



Hewlett Packard
Enterprise

Family guide

Check if the document is available
in the language of your choice.



HPE APOLLO AND HPE CRAY SYSTEMS

A supercomputer for data centers of all sizes



TABLE OF CONTENTS

3	DENSITY-OPTIMIZED SOLUTIONS FOR HIGH-PERFORMANCE COMPUTING AND ARTIFICIAL INTELLIGENCE
3	ENABLING THE EXASCALE ERA THROUGH PURPOSE-BUILT COMPUTE PLATFORMS
4	TYPICAL HPC AND AI USE CASES
5	THE HPE APOLLO SYSTEMS FAMILY
6	HPE APOLLO 20 SYSTEMS
7	TECHNICAL SPECIFICATIONS: HPE APOLLO 20 SYSTEM
8	HPE APOLLO 2000 SYSTEMS
9	TECHNICAL SPECIFICATIONS: HPE APOLLO 2000 SYSTEM
11	HPE APOLLO 6500—MASSIVE GPU COMPUTE
12	TECHNICAL SPECIFICATIONS: HPE APOLLO 6500 GEN10 SYSTEM
13	HPE CRAY SUPERCOMPUTER: PERFORMS LIKE A SUPERCOMPUTER, RUNS LIKE A CLOUD
14	TECHNICAL SPECIFICATIONS: CRAY
15	SOFTWARE ECOSYSTEM FOR HPE APOLLO SYSTEMS
16	HPE POINTNEXT SERVICES



Different data assets have different value

- Mission-critical data must be stored for a long time—e.g., customer data
- Some data must be stored for regulatory compliance—e.g., email archiving
- Some data has temporal value and can be discarded after its value diminishes—e.g., Social Media feeds

DENSITY-OPTIMIZED SOLUTIONS FOR HIGH-PERFORMANCE COMPUTING AND ARTIFICIAL INTELLIGENCE

The HPE Apollo Family is designed to deliver efficient rack-scale solutions for your high-performance computing (HPC) and artificial intelligence (AI) workloads. With rack-scale efficiency, the HPE Apollo Systems Family delivers excellent business benefits:

- Delivers just the right amount of performance and efficiency with systems that are optimized for specific workloads
- Accelerates time-to-value by reducing implementation time
- Provides architectural flexibility with both scale-up and scale-out solutions
- Helps you reduce capital and operating expenditures (CAPEX and OPEX)
- Gives you peace of mind with complete HPE service and support offerings

The scale-out compute portion of the HPE Apollo System portfolio includes the HPE Apollo 2000 System for hyperscale and general-purpose scale-out computing, the HPE Apollo 6500 System and HPE Apollo 20 for accelerated compute and compute built on emerging technology HPE Cray EX for supercomputing.

This complete range of offerings makes highly dense storage, high performance, and rack-scale efficiency available to organizations of all sizes, with a tiered approach that provides a simplified starting point. HPE Apollo Systems are the logical choice for companies seeking to utilize HPC, AI, and supercomputing.

ENABLING THE EXASCALE ERA THROUGH PURPOSE-BUILT COMPUTE PLATFORMS

The ExaScale era is not just about a few large systems that only the largest national labs will purchase. It is about a new era of computing requirements that have emerged from the necessity to extract the most insight and value out of rapidly growing data.

Traditional data processing technologies are no longer adequate. It requires applying a broader set of analytics methodologies to your data in a coordinated fashion. Translates into using a varying combination of HPC, analytics, and AI workloads in a single workflow—often in real-time.

But winning in this new era requires more than just throwing resources at the problem. It requires taking an intelligent and nuanced approach to your data assets. An approach that supports both scale-out and scale-up architectures to provide capacity and performance scalability—one that recognizes that different data assets have different values, and optimizes investments across scale, performance, and cost-efficiency requirements for each type of data so you can manage the volume, velocity, and variety of data in the most efficient way possible. In other words—workload-optimized compute solutions.

Frontier is Oak Ridge National Lab's new supercomputer for the ExaScale era which will out perform today's top 200 supercomputers combined.



Becoming a data-driven organization

The data-driven organization is one that utilizes data in ways that were impossible just a few years ago, due to expense, space limitations, or lack of compute capacity. It recognizes that there is “no one size fits all” approach to managing, processing, and storing data, and leverages different technologies—each specifically optimized across scale, performance, and cost-efficiency attributes—to deliver a specific value proposition for each type of data.

