Business white paper

A new standard for CAE and AI solutions in manufacturing

Accelerating product development and enhancing competitiveness using HPE systems powered by 4th and 5th Gen AMD EPYC processors and AMD Instinct accelerators

Contents

Manufacturers must deliver ever more complex products, get to market faster, and continuously innovate and improve product quality with limited resources.

Design teams frequently run over thousands of simulated scenarios, putting enormous stress on the HPC infrastructure.

Meeting manufacturing challenges with advanced simulation methods

As manufacturers of all sizes struggle with cost and competitive pressures, and as products become more complex and capable, advanced simulation methods using computer-aided engineering (CAE) tools have become essential to getting quality products to market faster.

Advanced simulation can help manufacturers better understand and test their products by creating detailed digital prototypes. By simulating the behavior of these prototypes before they go into production or into customers' hands, manufacturers can identify potential problems early in the process. This helps reduce risk, save money and time, and improve product quality. Additionally, CAE tools can be used to quickly develop cost-effective designs that meet customer requirements, considering factors such as materials selection, safety regulations, and overall cost.

CAE tools can also help drive innovation while mitigating risks and costs associated with potential product failures, such as warranty, recall costs, and potential litigation. Figure 1 shows how a CAE solution supporting multiple engineering disciplines can help address manufacturing business challenges

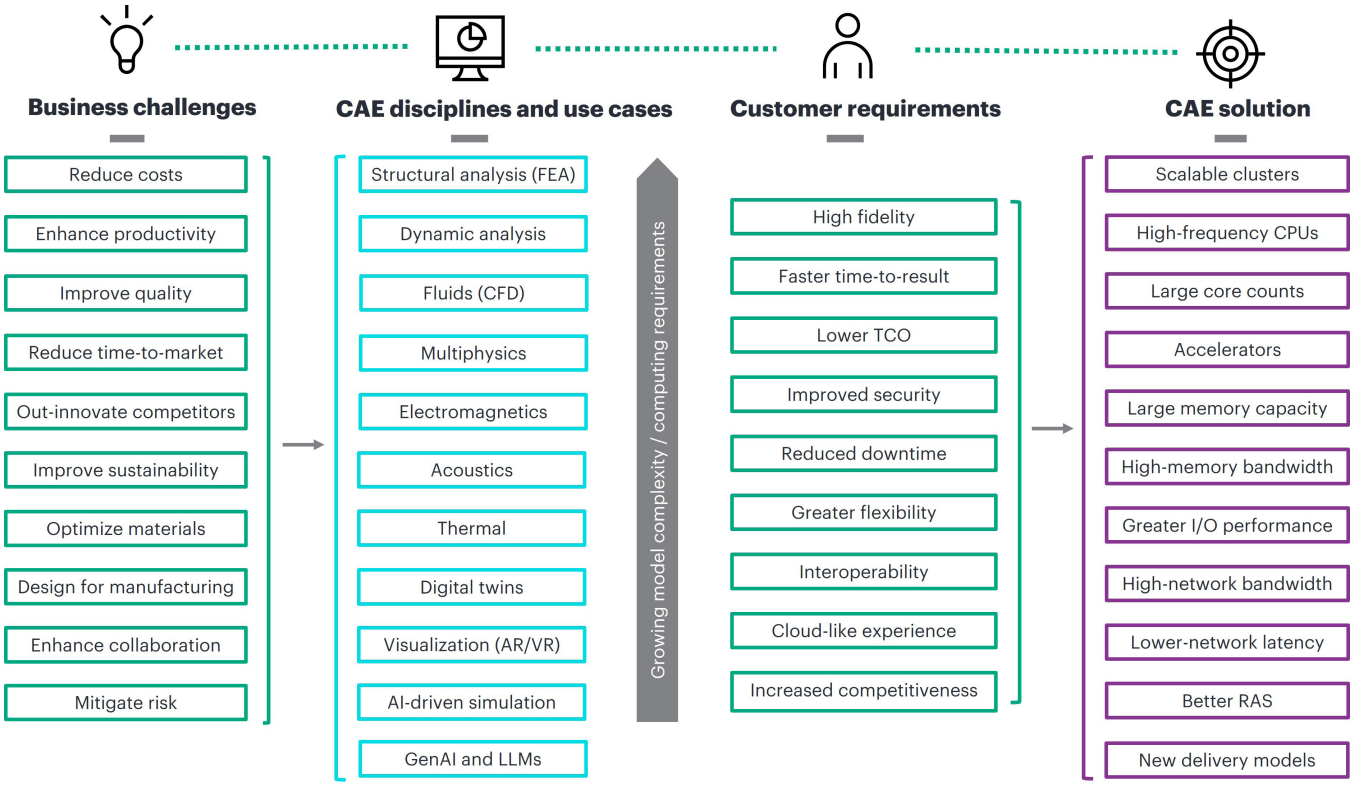


Figure 1. Manufacturing business challenges and customer requirements addressed with CAE

HPE is the market leader in HPC systems, with a 31.5% market share.[1]

Customer requirements for CAE simulation environments

Advanced simulation tools are used across multiple industries from automotive manufacturing to aerospace to consumer products to industrial manufacturing. Engineers rely on CAE tools for applications that include virtual testing for aerodynamics, assessing the durability of materials, ensuring interoperability between parts and components, and optimizing product quality.

To meet these and other challenges, manufacturers need a reliable, secure high performance computing (HPC) environment that can help them:

— Achieve faster time to results for large simulations
— Support larger, more complex models
— Accommodate a wider variety of advanced simulation workloads
— Reduced complexity and total cost of ownership (TCO)

If this was not already challenging enough, the need for high performance hardware and storage infrastructure continues to grow. Several factors are driving an increasing need for advanced computer-based simulation, including:

1. **Materials optimization**—The use of more exotic materials and honeycombed structures to reduce the weight and cost of designs
2. **Sustainability**—Demand for greener eco-friendly products, electric/hybrid vehicles, and higher levels of product and component recyclability and reuse
3. **Compliance**—The need to comply with various industry regulations related to product safety, emissions, and fuel-efficiency standards that vary by jurisdiction
4. **Internet of Things (IoT)**—Smarter, more connected products incorporating a variety of sensors
5. **Digital twins**—The use of virtual representations of physical objects that mimic their real-world counterparts

[1] "Hyperion Research HPC Market Update", Hyperion Research, page 10, November 2022

For example, manufacturers must allocate significantly more resources to simulating battery technologies in electric and hybrid vehicles to ensure optimal performance and safety. With the proliferation of sensors and smart connected devices in modern designs, engineers face new requirements related to spectrum management. They also need to worry more about electromagnetic compatibility (EMC) and electromagnetic interference (EMI).

As models grow in complexity, memory and available cache have emerged as essential requirements for the processors and servers that power advanced simulation. Models are frequently comprised of millions of elements or cells. The more of a model that fits in CPU cache or memory, the faster a simulation can be performed. Engineering solvers typically perform calculations in parallel, exploiting modern multicore processor designs to help maximize throughput and get results more quickly.

Of the advanced CAE disciplines and use cases described in Figure 1, multiphysics is among the most demanding as it involves simultaneous simulation of different aspects of a physical system. Multiphysics combines multiple CAE disciplines, including structural analysis, fluid mechanics, mechanical dynamics, and electromagnetics. Engineers frequently use iterative techniques such as design of experiments (DOE) to explore the relationship between multiple design parameters to optimize complex systems.

AI-augmented CAE simulation

Increasingly, manufacturers are leveraging machine learning (ML) and artificial intelligence (AI) to improve the efficiency of their CAE simulations. By collecting and learning from prior simulation data, AI-enabled CAE tools can generate models to rapidly test design alternatives without the computational constraints and complexities involved with traditional approaches. AI-augmented simulations can significantly boost productivity during design explorations involving high-fidelity simulations. For example, engineers can accelerate various types of simulations, such as automotive crash tests, external aerodynamics, battery thermal management, and antenna array systems.

