCIe slots. MI&S sees it as focused on accelerated workload support.

The R670 and R770 servers' DC-MHS architecture supports the DC-SCM (secure control module) for hosting the iDRAC10 integrated Dell Remote Access Controller. The iDRAC10 contains Dell IP that is built on top of OpenBMC, which is based on an open-source firmware stack for baseboard management controllers (BMCs). iDRAC10 enables core, enterprise, and datacenter operators to manage their heterogeneous server environments centrally.

In the case of the R670 and R770, iDRAC10 automates the entire server lifecycle, including deployment, configuration, monitoring, and updates. Further, iDRAC10 is an innovative agent-free architecture that ensures consistent management and offers comprehensive high-speed Telemetry Streaming. iDRAC10 features PowerEdge's Cyber Resilient Architecture that uses an immutable silicon root of trust (RoT) to authenticate firmware booting and updates.

While Dell has not publicly commented on how MHS impacts its full server portfolio, MI&S believes the company should (and will) consider adopting this HPM standard across its entire portfolio. By further embracing HPM, Dell can respond to the needs of the market in terms of time-to-value far more quickly. Further, this gives Dell first-mover status in an OEM server market that generally takes years to move from generation to generation.

DELL AND INTEL LEAD THE OPEN STANDARDS

As Dell is the first major OEM to support DC-MHS, Intel is the first CPU vendor to embrace this specification. The Intel Xeon 6 processor with E-cores is the underlying CPU architecture that populates the R670 and R770 server platforms.

When Intel launched its Xeon 6 family of processors, the company delivered two distinct performance bands—the Xeon 6 processor with P-cores for highly performant workloads (codenamed Granite Rapids) and the Xeon 6 processor with E-cores for traditional virtualized and highly dense platforms (codenamed Sierra Forest). Given the nature of CSPs and enterprise customers deploying cloud-like infrastructure, the first CPU in the Dell R670 and R770 is the Xeon 6 processor with E-cores. This CPU delivers the density that enables customers to maximize rack-space efficiency without sacrificing performance. While Intel refers to it as “little cores,” MI&S believes the Xeon 6 with E-cores can support the non-high-performance workloads that many enterprises deploy to the cloud. Indeed, Xeon 6 with E-cores ships with a number of integrated acceleration engines that enable the offload of CPU-intensive tasks such as data streaming and in-memory analytics.

Intel Xeon 6 with E-cores delivers integrated security through its software guard extensions (Intel SGX) and trusted domain extensions (Intel TDX) security solutions. While Intel SGX supports applications and data by creating a trusted execution environment at the process level, Intel TDX isolates virtual machines (VMs) from the hypervisor, host OS, and other VMs. This is critically important for multi-tenant environments such as the cloud.

In early 2025, Intel will release its Xeon 6900-series processor with E-cores. This version will include all of the processor’s existing capabilities, with a doubling of the core count—up to 288 cores. This will be the highest x86 core count available on the market and should deliver significant value to the CSP and enterprise market.

Finally, although it has not been publicly discussed by Dell or Intel, it is not a stretch to think that we may see the Intel Xeon 6 processor with P-cores supported in the R670 and R770 servers. This performance-focused CPU is optimized for the data-driven workloads that are increasingly important as generative AI moves from proof of concept projects to production in the enterprise. The Intel AMX acceleration engine combined with highly performant cores drives extreme performance across the AI data pipeline.

MHS Implications for the Market

It is important to consider the market impact of MHS on both the supplier and customer levels. Looking at this from the supplier's perspective, MHS should bring significant benefits if it is fully embraced. Simplifying the fundamental and time-consuming processes of laying out motherboards and defining interfaces will cut considerable costs and time in the platform design process. In turn, this can enable server vendors to accelerate their timelines for delivering highly optimized, highly tuned platforms to meet emerging workloads and deployment scenarios. This is evidenced by the speed at which Dell developed its Xeon 6 processor with E-cores-based R670 and R770 platforms. If the company adopted this approach across the entirety of its portfolio, one can imagine the efficiencies that would be gained in the long term.

From the customer's perspective—and it's important to separate hyperscalers from enterprises here—a fully enabled and aligned supply chain adhering to the MHS specification would mean that access to the latest chip, memory, and I/O technologies happens faster than ever. In concept, the result should be that the most optimized and efficient servers will populate datacenter racks—all deployed, managed, and secured centrally.

MI&S regards the MHS specification as enabling enterprise IT organizations to achieve hyperscale-like capabilities in the datacenter without introducing risk or potentially higher TCO. However, this can be achieved only if the server market follows Dell and Intel's lead.

Bringing Value to Both CSPs and Enterprise IT

For more than a decade, OCP has worked to deliver on behalf of the hyperscale community through the creation of standards for the hardware and software ecosystems. Indeed, its purview has evolved from providing standards only for server designs to standards for deploying, managing, and securing infrastructure across the data estate—from the datacenter to the edge and beyond.

As the cloud operating model is employed to support a business environment moving at breakneck speed —and powered by workloads with unique and complex computational needs—the rate of hardware innovation has not kept pace. As a result, we have seen CSPs and enterprise IT organizations deploy servers and other infrastructure of lower quality to meet short-term needs. This is a dynamic that is introduced and reinforced with every new technology trend.

Recognition of these shortcomings has spurred OCP-MHS and the HPM specification. Through the HPM spec, server infrastructure vendors can turn concepts into products faster and more reliably. For CSPs, this can mean the ability to stand up a differentiated service faster than the competition. And for enterprise IT, this delivers the complete cloud operating model—from silicon and hardware to software stacks—to companies of all sizes.

MI&S sees the obvious appeal of MHS for the CSP community. In many ways, this is an improvement on a delivery model that has continually evolved for many years. Although this is valuable, it is not novel.

However, we also see considerable value for the enterprise IT community, even if OCP-MHS isn't recognized by all of those companies as a body and specification. The ability to more readily consume highly optimized, highly secure technology from the likes of Intel through server platforms as new chip, memory, and I/O advancements are introduced into the market is invaluable, especially as enterprise IT increasingly evolves into a service provider to the business units it supports.

Dell and Intel enjoy first-mover status in this movement and should benefit from delivering the R670 and R770 platforms on Xeon 6 processor with E-cores. While we expect others to follow, the leadership and experience gained by moving first should enable both companies to extend this position while the market plays catch-up.

For more information on the Dell R670 and R770 platforms, please follow this link.
For more information on Intel Xeon 6 processors, please follow this link.

Important Information About This Paper

Contributor
Matt Kimball, Vice President and Principal Analyst, Datacenter Compute and Storage

Publisher
Patrick Moorhead, CEO, Founder and Chief Analyst at Moor Insights & Strategy

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

Citations
Accredited press and analysts can cite this paper, but it must be cited in context, displaying the author's name, title, and “Moor Insights & Strategy.” Non-press and non-analysts must receive prior written permission from 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 without Moor Insights & Strategy's prior written permission.

Disclosures
Dell Technologies commissioned this paper. Moor Insights & Strategy provides research, analysis, advice, and consulting to many of the high-tech companies mentioned in this paper. No employees at the firm hold equity positions with any of the 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 regarding such information’s accuracy, completeness, or adequacy and shall have no liability for errors, omissions, or inadequacies. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The views expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators, not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on the future, it 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 this document’s publication date. Please remember 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.