In short, the data-driven organization leverages workload-optimized compute solutions to turn data into an asset that directly impacts the bottom line—through initiatives that empower faster decision-making, improved operational efficiency, and direct content monetization.

The data-driven organization is here, and it's powered by HPE density optimized systems.

TYPICAL HPC AND AI USE CASES

Supercomputing

Supercomputers enable the world's leading research scientists and mathematicians to solve the world's most complex problems. But while the human imagination is limitless, the massive space and energy requirements of contemporary supercomputers, combined with the drastic slowing of processor speed increases, are slowing the pace of research.

Hewlett Packard Enterprise, the leader in HPC, is breaking through traditional supercomputing barriers by rethinking cooling to enable high-density, energy-efficient supercomputing solutions. A prime example: the DOE's Lawrence Livermore National Laboratory (LLNL) named El Capitan, which will be built on the HPE Cray EX supercomputer technology. It will be managed and hosted by LLNL for use by the three NNSA national laboratories: LLNL, Sandia National Laboratories and Los Alamos National Laboratory. The system will enable advanced simulation and modeling to support the U.S. nuclear stockpile and ensure its reliability and security. This system is 10 times faster than today's most power supercomputer and will hit a performance record of 2 exaFLOPS.

Divisional and departmental HPC

The demand for more compute performance for applications used by engineering design automation (EDA), financial risk modeling, life sciences, and other modern workloads is relentless. Performance gains from microprocessors have been limited by power and manufacturability challenges that affect scalability. Given today's financial, power, and space constraints, Hewlett Packard Enterprise has taken a new approach: thinking beyond just the server and optimizing performance at the rack level to get the most out of the entire system infrastructure.





The HPE Apollo and HPE Cray Systems supports a full complement of HPC options

Optimize HPC clusters for many different applications including product design and testing simulation, financial risk modeling and Monte Carlo simulation, and scientific research modeling.

HPC options include top bin CPUs, fast memory, integrated accelerators (GPUs or coprocessors), and fast cluster fabrics and I/O interconnections, making it easy for you to achieve the right performance, and price/performance for your HPC workloads.

Scale-out data center

You need to deploy additional compute power for cloud, web-based applications, web hosting, and other workloads to speed research and get to market faster, but space and resource restrictions are getting in the way. HPE Apollo systems provide a bridge from traditional to scale-out architecture so you can achieve the power of HPC systems with the space and cost savings of density-optimized infrastructure—without disruption.

THE HPE APOLLO SYSTEMS FAMILY

HPE Apollo 20 System: accelerated compute with flexible options

Purpose-built, high density platform delivering advanced performance for memory and compute intensive HPC and AI workloads with innovative air and liquid cooled options.

HPE Apollo 2000 System: density optimized, scale-out compute

Deploy more compute power to reap the benefits of cloud business, deploy web-based applications, and increase HPC power to speed research and bring new products and services to market faster—within space and resource boundaries.

The HPE Apollo 2000 Gen10 and Gen10 Plus Systems provide a bridge to scale-out architecture for traditional data centers so you can achieve the space and cost savings of density-optimized infrastructure in a non-disruptive manner.

HPE Apollo 6500 System: high density GPU compute

The HPE Apollo 6500 Gen10 system solves problems faster with up to 15 TFLOPS of single-precision performance per 2U node. The HPE Apollo 6500 increases your return on IT investment by accelerating the performance of your data center workloads with up to eight GPUs or coprocessors. With high-powered analysis and prediction, you will solve your most demanding AI and HPC problems in the shortest time.

HPE Cray supercomputer: Performs like a supercomputer, runs like a cloud

In the HPE Cray supercomputer, hardware and software innovations tackle system bottlenecks, manageability, and job completion issues that emerge or grow when core counts increase, compute node architectures proliferate, and workflows expand to incorporate AI at scale. It eliminates the distinction between clusters and supercomputers with a single new system architecture, enabling a choice of computational infrastructure without tradeoffs.



HPE APOLLO 20 SYSTEMS

Accelerating performance to support diverse workloads

Enterprise HPC and AI application require high memory bandwidth, low latency and high performance and the 2P HPE Apollo 20 System delivers to meet these needs. This 2P system with independently serviceable nodes, supports both compute and memory bound applications with 32, 48, or 56 cores and 12 DIMMs per socket.