AI features are increasingly embedded directly into the products offered by leading CAE independent software vendors (ISVs), helping engineers take advantage of AI-augmented simulation without the need for coding or deep data science skills. Examples of AI-augmented tools include Ansys® SimAI™, Altair® DesignAI™, and Altair HyperStudy®. As AI-augmented simulation, also known as intelligent simulation, becomes increasingly prevalent in modern environments, new generative AI (GenAI) techniques are anticipated to further enhance efficiency and competitiveness in manufacturing operations.

GenAI in manufacturing

GenAII has the potential GenAI has the potential to improve productivity in multiple areas. These include product development, product planning and procurement, and various aspects of production. Specific examples of where GenAI can help include creating new product concepts and engineering drawings to reduce prototyping, predicting product-market fit with qualitative consumer/market data, and adjusting production orders in real time based on sensors and order-tracking data. As CAE design centers tend to be centers of excellence for managing HPC infrastructure, many new AI-driven applications supporting these use cases will likely find their way into converged HPC and AI data centers.

To meet these evolving requirements, manufacturers are increasingly looking for servers with:

- High-frequency processors with high core counts to complete more simulations faster
- Large memory capacity, high memory bandwidth, and large amounts of L3 cache to help maximize throughput
- Accelerators for both engineering simulation and AI model training
- High I/O performance for storage and network connections
- Low network latency and high network bandwidth to enable parallel simulations and model training
- Reliability, availability, and serviceability (RAS) to help maximize productivity and minimize downtime costs

Advances in HPC and AI bring new challenges

Besides raw CPU performance, energy efficiency is another important consideration for manufacturers. To reduce their carbon footprint and meet sustainability goals, data center managers need servers that deliver maximum throughput per watt to help minimize power and cooling requirements. They may also need dense, energy-efficient designs that help minimize cooling and data center space requirements.

As the market leader in HPC systems with a 31.5% market share,[2] Hewlett Packard Enterprise delivers one of the industry's most comprehensive CAE solution portfolios across compute, interconnect, software, storage, and services delivered on-premises, hybrid, or as a service. By teaming up with AMD, HPE offers a comprehensive ecosystem that provides exceptional performance, flexibility, and choice across the full range of CAE and AI applications.

[2] "Hyperion Research ISC24 Market Update," May 2024

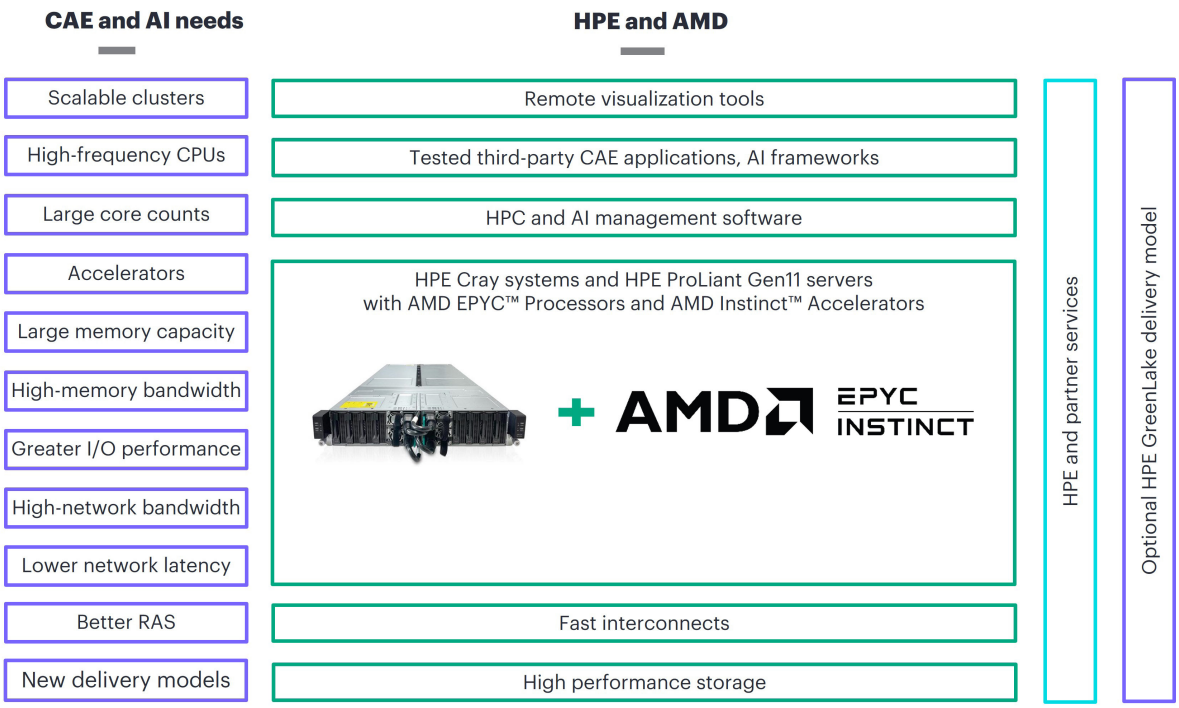


Figure 2. A high-level view of the HPE and AMD CAE solution portfolio

HPE has you covered with multiple deployment options for manufacturers, from on-premises infrastructure to colocation to private and hybrid clouds.

The HPE and AMD solution for CAE

HPE provides a complete end-to-end solution stack, shown in Figure 2, that is flexible and customizable to meet nearly any manufacturing clients' business requirements. This stack provides an extensive portfolio of supported and tested third-party CAE application software. It also encompasses various system software, server, and storage platforms, delivered with high-value HPE and partner services and remote visualization tools to support CAE and AI workloads.

Most large CAE teams run their advanced engineering simulations in-house using on-premises infrastructure. This can be advantageous for several reasons. It is generally cost-efficient for organizations with sustained workloads that can keep their infrastructure highly utilized. It also helps ensure that sensitive intellectual property is protected and that organizations can manage costs through their product's lifecycle.

For organizations that prefer colocation solutions or hybrid cloud delivery models, HPE has you covered. The same powerful CAE solution can optionally be deployed on HPE GreenLake for customers who prefer a managed service with a flexible cloud delivery model. These deployment options provide manufacturers with flexibility as their requirements evolve. Components of the HPE and AMD solution for CAE are described as follows:

Remote visualization: Enhances security by keeping critical data within the data center; boosts productivity and collaboration with any time, any place / any location access to graphic-intensive models; helps lower costs by centralization, improving system manageability, and helping optimize software and hardware resource utilization; and promotes retention of highly skilled staff with better work-life balance and location flexibility. HPE offers a variety of remote visualization and virtual desktop infrastructure (VDI) solutions built using the latest 4th and 5th Gen AMD EPYC™ processors.[3] For CAE customers with demanding visualization requirements, NICE Desktop Cloud Visualization (DCV) provides efficient and optimized remote access to graphic-intensive 3D applications, including all the major CAE pre and postprocessing software.[4]

[3] "Elevate your virtual desktop experience with HPE ProLiant and AMD," HPE, 2024

[4] "NICE Software Desktop Cloud Visualization and EnginFrame.," HPE, 2021 DCV and EnginFrame are offered by NICE software, now part of Amazon Web Services.

CAE applications: HPE and AMD have strong partnerships with multiple ISVs. HPE and AMD have computer scientists and application engineers who help ISVs test and optimize their applications for HPE servers and AMD processors. Major CAE applications supported and tested include Altair® Radioss®, Altair® AcuSolve®; Ansys® Fluent®, Ansys® Mechanical™, Ansys® LS-DYNA®, and Ansys® CFX®; MSC® Nastran®; Siemens® STAR-CCM+®; OpenFOAM®; Dassault Systèmes SIMULIA® Abaqus® FEA; and ESI® Virtual Performance Solutions® (VPS) to name just a few.

AI frameworks and data science tools: While some design centers will leverage the AI-augmented simulation capabilities offered by CAE tool vendors, others may need to develop the capacity to train models in-house. For example, applications such as advanced driver assistance (ADAS) and autonomous vehicle (AV) systems require enormous amounts of training data (either collected by sensors or obtained through virtual drive simulations) to refine models continuously and help ensure that these features perform flawlessly in response to adverse weather conditions, road hazards, and various other driving situations. Manufacturers can use high performance HPE infrastructure powered by AMD processors and accelerators to run advanced ADAS/AV software solutions for training and simulation. They can also run a wide range of other AI model training frameworks, including TensorFlow™, PyTorch™, or the HPE Machine Learning Development Environment Software.