AI Acceleration built in

The HPE Apollo 20 System, built on the Intel® Xeon® Platinum 9200 Series processors, is the only HPE Apollo platform that takes advantage of the Intel® Deep Learning Boost (Intel DL Boost) instruction to deliver exceptional AI performance with up to ~2.5x faster inference over the Intel Xeon Platinum 8180 processor. Intel DL Boost is a group of acceleration features that increases in data parallelism to enhance performance of targeted workloads. The rapid proliferation of AI inference services, the need for these services to render results quickly, and the tendency for increasingly complex deep learning applications to be processor-intensive are helping drive unprecedented demand for high-performance, low-latency compute.

Flexible platform to meet data center needs

Built to support both liquid-cooled and air-cooled options, the HPE Apollo 20 System takes advantage of the Hewlett Packard Enterprise experience in HPC cooling technologies as workloads continue to push power and density. Both versions are supported and managed by a comprehensive HPC software stack include HPE Performance Cluster Manager, as well as worldwide service and support through HPE.

Key features and benefits

High density system purpose built for memory and compute bound applications

- Up to four independently serviceable servers in 2U chassis
- Hot-swappable storage, fans, and power supplies for maximum uptime (model specific)

Standardized scale-out platform for diverse workloads

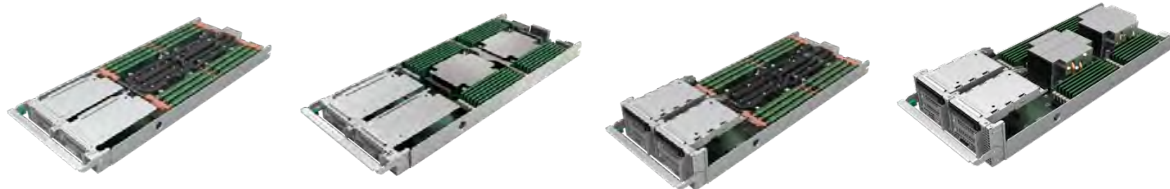
- 1U and 2U air-cooled and liquid-cooled trays
- Supports up to 350W in 2U air-cooled chassis or up to 400W with liquid-cooled versions

Advanced management and support integrated and validated high-performance software environment

- Global HPE Pointnext Services support
- HPE factory integration and testing tailored to your needs



TECHNICAL SPECIFICATIONS: HPE APOLLO 20 SYSTEM



Maximum number	1U half width—up to four per chassis	1U half width—up to four per chassis	2U half width—up to two per chassis	2U half width—up to two per chassis
Cooling	Direct-to-chip liquid cooling via cold plates	High-velocity air cooling	Direct-to-chip liquid cooling via cold plates	High-velocity air cooling
Processor family	Up to two Intel Xeon Platinum 9200 processors	Up to two Intel Xeon Platinum 9200 processors	Up to two Intel Xeon Platinum 9200 processors	Up to two Intel Xeon Platinum 9200 processors
Processor	56C 400W, 48C 350W, 32C 250W	32C 250W	48C 350W, 32C 250W	48C 350W, 32C 250W
PCIe Gen 3	Two low-profile PCIe cards through riser slot 1 and 2 riser cards	Two low-profile PCIe cards through riser slot 1 and 2 riser cards	Four low-profile PCIe cards through riser slot 1 and 2 riser cards	Four low-profile PCIe cards through riser slot 1 and 2 riser cards
Hot-swap storage	None	None	2x U.2 2.5" SSDs	2x U.2 2.5" SSDs
NVMe storage	2 M.2 per node	2 M.2 per node	2 M.2 & 2 U.2 per node	2 M.2 & 2 U.2 per node
Storage controller	Virtual RAID on Chip (VROC)	Virtual RAID on Chip (VROC)	Virtual RAID on Chip (VROC)	Virtual RAID on Chip (VROC)
System management	HPE Performance Cluster Manager	HPE Performance Cluster Manager	HPE Performance Cluster Manager	HPE Performance Cluster Manager
Size	2U: 17.39" wide x 34.06" deep	2U: 17.39" wide x 34.06" deep	2U: 17.39" wide x 34.06" deep	2U: 17.39" wide x 34.06" deep
Power supplies	3x hot-swap 2100W (Platinum) or 1600W (Titanium) PSUs	3x hot-swap 2100W (Platinum) or 1600W (Titanium) PSUs	3x hot-swap 2100W (Platinum) or 1600W (Titanium) PSUs	3x hot-swap 2100W (Platinum) or 1600W (Titanium) PSUs



HPE APOLLO 2000 SYSTEMS

Density optimized scale-out compute

To reap the benefits of cloud, web-based applications, and high-performance computing, you need to deploy more compute power—but you need to do so within space and resource constraints. The HPE Apollo 2000 System is a bridge to scale-out architecture for traditional rack-server data centers. It allows you to achieve the space and cost savings of density-optimized infrastructure in a non-disruptive manner.

Configuration flexibility for a variety of workloads

The HPE Apollo 2000 System is a very dense solution that packs a lot of performance and workload capacity into a small amount of data center space—making it ideal for your space-constrained data center or remote site. In fact, four independent hot-pluggable HPE Apollo 2000 servers in a single 2U chassis provide 2X the performance density of standard 1U servers—four servers in 2U vs. 4U of rack space—at a comparable cost.

Flexible configuration options make the HPE Apollo 2000 System a great fit for a variety of workloads, including HPC clusters. The ability to mix and match servers in the same chassis and the unique drive mapping flexibility allow you to create optimized configurations for many applications. Chassis, or groups of chassis, can be custom-configured to act as affordable, modular, 2U building blocks for specific implementations at scale—and for future growth.

Integrate seamlessly—and painlessly—into your data center

The HPE Apollo 2000 System is designed to be deployed in traditional enterprise data centers, without disruption or the need to change anything in your environment. The system can be managed at the individual server level with the same hardware and software tools and the same service procedures and practices used with traditional rack servers, significantly reducing the cost of change.

The HPE Apollo 2000 System has the right characteristics and delivers the right value to make it your enterprise bridge to efficient, space-saving, scale-out architecture.

Key features and benefits

Density-optimized for traditional data centers

- Up to four powerful servers in 2U chassis—2X the density of 1U servers
- Traditional racks and cabling for existing data centers
- Cost-effective in any configuration

Configuration flexibility for variety of workloads

- Mix and match servers for workload optimization (Gen10)
- HPC performance with accelerators, top bin CPUs, and fast HPC clustering
- Storage flexibility and a broad range of I/O options for workload optimization

Secure and manageable at scale—it's HPE ProLiant

- Secure from the start with iLO 5 and Silicon Root of Trust which provides a series of trusted handshakes from lowest level firmware to BIOS and software to ensure a known good state.
- Management comes easy with HPE Performance Cluster Manager which is a fully integrated system management solution offering all the functionalities you need to manage your Linux®-based high performance computing (HPC) clusters
- Other system management tools like HPE Apollo Platform Manager for advanced power and server control or HPE OneView for integration into overall enterprise infrastructure.



TECHNICAL SPECIFICATIONS: HPE APOLLO 2000 SYSTEM

HPE Apollo 2000 Gen10 System



HPE ProLiant XL170r Gen10 Server: 1U node



HPE ProLiant XL190r Gen10 Server: 2U node

Maximum number	1U half width—Up to four per chassis	2U half width—Up to two per chassis
Processor	Up to two Intel Xeon Scalable processors per server node, up to 150W	Up to two Intel Xeon Scalable processors per server node, up to 150W
Memory	Supports up to 2666/2933 MT/s DDR4 SmartMemory; 6 channels, 16 slots	Supports up to 2666/2933 MT/s DDR4 SmartMemory; 6 channels, 16 slots
Network module	2 I/O slots for a choice of networking and clustering options including: 1 Gb, 10 Gb, 100 Gb Ethernet, 100 Gb/s EDR InfiniBand, 100 Gb/s Omni-Path, Fibre Channel and options for either 1 PCIe (x16) slot plus a FlexibleLOM or 2 PCIe (x16) slots	2 I/O slots for a choice of networking and clustering options including: 1 Gb, 10 Gb, 100 Gb Ethernet, 100 Gb/s EDR InfiniBand, 100 Gb/s Omni-Path, Fibre Channel and options for either 1 PCIe (x16) slot plus a FlexibleLOM or 2 PCIe (x16) slots
Storage	Up to 24 drives per node Dual SATA host based M.2 2242 NGFF SSDs—internal Hot-plug HDD support Internal USB port Hard drive mapping feature on r2800 chassis	Up to 24 drives per node Dual SATA host based M.2 2242 NGFF SSDs—internal Hot-plug HDD support Internal USB port Hard drive mapping feature on r2800 chassis
Storage controller	(1) HPE Smart Array S100i; optional: HPE Smart Array PCIe card optional: HPE Smart Array Controller	(1) HPE Smart Array S100i; optional: HPE Smart Array PCIe card optional: HPE Smart Array Controller
Supported accelerators	N/A	NVIDIA® Tesla T4, M10, P40, V100, V100s, RTX 4000, RTX 8000
Management interface options	HPE iLO (iLO 5) HPE Apollo Platform Manager HPE Performance Cluster Manager for cluster deployments HPE OneView only for integration into overall enterprise infrastructure	HPE iLO (iLO 5) HPE Apollo Platform Manager HPE Performance Cluster Manager for cluster deployments HPE OneView only for integration into overall enterprise infrastructure