HPC architecture: 4th and 5th Gen AMD EPYC processors provide industry performance leadership for many of the most challenging workloads in CAE. For example, based on AMD testing on reference platforms, a 2P 64-core 5th Gen AMD EPYC 9575F-based system (128 total cores) yields approximately 1.57x the composite arithmetic mean performance compared to the latest generation of alternative processors running select tests with Ansys Fluent v241.[5] Systems with superior performance can help manufacturers improve simulation throughput and engineering productivity.

HPC and AI system software: HPE offers CAE customers a complete and modular software portfolio consisting of HPE developed software solutions combined with best-of-breed solutions from business partners and open-source software providers. This portfolio is validated, integrated, and performance-enhanced by HPE so that manufacturers can select the right software mix for their CAE applications—all from one source, including operating system, cluster management, job schedulers and resource managers, HPC tools and libraries, and more.

Optimized libraries: AMD Optimizing CPU Libraries (AOCL) are a set of numerical libraries optimized for the AMD EPYC processor family. They have a simple interface that takes advantage of the latest hardware innovations to accelerate the development and performance of CAE applications. These libraries can take advantage of hardware-level features such as AVX-512 instructions, available in the latest 4th and 5th Gen AMD EPYC processors, to accelerate CAE ISV applications and in-house developed code able to use these extended instructions.

AMD ROCm™: AMD ROCm is an open software stack with programming tools, compilers, libraries, and runtimes for software development on select AMD GPUs and APUs. The extensive set of AMD ROCm libraries makes it easy for manufacturers with advanced computing requirements to get accelerated applications deployed quickly. At the heart of AMD ROCm, is the Heterogeneous-computing Interface for Portability (HIP)—an API and runtime environment that enables developers to create applications that are portable across GPUs from multiple manufacturers.[6] HIP makes it straightforward for developers to adapt NVIDIA® CUDA code to run on the AMD Instinct accelerators. Because AMD ROCm is freely available and open source, developers have quickly embraced the platform, leading to an extensive ecosystem of AI models and CAE applications optimized for AMD-powered servers.[7]

HPC compute nodes: HPE provides multiple high performance server options to act as HPC compute nodes depending on customer requirements. For most commercial CAE environments, HPE Cray XD2000 systems or HPE ProLiant Gen11 servers will be good choices, as they easily fit into existing data center environments. Supercomputing facilities requiring the highest performance levels in 100% liquid-cooled configurations may be interested in the HPE Cray EX2500 and HPE Cray EX4000 supercomputers.

[5] amd.com/en/legal/claims/epyc.html#q=9xx5-033

[6] For details on the HIP see the AMD ROCm Software documentation.

[7] See github.com/ROCm/ROCm.

HPE servers and systems suitable for CAE simulation and AI workloads are shown in Table 1.

Table 1. AMD-powered systems and servers for CAE environments

Server family	HPE ProLiant Gen 11	HPE Cray XD2000	HPE Cray XD675	HPE ProLiant Compute XD685	HPE Cray EX2500/ EX4000
Servers	HPE ProLiant DL325 Gen11 Server HPE ProLiant DL365 Gen11 Server HPE ProLiant DL385 Gen11 Server	HPE Cray XD225v Server HPE Cray XD295v Server	HPE Cray XD665 Server HPE Cray XD675 Server	HPE ProLiant Compute XD685	HPE Cray EX425 Blade HPE Cray EX4252 Blade HPE Cray EX235n Blade HPE Cray EX235a Blade HPE Cray EX255a Blade
Processor technology	4th, 5th Gen AMD EPYC CPUs	4th, 5th Gen AMD EPYC CPUs	4th Gen AMD EPYC CPUs	5th Gen AMD EPYC CPUs	3rd, 4th Gen AMD EPYC CPUs
Accelerator technology	NVIDIA or AMD Instinct GPUs (select models)	NVIDIA or AMD Instinct GPUs	NVIDIA or AMD Instinct GPUs[8]	AMD Instinct GPUs	NVIDIA or AMD Instinct GPUs/APUs
Packaging	Standard 1/2U rackmount servers	Dense, multiserver chassis	Dense GPU server chassis	Dense GPU server chassis	Dense, multiserver chassis, blade-based architecture
Management	HPE iLO 6, DMTF Redfish HPE Converged infrastructure management software	DMTF Redfish HPE Performance Cluster Manager	DMTF Redfish HPE Performance Cluster Manager	HPE iLO HPE Performance Cluster Manager	HPE Cray System Management HPE Performance Cluster Manager
Interconnects	Ethernet, InfiniBand, HPE Slingshot	Ethernet, InfiniBand, HPE Slingshot	Ethernet, InfiniBand, HPE Slingshot	Ethernet, InfiniBand	HPE Slingshot
Cooling options	Air cooling or direct liquid cooling (DLC)	Air cooling or DLC	Air cooling or DLC	Air cooling or DLC	Liquid cooling through cooling distribution units[9]
Ideal for	Traditional air-cooled high performance server building blocks	Dense HPC data center environments that require the latest high performance processors and accelerators	Dense HPC data center environments with multiple GPUs for CAE apps or AI model training and inference	Dense HPC data center environments with multiple GPUs for CAE apps or AI model training and inference	HPC centers that require the highest levels of supercomputing performance

[8] The 4U HPE Cray XD665 supports NVIDIA H100 GPUs only. The 8U HPE Cray XD675 supports AMD Instinct MI300X OAM GPUs only.

[9] In the HPE Cray EX2500, parts of the system may be air-cooled in specific circumstances. "HPE Cray Supercomputing EX QuickSpecs," HPE, 2024

HPE ProLiant Gen11 servers: For CAE customers who prefer 1U, single-processor systems, the HPE ProLiant DL325 Gen11 Server is an excellent solution. This server has modest power and cooling requirements and fits easily into most data center environments. The HPE ProLiant DL365 Gen11, a 1U dual-processor server, and the HPE ProLiant DL385 Gen11, a 2U dual-processor server, are excellent choices for CAE workloads that require substantial memory capacity.

HPE ProLiant servers

HPE ProLiant DL325 Gen11 Server

- 1U 1P server
- Single 4th or 5th Gen AMD EPYC processor, up to 160 cores[10]
- Up to 3 TB DDR5, CXL 1.1 supported
- Up to 4 LFF HDD/SSD; SAS/SATA or 10 SFF HDD/SSD; SAS/SATA/NVMe;
- Up to 20 EDSFF E3.S; NVMe

HPE ProLiant DL325 Gen11 Server front view

HPE ProLiant DL365 Gen11 Server

- 1U 2P server
- Up to 2x 4th or 5th Gen AMD EPYC processors with up to 160 cores[11]
- Up to 6 TB DDR5, CXL 1.1 supported
- Up to 10 SFF HDD/SSD; SAS/SATA/NVMe; up to 20 EDSFF E3.S
- Up to 2 x16 PCIe Gen5 slots; up to 2 x8 OCP3.0 slots (upgradable to x16)

HPE ProLiant DL365 Gen11 Server front view

HPE ProLiant DL385 Gen11 Server

- 2U 2P server
- Up to 2x 4th or 5th Gen AMD EPYC processors with up to 160 cores[12]
- Up to 6 TB DDR5, CXL 1.1 supported
- Up to 12 LFF HDD/SSD; SAS/SATA; Up to 24 SFF HDD/SSD; SAS/SATA/NVMe; Up to 36 EDSFF E3.S; up to 48 SFF HDD/SDD
- Up to 8 PCIe Gen5 slots; up to 2 x8 OCP3.0 slots (upgradable to x16)

HPE ProLiant DL385 Gen11 Server front view

Figure 3. HPE ProLiant servers for CAE workloads

The HPE ProLiant DL325, HPE ProLiant DL365, and HPE ProLiant DL385 Gen11 Servers pictured in Figure 3 support the latest 4th or 5th Gen AMD EPYC processors. For the most compute-intensive CAE applications, HPE recommends that customers use 5th Gen AMD EPYC processors with between 32 and 64 cores per socket. These processors exhibit excellent performance even at higher core counts because they support 12 independent memory channels. HPE ProLiant systems running 5th Gen AMD EPYC processors require the latest 6400 MT/s DIMMs, offering up to 25% greater memory bandwidth than 4800 MT/s DIMMs.[13]

[10, 11, 12] HPE plans to release HPE ProLiant servers with 160 core 5th Gen AMD EPYC 9845 processors in early 2025.