HPE Apollo 2000 System Chassis options



HPE Apollo r2200 Gen10 Chassis



HPE Apollo r2600 Gen10 Chassis



HPE Apollo r2800 Gen10 Chassis

Storage configuration	Up to 12 LFF hot-plug SAS/SATA HDD/SSD, allocated equally across server nodes	Up to 24 SFF hot-plug SAS/SATA HDD/SSD, allocated equally across server nodes or 16 SFF SAS/SATA + 8 NVMe HDD	Up to 24 SFF hot-plug SAS/SATA HDD/SSD; supports flexible drive mapping enabling custom drive allocations to match workloads giving you flexible storage density for various applications
Size	2U: 17.64" wide x 31.21" deep	2U: 17.64" wide x 29.61" deep	2U: 17.64" wide x 29.61" deep
Power supplies	800W, 1600W or 1800–2200W Platinum Power Supplies, N+1 redundancy option	800W, 1600W or 1800–2200W Platinum Power Supplies, N+1 redundancy option	800W, 1600W or 1800–2200W Platinum Power Supplies, N+1 redundancy option



TECHNICAL SPECIFICATIONS: HPE APOLLO 2000 SYSTEM (CONTINUED)

HPE Apollo 2000 Gen10 Plus System



HPE ProLiant XL225n Gen10 Plus System

Maximum Number	1U half width—Up to four per chassis
Processor	Up to two AMD EPYC 7002 Series Processors per server node, 240+W
Memory	Supports up to 3200 MT/s DDR4 SmartMemory 8 channels, 16 slots
Network Module	Choice of Ethernet (1 Gb to 100 Gb), InfiniBand (100 Gb HDR or 200 Gb HDR) InfiniBand 1x 1 Gb + OCP 3.0 Form Factor
Storage	Dependent on chassis selection n2400 zero drives supported n2600 up to 24 SFF HDD/SSD or up to 8 NVMe SFF SSD Optional OS boot device: 2 internal M.2
Storage controller	HPE Smart Array S100i; optional HPE Smart Array PCIe card
Supported accelerators	N/A
Management Interface	HPE Performance Cluster Manager (HPCM), iLO Advanced (optional), Rack Consolidation Module (RCM) (optional) Two 1600W/1800W–2200W/3000W HPE Apollo Platform Manager option for rack level management
Cooling	(5) Single rotor fans (standard) and an additional 2 rotor fans can be added for additional cooling
Warranty (parts, labor, on-site support)	3/3/3

HPE Apollo 2000 Plus System Chassis options



HPE Apollo n2400 Gen10 Plus Chassis



HPE Apollo n2600 Gen10 Plus Chassis

Storage configuration	No drives supported with this chassis; optional M.2 SSD storage can be used for OS boot Utilizing a zero drive chassis enables support of CPUS over 240W+	Up to 24 SFF hot-plug SAS/SATA HDD/SSD, allocated equally across server nodes or up to 8 NVMe SSD
size	2U:17.32" wide x 36.38" deep	2U:17.32" wide x 36.38" deep
Power supplies	Two 1600W/1800W–2200W/3000W HPE Apollo Platform Manager option for rack level management	Two 1600W/1800W–2200W/3000W HPE Apollo Platform Manager option for rack level management





HPE APOLLO 6500—MASSIVE GPU COMPUTE

Do you need to rapidly transform massive data streams into actionable results? Are you running massively parallel data models that require more processing horsepower? Would you like to deliver the maximum amount of performance from each of your HPC racks?