[13] HPE ProLiant DL3x5 servers with 5th Gen AMD EPYC processors require 6400 speed-rated DIMMs but operate at 6000 MT/s, 25% faster than 4800 MT/s DIMMs supported by 4th Gen AMD EPYC processors.

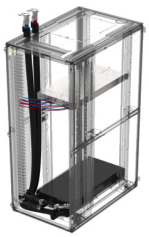
Customers running these servers have their choice of operating environments. Supported environments include Red Hat® Enterprise Linux® (RHEL), SUSE Linux Enterprise Server (SLES), and Windows Server.[14] HPE ProLiant embedded server management, HPE iLO, enables you to securely configure, monitor, and update your servers from anywhere.

HPE Cray XD2000 system: The HPE Cray XD2000 system is a dense, multiserver platform that packs exceptional performance and workload flexibility into a small data center space while delivering the efficiencies of a shared infrastructure. It is an excellent choice for midsized or larger CAE environments that require performance, scalability, density, serviceability, and ease of management.

Each HPE Cray XD2000 2U chassis supports up to four HPE Cray XD225v 1U Servers supporting either 4th or 5th Gen AMD EPYC processors. Alternatively, each chassis can support two half-width 2U 2P HPE Cray XD295v Servers. Both servers can be serviced without impacting the operation of other servers in the same chassis, increasing server availability. The HPE Cray XD2000 delivers up to 4x the density of a traditional rackmount 2U server in standard racks and provides rear-aisle serviceability access.[15]

HPE Cray XD2000 system
XD225v (1U) & XD295v (2U) specifications

Density-optimized HPE Cray XD2000 chassis supporting up to 4x HPE Cray XD225v 1U Servers or 2x HPE Cray XD295v 2U Servers



HPE data center rack with optional DLC

Processor	Up to two 4th or 5th Gen AMD EPYC processors—up to 160 cores / 320 threads per socket[16]
OS support	Windows, RHEL, SLES, Ubuntu, HPE Cray OS, VMware®
Memory	12 channels DDR5 per CPU @ 4800/6000 MT/s (up to 24 total DIMMs for 2P)[17]
Nodes	1U/2P half-wide; up to 4 per chassis 2U/2P half-wide (GPU support); up to 2 per chassis
PCIe lanes	65 PCIe5 lanes with 4-link XGMI + 6 bonus PCIe3 lanes per CPU
Storage	Supports up to 2x NVMe SSDs per node (4 for 2U) and/or M.2 2280/22110
Expansion slot	2x PCIe Gen5 x16 slot 1x PCIe Gen5 x16 MCIO slot 1x PCIe Gen5 x16 MCIO cable slot 1x PCIe Gen3 x4 M.2 Mezz
Power supply (Hot plug)	CRPS PSU support 2400W 240V
Fans	4x 4056 fans per node
Cooling	Air-cooled or optional DLC

Figure 4. HPE Cray XD2000 systems for CAE workloads.

[14] Operating systems and virtualization software support for HPE ProLiant servers are designed for seamless integration with partner operating systems and virtualization software. By collaborating closely with our partners, we ensure that their products are optimized, certified, and fully supported within your HPE server environment. Access the certified and supported servers for each of the operating system and virtualization software: HPE Servers Support & Certification Matrices—hpe.com/us/en/collaterals/collateral.a50010841enw.html

[15] The HPE Cray XD2000 system supports four HPE Cray XD225v 1U Servers in two rack units, whereas some servers require 2 rack units for a single server. "HPE Cray XD2000 QuickSpecs," HPE, 2024

[16] HPE plans to release HPE Cray XD2000 systems with 160 core 5th Gen AMD EPYC 9845 processors in 2025.

[17] The HPE Cray XD225v and HPE Cray XD295v Servers with 5th Gen AMD EPYC processors require 6400 MT/s rated DIMMs but operate at 6000 MT/s.

HPE Cray XD2000 systems offer a complete, scalable solution for HPC customers. They feature flexible power and cooling options, including DLC, delivering superior performance while reducing TCO. Up to 20 HPE Cray XD2000 chassis can be installed in either 42U or 48U HPE standard racks, delivering up to 80 servers and 160x 4th or 5th Gen AMD EPYC processors per data center rack, subject to power and cooling considerations.

Customers who wish to increase compute density can deploy HPE Cray XD2000 systems with up to four 1U HPE Cray XD225v Servers per chassis. For customers running accelerated workloads, up to two HPE Cray XD295v Servers can be installed in each chassis supporting the latest GPU accelerators, including NVIDIA H100 PCIe, NVIDIA H100 NVL, NVIDIA L40s PCIe, or AMD Instinct MI210 PCIe accelerators.

Manufacturers can also use HPE Performance Cluster Manager, a fully integrated system management solution offering all the functionality needed to manage HPE Cray XD2000 clustered servers.[18] HPE Performance Cluster Manager aggregates system management and remote management information from the standard DMTF Redfish API supported on the HPE Cray XD2000. The software also provides system setup, hardware monitoring and management, image management and updates, power management, and integrations with ISV and open-source software solutions.

Tied into the silicon root of trust from HPE is the AMD Secure Processor, a dedicated security processor embedded in the AMD EPYC system on a chip (SoC). These features, along with secure recovery and firmware runtime validation of HPE at startup, help limit security breaches and system disruption if system code is compromised.

HPE Cray XD675 Servers and HPE ProLiant Compute XD685: The HPE Cray XD675 Server is an excellent solution for GPU-intensive deep learning and GenAI workloads. Built using the same technologies that power the world's most powerful and efficient supercomputers, it supports 4th Gen EPYC CPUs with multiple GPUs and accelerated CAE applications or large AI model training. Details about the HPE Cray XD675 Server are provided as follows.

[18] "HPE Performance Cluster Manager QuickSpecs," HPE, 2024

HPE Cray XD675 Server
(8U) server specification.

HPE Cray XD675 8U system chassis front view with 2x 4th Gen AMD EPYC CPUs and up to 8x AMD MI300X OAM GPUs, multiple storage options

Processor	Up to two 4th Gen AMD EPYC processors—up to 96 cores / 320 threads per socket
Accelerator	8x AMD Instinct MI300X OAM GPUs
OS support	Windows, RHEL, SLES, Ubuntu, HPE Cray OS, VMware
Memory	Support for up to 24x DDR5 5600 MT/s DIMMs
Nodes	8U server nodes installed in HPE Cray XD6500 rack.
PCIe lanes	65 PCIe5 lanes with 4-link XGMI + 6 bonus PCIe3 lanes per CPU
Storage	Supports up to 2x NVMe SSDs per node (4 for 2U) and/or M.2 2280/22110
Expansion slot	5x full-height, half-length PCIe Gen5.0 2x OCP 3.0 expansion slot with embedded
Power supply (Hot plug)	3000W Platinum Kit (6x 3000W + 2x 2400W) and 3000W Titanium Kit (6x 3000W + 2x 2700W)
Fans	12x 801105 fans per node
Cooling	Air-cooled or optional DLC

Figure 5. HPE Cray XD675 Server for CAE workloads

For customers that require the highest levels of accelerated performance, the 5U HPE ProLiant Compute XD685 Server is a similar GPU-dense system that supports the latest 5th Gen AMD processors and your choice of AMD Instinct MI300X or MI325X OAM GPUs.

Storage and networking: Fast I/O is critical for CAE and AI workloads to help ensure that file and network I/O do not become bottlenecks during simulation. HPE Cray XD2000 systems, HPE Cray XD675 Servers, HPE ProLiant Compute XD685 Servers, and HPE ProLiant Gen11 servers offer PCIe® 5.0, providing twice the throughput of the previous generation PCIe Gen4.[19] HPE offers a variety of high performance PCIe options, including 200 Gbps HDR and NDR InfiniBand adapters,[20] multiport 100GbE adapters, and high performance NVMe SSDs. Customers can optionally use 200 Gbps HPE Slingshot Ethernet adapters for added network performance. Multiple storage options are available inside the chassis, ranging from 0 to 24 (SFF) SAS/SATA drives.

For storage requirements that exceed available capacity inside the chassis, HPE offers two options for a shared external parallel file system for CAE and AI clusters, depending on the customers' preference and scalability requirements:

— Cray ClusterStor E1000 Storage systems—A rack-scale HPE storage product available in all-flash, disk, or hybrid configurations that embeds the open-source Lustre® file system with full enterprise support for both hardware and software from HPE.

— HPE Cray Storage Systems C500—An entry-level converged version of the Cray ClusterStor E1000 Storage system that delivers impressive performance and is ideal for entry and midrange HPC/AI clusters in CAE environments with limited storage budgets. This unique, hybrid architecture exploits the strengths of different storage media (NVMe, SSD/SAS, HDD), enabling superb performance and capacity starting in a 3U rack configuration.

Services: HPE and its partners offer a spectrum of services to meet manufacturing CAE requirements—from application tuning to more integrated advisory service offerings such as project and cluster management, on-site consulting, technical account management, and solution architecture consulting.

Optional direct liquid cooling

CAE simulation and AI are among the most compute-intensive applications. Not surprisingly, with the rapid growth of AI, data center power consumption, and associated energy costs continue to increase. According to the International Energy Agency (IEA) Electricity 2024 report, data centers globally used 2% of all electricity in 2022, and the IEA predicts that energy requirements could double by 2026.[21] Many data centers are not equipped to support the cooling requirements of modern CPUs and GPUs.

For customers with suitably equipped data centers, HPE Cray XD2000 systems offer plug-and-play support for DLC. The DLC option allows customers to increase power density and data center efficiency.[22] HPE server racks connect directly to facility water supplies without the need for secondary plumbing. Options are available for CPU only or CPU plus memory cooling. While air cooling is fine for most applications using the latest AMD EPYC processors described in this document for specific dense configurations involving AMD EPYC CPUs with per-socket TDPs of 300 watts or more, HPE may recommend the HPE Cray XD2000 system with the DLC option.[23]

[19] PCIe 5.0 supports up to 32 GT/sec. PCIe 4.0 supports up to 16 GT/sec.

[20] HPE InfiniBand adapters are based on standard Mellanox ConnectX-6 technology.

[21] "Electricity 2024 - Analysis and forecast to 2026," IEA, 2024

[22] "Liquid cooling: a cool approach for AI," HPE, August 2024

[23] "HPE Cray XD2000 QuickSpecs," HPE, 2024

A four-node HPE Cray XD2000 system powered by 5th Gen AMD EPYC processors delivered up to 1.25x the throughput of an equivalent system powered by 4th Gen AMD EPYC processors running a mix of Ansys CAE workloads.[24]

Industry-leading performance

HPE Cray XD2000 systems and HPE ProLiant Gen 11 servers can give manufacturers a competitive advantage. In October 2024, HPE undertook a comprehensive series of internal tests involving popular CAE workloads. The tests evaluated an HPE XD2000 system powered by the latest 64-core 5th Gen AMD EPYC 9555 processors compared to a similarly configured system powered by 4th Gen AMD EPYC 9554 processors.[25]

Four different Ansys applications (Ansys LS-DYNA, Ansys Fluent, Ansys CFX, and Ansys Mechanical), each involving multiple models, were tested.[26] The average performance improvement was calculated for each application across all the models run.

The results are compelling, as shown in Figure 6. For the explicit finite element analysis (FEA) workloads running Ansys LS-DYNA, the latest HPE Cray XD2000 systems powered by 5th Gen AMD EPYC processors delivered up to a 1.25x performance improvement vs. an equivalent system powered by 4th Gen AMD EPYC processors. Similar improvements were observed for the other Ansys solvers as well.

HPE server CAE benchmarks—HPE Cray XD2000 systems with 4th Gen AMD EPYC vs. 5th Gen AMD EPYC CPUs

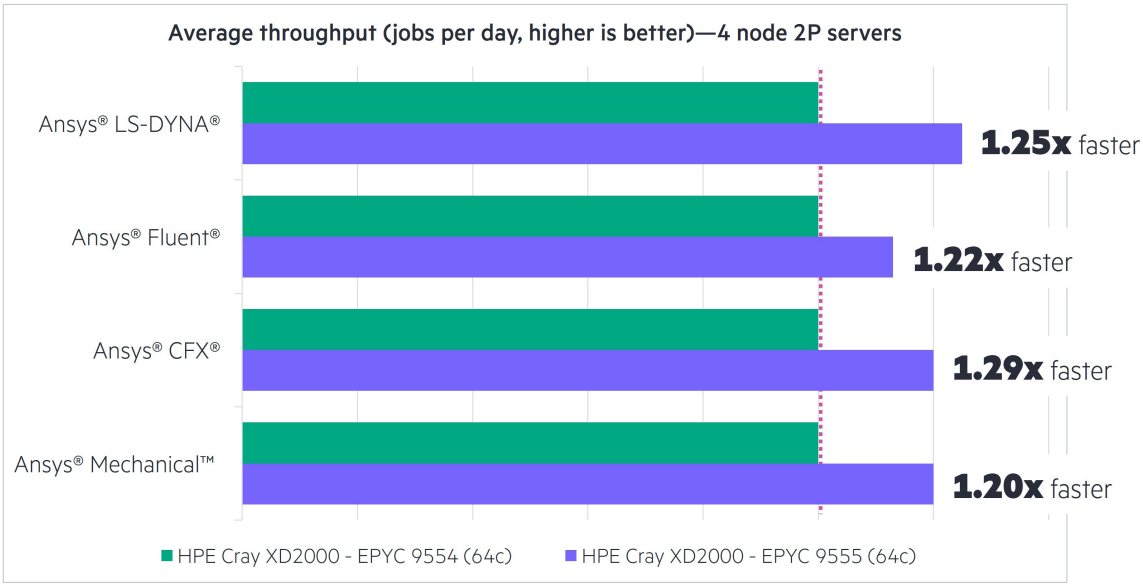


Figure 6. Comparing similar HPE Cray XD2000 systems with AMD EPYC 4th and 5th Gen CPUs across standard Ansys workloads

[24] Detailed results are available from HPE with a nondisclosure agreement (NDA)

[25] HPE internal testing—The Ansys tests were conducted in October 2024, comparing performance of an HPE Cray XD2000 system running 8x 4th Gen AMD EPYC 9554 processors (4x HPE Cray XD225v Servers) with a base clock of 3.10 GHz compared to an identical system running 8x 5th Gen AMD EPYC 9555 processors with a base clock of 3.20 GHz. Results may vary based on factors such as silicon version, hardware and software configuration, memory DIMMs, and driver versions.

[26] The Ansys Fluent test involved nine different models; The Ansys CFX test involved five different models; The Ansys LS-DYNA tests involved two models; The Ansys Mechanical test involved three models.

The "EPYC" advantage

First introduced in June 2017, AMD EPYC processors combine high core counts, large memory capacity, extreme memory bandwidth, large cache sizes, and massive I/O with the right ratios to enable exceptional performance for technical workloads. For CAE customers, this can translate into faster, more thorough simulations of larger, more complex models, helping deliver a significant competitive advantage.

AMD EPYC 9005 processors are the newest generation of the powerful and efficient AMD EPYC processor family for servers that have set hundreds of performance and efficiency world records.[27] Advancements in the AMD EPYC 9005 processor family are enabled by the highly efficient Zen 5 processor core architecture and advanced microprocessor process technologies to better meet the needs of the modern CAE and AI data centers. The complete line of processor offerings includes a wide range of core counts (up to 192 cores, 384 threads per CPU), frequencies (up to 5 GHz), cache capacities, energy efficiency levels, and competitive cost points—all complemented by the familiar x86 software compatibility that allows AMD EPYC 9005-based servers to readily support almost any business need.