The HPE Apollo 6500 System provides the tools and the confidence to deliver high-performance computing (HPC) innovation. With up to 125 TFLOPS of single precision performance per server with support for up to 8 GPUs, the HPE ProLiant XL270d Gen10 Accelerator Tray delivers the performance you need. The system supports NVLink, a high-bandwidth, energy-efficient interconnect that enables ultra-fast communication between the CPU and GPU, and between GPUs. To deliver superb performance for your deep learning workloads.

The system consists of three key elements:

- The HPE ProLiant XL270 Gen10 Server tray
- The HPE Apollo 6500 Chassis
- The HPE Apollo 6500 System: your next accelerated computing solution.

High-bandwidth, low-latency networking is tightly coupled to the accelerators allow you to take full advantage of your network. And the four x16 PCIe Gen3 slots add to your flexibility when choosing high-speed fabrics.



What's new?

- The HPE ProLiant XL270d Gen10 server utilizes the latest 2nd generation Intel® Xeon® Scalable processors including the new Intel® Xeon® Gold 6200R processors.
- Designed for thermal excellence in the enterprise data center, with up to 205W 2nd generation Intel Xeon Scalable processor family and a broad range of inlet air temperatures for easy deployment.
- Eight GPUs per server for faster and more economical deep learning system training compared to more servers with fewer GPUs each. Get more done, in less time.
- NVLink connects GPUs at up to 300 Gb/s for one of the world's most powerful computing servers. AI models that would consume days or weeks can now be trained in a few hours or minutes.
- Enterprise RAS with HPE iLO 5, easy access modular design, and 2+2 power supplies.
- Secure from the start with HPE iLO 5 and industry leading firmware security from Silicon Root of Trust



Features

- Flexible configuration for the most demanding high-performance computing workloads. The HPE Apollo 6500 System supports up to eight 300W GPU or coprocessors delivering increased performance.
- For workloads optimized for high peer-to-peer communication among the accelerators, place four (4) GPU on a single high-speed PCIe switch, and two banks to a CPU for eight (8) GPU per CPU.
- For workloads requiring higher CPU to GPU communications, choose our four (4) GPU per CPU configuration.
- The HPE ProLiant XL270 Gen10 Server supports industry standard 1st and 2nd Generation Intel Xeon Scalable processors, solid state drives (SSD) and NVMe drives and 128 GB DDR4 2,933 MHz memory for blazing performance.
- Up to 24 HPE DDR4 2,933 MHz Memory Modules per HPE ProLiant XL270 Gen10 Server for faster performance with data-intensive application workloads.
- High-bandwidth, low-latency networking between Accelerator Nodes.
- The HPE Apollo 6500 System includes two low profile PCIe Gen3 x16 slots to enable your choice of high-speed fabrics.
- Choice of accelerator topologies to suit your specific workload: NVLink, PCIe 4:1 or PCIe 8:1.
- Supports GPUDirect with four (4) GPUs per HPE InfiniBand Adapter.

TECHNICAL SPECIFICATIONS: HPE APOLLO 6500 GEN10 SYSTEM



HPE ProLiant XL270d Gen10 Server

Rack	Optimized for 1,200 mm deep racks
Chassis	HPE Apollo d6500 Chassis (4U, 2 server trays, up to 16 GPU/chassis) 6.96 x 17.638 x 37 in.
Processor	2nd Generation Intel Xeon Scalable processors up to 28 cores and 205W TDP Select 8100/8200, 6100/6200, 5100/5200, and 4100/4200 processors
Power	Up to 4 HPE 2200W Platinum 80 Plus (2+2)
Memory	24 2,933 MT/s DDR4 DIMMs, 3 TB max (24x 128 GB LRDIMM)
Network options	Embedded Ethernet adapter with 4 ports and/or optional HPE FlexibleLOM and PCIe adapters for high speed networking. Additionally, 4 x16 PCIe Gen3 on GPU module for high speed fabrics such as InfiniBand and Intel Omni-Path Architecture
I/O slots	4 x16 PCIe Gen3 from GPU Module (LP PCIe for PCIe, FHHL for SXM-2) for InfiniBand, Omni Path Architecture or high-speed Ethernet fabric 1 x16 FHHL PCIe Gen3 on system board module
Storage	Up to 16 SFF hot plug SAS/SATA/SSD or up to 4 NVMe SSD storage
Management interface options	Standard: HPE iLO 5 Management, Optional: iLO Advanced, HPE OneView Advanced
Warranty (parts, labor, on-site support)	3/3/3



HPE CRAY SUPERCOMPUTER: PERFORMS LIKE A SUPERCOMPUTER, RUNS LIKE A CLOUD

HPE Cray EX an entirely new design created to address the challenges of the ExaScale era. Hardware and software innovations tackle system bottlenecks, manageability, and job completion issues that emerge or grow when core counts increase, compute node architectures proliferate, and workflows expand to incorporate AI at scale. It eliminates the distinction between clusters and supercomputers with a single new system architecture, enabling a choice of computational infrastructure without tradeoffs.