5th Gen AMD EPYC processors offer several advantages over the previous generation and competing processors:

- Up to 192 cores, 50% more than the 4th Gen top-of-stack AMD EPYC 9754 processor[28]
- Up to a geomean 37% instruction-per-clock (IPC) uplift single threaded uplift over the previous generation for HPC workloads[29]
- 25% greater memory throughput compared to the previous generation[30]
- Advanced chip-level security feature enhancements (SME, SEV-ES, SEV-SNP)

While performance is important, energy efficiency and data center sustainability are also critical. AMD EPYC processors power the industry's most energy-efficient x86 servers, delivering exceptional performance and helping lower energy consumption.[31] In the SPECpower® benchmark widely used to measure energy efficiency, a dual-socket system powered by 4th Gen 128-core AMD EPYC 9754 processors delivered approximately 2.25x the energy efficiency of a dual-socket system powered by competing processors.[32] This means manufacturing customers can achieve dramatic performance and efficiency gains by carefully selecting processor technologies for their server infrastructure.

An ideal architecture for CAE and AI workloads

The unique architecture shown in Figure 7 is the key to the AMD EPYC processor's throughput advantage. While a 128-core AMD EPYC 9745 with 8 core complex dies (CCDs) and 256 MB of L3 cache is shown in the illustration, a wide range of SKUs are available to support different workloads. Depending on the SKU, 5th Gen AMD EPYC processors feature a single I/O die and up to 16 CCDs. The advanced TSMC N3 (3 nm) manufacturing process used for the CCDs enables clock frequencies to scale up to 5.00 GHz max boost for selected AMD EPYC CPUs, delivering exceptional single-threaded performance for CAE and AI workloads.[33]

While other processors share relatively small amounts of L3 cache across multiple cores, 5th Gen AMD EPYC processors offer up to 512 MB of L3 cache per socket. Each core on a core complex (CCX) has a direct path to its associated L3 cache to speed throughput and help reduce latency.[34] This combination of large amounts of L3 cache per core, direct channels to cache, multiple memory channels, and fast memory transfer speeds helps deliver exceptional throughput.

[27] See AMD EPYC processor world records.

[28] The top-of-stack AMD EPYC 9965 provides 192 cores. The 4th Gen AMD EPYC 9754 provides 128 cores, a 50% increase.

[29] amd.com/en/legal/claims/epyc.html#q=9xx5-001.

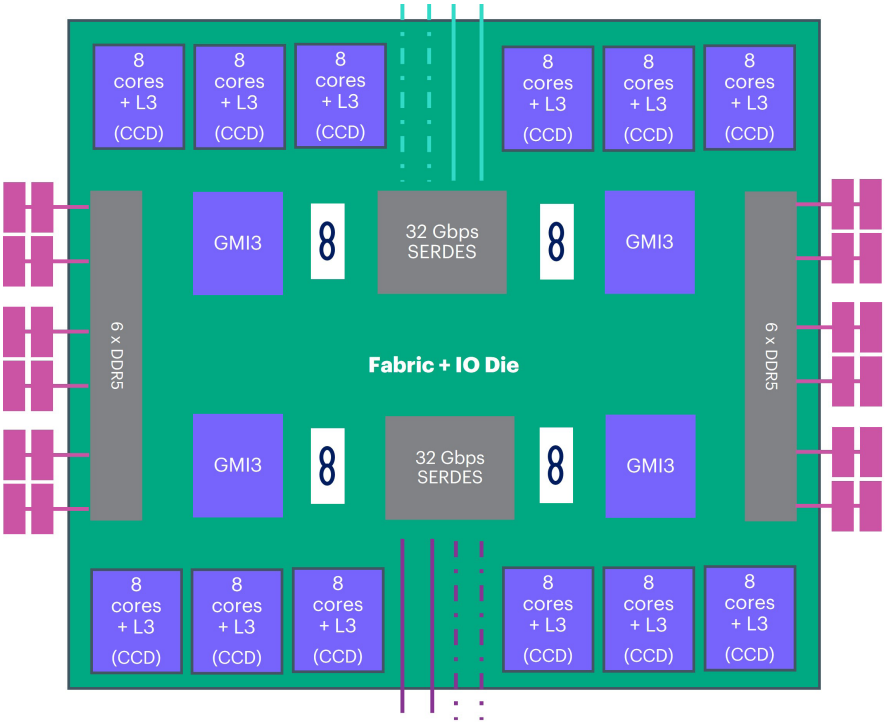
[30] 5th Gen AMD EPYC processors support DDR5 DIMMs operating at 6000 MT/s vs 4800 MT/s for the previous generation.

[31] EPYC-028D: SPECpower_ssj® 2008, SPECrate 2017_int_energy_base, and SPECrate 2017_fp_energy_base based on results published on SPEC's website as of 2/21/24. VMmark® server power-performance / server and storage power-performance (PPKW) based results published at vmware.com/products/vmmark/results3x.1.html?sort=score. The first 105 ranked SPECpower_ssj®2008 publications with the highest overall efficiency overall ssj_ops/W results were all powered by AMD EPYC processors. For SPECrate 2017 Integer (Energy Base), AMD EPYC CPUs power the first 8 top SPECrate®2017_int_energy_base performance/system W scores. For SPECrate®2017 Floating Point (Energy Base), AMD EPYC CPUs power the first 12 SPECrate®2017_fp_energy_base performance/system W scores. For VMmark server power-performance (PPKW), have the top 5 results for 2- and 4-socket matched pair results outperforming all other socket results and for VMmark server and storage power-performance (PPKW), have the top overall score. See amd.com/en/claims/epyc4#faq-EPYC-028D for the full list. For additional information on AMD sustainability goals see: amd.com/en/corporate/corporate-responsibility/data-center-sustainability.html. More information about SPEC® is available at spec.org. SPEC, SPECrate, and SPECpower are registered trademarks of the Standard Performance Evaluation Corporation. VMmark is a registered trademark of VMware in the US or other countries.

[32] SP5-011F: SPECpower_ssj®2008 comparison based on published 2P server results as of 1/12/2024. Configurations: 2P 128-core AMD EPYC 9754 (36,210 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2024q1/power_ssj2008-20231205-01347.html) is 2.25x the performance of best published 2P 64-core Intel Xeon Platinum 8592+ (16,106 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2024q1/power_ssj2008-20231205-01349.html) SPEC® and SPECpower_ssj® are registered trademarks of the Standard Performance Evaluation Corporation. See spec.org for more information.

[33] Refers to the 5th Gen 9575F part. Max. Boost for AMD EPYC processors is the maximum frequency achievable by any single core on the processor under normal operating conditions for server systems. (EPYC-18).

[34] CCX is a term used in AMD CPUs that stands for core complex. Up to 8 Zen 5 cores or 16 Zen 5c cores can be combined with 32 MB of shared L3 cache to create a CCX. These core complexes are fabricated onto a die (CCD). The number of CCDs per processor varies. See 5th Gen CPU cores on page 5 of the 5th Gen AMD EPYC Processor Architecture white paper.



5th Gen AMD EPYC processors*

- Next-generation 3 nm technology
- Up to 192 cores/384 threads[35]
- Up to 5.0 GHz max boost clock speed[36]
- Up to 512 MB L3 cache[37]
- 12 DDR5 memory channels per socket
- Up to 6000 MT/s DDR5
- 128 PCIe® Gen5 lanes per socket**
- CXL™ 1.1+ memory expansion
- Embedded AMD Secure Processor

* Specifications vary by product SKU
** Max of 160 PCIe lanes in 2P configurations

Figure 7. 5th Gen AMD EPYC 9745 (128 core) high-level architecture

AMD EPYC processors are the choice of next-generation exascale supercomputers.[38] With high core counts (up to 192 cores/processor), the latest 5th Gen AMD EPYC processors can help improve scaling and throughput and reduce simulation time. Despite their exceptional performance, they are affordable, fitting the budgets of small- or medium-sized CAE environments typically found among tier-1 and tier-2 suppliers.

For most CAE applications, 32-core or 64-core 4th and 5th Gen parts will provide optimal performance and value. Processors such as the 32-core 5th Gen AMD EPYC 9355 and 64-core AMD EPYC 9555 deliver leadership per-core performance while offering large amounts of L3 cache. Combining high clock frequencies, high memory bandwidth, and large amounts of L3 cache can help manufacturers improve performance and time-to-market and reduce TCO by helping deliver better performance with a smaller data center footprint. 5th Gen AMD EPYC processors are a good choice for CAE workloads because they offer:

- Exceptional performance to optimize CAE tool license efficiency
- Large L3 cache-per-core ratios for reduced analysis runtime
- 12x DDR5 high-speed memory channels supporting up to 6000 MT/s for exceptional memory throughput
- Dedicated L3 cache per CCD, enabling more concurrent analyses per server

[35] The specific EPYC SKUs offered will depend on the HPE server platform.

[36] 5.0 GHz boost speed is available only on the high-frequency 5th Gen AMD EPYC 9575F and AMD EPYC 9175F parts.

[37] 512 MB of L3 cache available on the 5th Gen AMD EPYC 9755 and AMD EPYC 9175F parts.

[38] AMD EPYC-based systems have been chosen as the basis of exascale supercomputers. Design wins include Frontier, a collaboration between the U.S. Department of Energy (DOE), ORNL, and El Capitan, a collaboration between U.S. DOE, LLNL, and HPE.

Table 2. AMD EPYC 9005 and 9004 series processors recommended for CAE workloads

EPYC model	Cores / threads	Base clock	Max. boost clock[39]	L3 cache	Power (watts)	Cooling	Workload type
5th Gen AMD EPYC processors							
9655	96/192	2.60 GHz	4.50 GHz	384 MB	400	Liquid (DLC)	Throughput
9575F	**64/128**	**3.30 GHz**	**5.00 GHz**	**256 MB**	**400**	**Liquid (DLC)**	**Core performance**
9555(P)	**64/128**	**3.20 GHz**	**4.40 GHz**	**256 MB**	**360**	**Liquid (DLC)**	**Throughput**
9355(P)	**32/64**	**3.55 GHz**	**4.40 GHz**	**256 MB**	**280**	**Air**	**Balanced**
9135	16/32	3.65 GHz	4.30 GHz	64 MB	200	Air	Balanced
4th Gen AMD EPYC processors							
9654	96/192	2.40 GHz	Up to 3.7 GHz	384 MB	360	Liquid (DLC)	Throughput
9554	**64/128**	**3.10 GHz**	**Up to 3.75 GHz**	**256 MB**	**360**	**Liquid (DLC)**	**Throughput**
9454	**48/96**	**2.75 GHz**	**3.8 GHz**	**256 MB**	**290**	**Air**	**Balanced**
9354	**32/64**	**3.25 GHz**	**3.8 GHz**	**256 MB**	**280**	**Air**	**Balanced**
9474F	**48/96**	**3.60 GHz**	**Up to 4.1 GHz**	**256 MB**	**360**	**Liquid (DLC)**	**Core performance**
9374F	**32/64**	**3.85 GHz**	**Up to 4.3 GHz**	**256 MB**	**320**	**Air**	**Core performance**
9274F	**24/48**	**4.05 GHz**	**Up to 4.30 GHz**	**256 MB**	**320**	**Air**	**Core performance**
9174F	16/32	4.10 GHz	Up to 4.40 GHz	256 MB	320	Air	Core performance

Table 2 shows a selection of the processor SKUs available for HPE Cray XD2000 systems and HPE ProLiant Gen11 servers.[40] For CAE workloads, the bolded processors are often considered optimal to achieve a good balance of clock speed, per core throughput, and available L3 cache. Customers can consult their HPE representative to discuss the best CPU profile depending on their workloads and goals.

[39] Max Boost speed for AMD EPYC processors is the maximum speed achievable by any core on the processor under normal operating conditions for server systems. (EPYC-18)

[40] For a complete list of available processors, see the HPE Cray XD2000 QuickSpecs and the HPE ProLiant DL385 Gen11 QuickSpecs

AMD Instinct accelerators

AMD Instinct accelerators are designed to deliver leadership performance for AI and HPC applications optimized to take advantage of GPUs. Customers can choose appropriate HPE servers depending on the AMD Instinct accelerator they require, as illustrated in Table 3.

Table 3. Supported AMD Instinct accelerators by HPE Server

AMD Instinct accelerator	MI210X	MI300X	MI325X
Supported HPE Server	HPE Cray XD2000 system (XD295v Server)	HPE Cray XD675 Server, HPE ProLiant Compute XD685	HPE ProLiant Compute XD685
GPUs/servers (up to)	2	8	8
FP32 vector performance (peak)	45.3 TFLOPs	163.4 TFLOPs	163.4 TFLOPs
Connection	PCIe	OAM Infinity Fabric	OAM Infinity Fabric
GPU compute units	104	304	304
Memory	64 GB HBM2e	192 GB HBM3	256 GB HBM3E
Peak theoretical memory BW	1.6 TB/s	5.3 TB/s	6 TB/s

The AMD Instinct MI325X GPU accelerators, with 3rd Gen AMD CDNA™ architecture, deliver performance and efficiency for training and inference. With industry-leading memory capacity and 6 TB/s bandwidth, they optimize performance and help reduce TCO. Integrated with AMD ROCm software, they support key AI frameworks, simplifying deployment. These accelerators are compatible with AMD Instinct MI300X platform infrastructure, with strong Day 0 support from Open AI, PyTorch, Hugging Face, Databricks, Lamini, and many more.

For GPU-intensive CFD or model training workloads that benefit from multiple GPUs and large amounts of GPU memory, HPE recommends GPU-dense configurations such as the HPE Cray XD675 Server or HPE ProLiant Compute XD685 Server. Compared to competitive accelerators, the AMD Instinct MI325X GPU can deliver up to 2.4x the HPC performance (Peak TFLOPs).[41]

[41] MI325-008: Calculations conducted by AMD Performance Labs as of October 2nd, 2024 for the AMD Instinct™ MI325X (1000W) GPU designed with AMD CDNA™ 3 5nm | 6nm FinFET process technology at 2,100 MHz peak boost engine clock resulted in 163.4 TFLOPs peak theoretical double precision Matrix (FP64 Matrix), 81.7 TFLOPs peak theoretical double precision (FP64), 163.4 TFLOPs peak theoretical single precision Matrix (FP32 Matrix), 163.4 TFLOPs peak theoretical single precision (FP32), 653.7 TFLOPS peak theoretical TensorFloat-32 (TF32), 1307.4 TFLOPS peak theoretical half precision (FP16), Actual performance may vary based on final specifications and system configuration. Published results on NVIDIA H200 SXM (141GB) GPU: 66.9 TFLOPs peak theoretical double precision tensor (FP64 Tensor), 33.5 TFLOPs peak theoretical double precision (FP64), 66.9 TFLOPs peak theoretical single precision (FP32), 494.7 TFLOPs peak TensorFloat-32 (TF32), 989.5 TFLOPS peak theoretical half precision tensor (FP16 Tensor). TF32 Tensor Core performance were published by NVIDIA using sparsity; for the purposes of comparison, AMD converted these numbers to non-sparsity/dense by dividing by 2, and this number appears above. NVIDIA H200 source: nvdam.widen.net/s/nb5zzzsjdf/hpc-datasheet-sc23-h200-datasheet-3002446 and anandtech.com/show/21136/nvidia-at-sc23-h200-accelerator-with-hbm3e-and-jupiter-supercomputer-for-2024

AMD-powered HPE servers—ideal for CAE workloads

CAE applications such as computational fluid dynamics (CFD) and finite element analysis (FEA) have different computing requirements. CFD algorithm solutions are often memory-bound and benefit from servers with large amounts of memory, multiple memory channels, and large amounts of L3 cache.

Implicit FEA, such as structural or thermal analysis, involves computationally expensive sparse matrix inversion, which is typically limited by memory size and bandwidth. Explicit FEA problems, such as crash and transient nonlinear analysis, need high processor performance—these workloads benefit from higher core counts and high-frequency processors with large amounts of cache.

Depending on the type of CAE problem, mixing and matching large core-count processors with high frequencies, large amounts of L3 cache, high memory bandwidth, and massive I/O can help solve CAE problems more quickly. HPE servers with AMD EPYC processors deliver a broad range of unique choices for manufacturers to help optimize their CAE environments. High-core-count EPYC processors can deliver high throughput per node for CAE applications that benefit from multicore parallelism. Lower-core count processors with high frequencies and large L3 caches offer high performance per core, helping to utilize per-core software licenses efficiently.

Accelerating AI model training and inference

AMD EPYC processors can accelerate your entire AI journey, providing a host processor for GPU-accelerated ML and an efficient processor for AI inferencing. Servers powered by 4th and 5th Gen AMD EPYC processors and the latest AMD Instinct accelerators offer the bandwidth needed to train predictive models and arrive at the necessary parameter weights to make models function with the speed and accuracy you need. AMD Instinct accelerators provide exceptional AI performance compared to the competition. For example, based on AMD internal testing and looking at commonly used AI precisions, the AMD Instinct MI325X can provide up to 1.3x the performance compared to competing accelerators.[42]

Advancing sustainability initiatives

For data center operators, power consumption and associated carbon emissions are increasingly important considerations. Based on science-based goals aligned to a 1.5C˚ warning scenario, AMD has pledged to reduce emissions from its operations (Scopes 1 and 2) by 50% by 2030.[43] AMD offers a greenhouse gas (GHG) emissions TCO estimation tool that can be used to estimate the potential savings and emission reductions with various AMD EPYC CPUs.[44]

HPE is similarly committed to sustainability and has pledged to source 100% of its electrical energy supporting operations from renewable sources by 2025. HPE has also pledged to achieve net-zero GHG emissions across its value chain by 2040.[45]

[42] MI300-17: Measurements conducted by AMD Performance Labs as of November 11th, 2023 on the AMD Instinct™ MI300X (750W) GPU designed with AMD CDNA™ 3 5nm | 6nm FinFET process technology at 2,100 MHz peak boost engine clock resulted in 653.7 TFLOPS peak theoretical TensorFloat-32 (TF32), 1307.4 TFLOPS peak theoretical half precision (FP16), 1307.4 TFLOPS peak theoretical Bfloat16 format precision (BF16), 2614.9 TFLOPS peak theoretical 8-bit precision (FP8), 2614.9 TOPs INT8 floating-point performance. The MI300X is expected to be able to take advantage of fine-grained structure sparsity providing an estimated 2x improvement in math efficiency resulting 1,307.4 TFLOPS peak theoretical TensorFloat-32 (TF32), 2,614.9 TFLOPS peak theoretical half precision (FP16), 2,614.9 TFLOPS peak theoretical Bfloat16 format precision (BF16), 5,229.8 TFLOPS peak theoretical 8-bit precision (FP8), 5,229.8 TOPs INT8 floating-point performance with sparsity. Published results on NVIDIA H100 SXM (80GB) 700W GPU resulted in 989.4 TFLOPs peak TensorFloat-32 (TF32) with sparsity, 1,978.9 TFLOPS peak theoretical half precision (FP16) with sparsity, 1,978.9 TFLOPS peak theoretical Bfloat16 format precision (BF16) with sparsity, 3,957.8 TFLOPS peak theoretical 8-bit precision (FP8) with sparsity, 3,957.8 TOPs peak theoretical INT8 with sparsity floating-point performance. NVIDIA H100 source: resources.nvidia.com/en-us-tensor-core.

[43] In the AMD Corporate Responsibility Report and other AMD climate-related disclosures, AMD references this operational GHG reduction goal (market-based emissions) as a "science-based target" that is aligned with a 1.5˚C scenario. Per the California Climate and Carbon Disclosure Requirements (AB-1305), the basis for this statement is the AMD GHG goal is aligned with the Science-based Target initiative's (SBTi's) 1.5-degree minimum target ambition of 4.2% linear annual reduction. The SBTi criteria considers multiple climate scenario models from the IAMC and IEA. Interim progress by AMD toward the goal is disclosed in the AMD CR Report (page 35,37) as measured by metric tCO2e of Scope 1 and 2 GHG emissions (market-based). The AMD 2020 base year value is 61,754 metric tCO2e and the 2030 target year value is 30,877 metric tCO2e (a 50% reduction). AMD reports total Scope 1 and 2 GHG emissions for each interim year and undergoes third-party limited assurance of the Scope 1 and Scope 2 values, as well as the percentage completion toward the goal. For the most current AMD reported data, calendar year 2023, AMD reported 46,605 metric tCO2e Scope 1 and 2 emissions, approximately a 24.5% reduction from the 2020 baseline. The data received limited level assurance in accordance with the International Standard on Assurance Engagements (ISAE) 3000 Revised.

[44] AMD EPYC Bare Metal and Greenhouse Gas Emissions TCO Estimation Tool.

[45] "HPE Living Progress Report for Fiscal Year 2023," HPE, 2024

Cloud deployment options with HPE GreenLake

Digital transformation is driving new data-intensive workloads and the need for real-time analytics at an unprecedented scale. This has increased demand for HPC and CAE in cloud environments. HPE GreenLake democratizes HPC by allowing companies of all sizes to access highly performant infrastructure for simulation workloads. HPE is a leader in traditional HPC, and consumption-based solutions from HPE GreenLake are a natural next step. HPE GreenLake is a private cloud on-premises or colocation, turnkey solution fully managed and operated by HPE. CAE-aaS is built on the foundation of HPE GreenLake and makes it easier and faster for customers to deploy converged CAE/AI workloads on high performance clusters with predictable, transparent costs and continuous monitoring to enable capacity rightsizing plus the ability for capacity bursting on-site, and on-demand.

HPE can help find the optimal solution for CAE workloads

Customers run a wide variety of CAE workloads depending on their industry, mix of tools, and unique business requirements. HPE works closely with AMD and ISV partners to run a variety of internal benchmarks comparing systems based on AMD EPYC processors to competitive CPUs using standard CAE applications. HPE can help customers find the optimal solution for their needs by sharing details of standard benchmarks we have run. We can also assist customers who wish to benchmark their own application workloads. For additional details on available benchmarks or to inquire about running your own benchmarks, contact your HPE representative.

Reduce data center footprint and TCO

Key metrics for CAE data center managers include job turnaround time, simulations per rack, and throughput per kilowatt-hour (kWh). Industry-standard benchmarks are helpful, but what matters more is how an organization's unique mix of HPC applications performs in the real world. The benchmarks described previously show that manufacturers can significantly increase simulation throughput for various CAE workloads using HPE Cray XD2000 systems powered by 4th and 5th Gen AMD EPYC processors. This means that manufacturers can often achieve the same performance with fewer compute nodes resulting in savings across multiple dimensions, including infrastructure, power, cooling, and administrative costs.

Conclusion: Why HPE and AMD for manufacturing?

As a market-leading HPC and AI solution provider, HPE delivers a unified compute and storage portfolio designed to simplify system and data management, reduce costs and complexity, and scale to deliver the exceptional performance needed for the next generation of CAE solutions in manufacturing.

Worldwide, many manufacturing companies are already using these HPE CAE solutions. As CAE and AI become an even more integral part of the supply chain and product lifecycle, HPE Cray systems and HPE ProLiant servers powered by 4th and 5th Gen AMD EPYC processors can deliver excellent application performance to help manufacturers solve their most complex problems, innovate faster, and improve productivity and profitability.

Using HPE servers with the latest AMD EPYC processors, manufacturers can:

— Accelerate the design process to meet time-to-market pressures
— Reduce runtimes to help maximize productivity
— Realize higher throughput to improve design quality
— Rightsize infrastructure investments to optimize TCO

Learn more at

HPE.com/servers/CrayXD2000

AMD.com/en/products/processors/server/epyc/9005-series.html

HPE.com/partners/amd

HPE.com/info/manufacturing

Visit HPE.com

Chat now

a00130824ENW, Rev. 4

HEWLETT PACKARD ENTERPRISE

hpe.com