The new system hardware architecture designed to support the multiple processor and accelerator options available today and in the future. The architecture supports processors, accelerators and interconnect in both liquid and air-cooled cabinetry to meet end user requirements for density, performance, and efficiency.

For customers requiring the greatest performance, density, and efficiency for large-scale systems, the HPE Cray EX supercomputer is available in liquid-cooled cabinetry which supports direct liquid cooling of all components in a compact bladed configuration.

The initial air-cooled HPE Cray supercomputing solution is based on the HPE Apollo 2000 Gen10 plus with 4 dual socket AMD EPYC 7002 (Rome) nodes with one HPE Slingshot interconnect per node, and running HPE Cray software. Functionally the processor equivalent versions of the liquid-cooled and air-cooled solutions are identical with the same interconnect and software support.

The HPE Cray EX supercomputer solution includes an optimized HPC interconnect, HPE Slingshot, which demonstrates high-throughput and sustained performance at ExaScale, and has the ability to run diverse workloads simultaneously without interference, and the flexibility to enable new approaches to computing such as supercomputing via the cloud.

The HPE Slingshot network is built around our new 64 port, 12.8 Tb/s switch providing industry leading 200 Gb/s connectivity to endpoints. This high radix switch coupled with HPE's enhanced Dragonfly topology enables scaling to over 250,000 endpoints with a maximum of three switch-to-switch hops between endpoints. It also incorporates a host of new features to ensure packets are routed efficiently and network congestion is avoided. Based on the industry-standard Ethernet protocol, HPE Slingshot enables straightforward connectivity with standard data center environments, third-party storage devices, and can directly exchange IP/Ethernet traffic with the outside world.

The HPE Cray supercomputers are complete solutions with software and hardware that are tightly integrated and performance-tuned to offer the best system performance while bringing new standard in flexibility, manageability, and resiliency to supercomputing.

HPE Cray supercomputer software stack addresses the needs both system administrators and end users, including software developers and consists of:

HPE Cray System Management offers is built-for-scale system management solution offering administrators all functionalities they need to keep the systems healthy, utilized to the maximum and accommodating wide range of workload requirements via aaS experience.

HPE Cray Operating System is a suite of high-performance software is designed to run large, complex applications and scale efficiently.

Software developers can then use, **HPE Cray Programming Environment**—a fully integrated software development suite offering programmers comprehensive set of tools for developing, porting, debugging, and tuning of their applications so they can shorten application development time and accelerate their performance.

The whole software stack is supported by HPE Pointnext Services.



TECHNICAL SPECIFICATIONS: CRAY



HPE Cray supercomputer



HPE Cray EX supercomputer

	HPE Cray supercomputer	HPE Cray EX supercomputer
Infrastructure architecture	Standard 19" rack	Proprietary cabinet
Density/Scale	Up to 4 dual-socket nodes per 2U chassis	Up to 512 processors/256 nodes per cabinet
Node architecture	Up to two AMD EPYC 7002 series processors (4 per 2U chassis)	Up to two AMD EPYC 7002 series processors (4 nodes per compute blade)
Memory (type, channels, slots)	Up to 8 3200 MT/s DDR4 channels per socket; 1 DIMM per channel; up to 1024 GB per node	Up to 8 3200 MT/s DDR4 channels per socket; 1 DIMM per channel; up to 1024 GB per node
Storage	2 x 25" SATA or PCIe Gen3 NVMe SSD per node (up to 7.68 TB SSD's)	No local storage
GPU support	TBD	Q4 2020—Support for single socket AMD EPYC 7002 with 4 NVIDIA V100 GPUs with NVLink. 2 nodes per blade. Up to 128 GPUs per cabinet
Infrastructure management	HPE Cray System Management	HPE Cray System Management
OS support	HPE Cray operating system (OS)	HPE Cray operating system (OS)
Power supply-hot plug	HPE Apollo 2000 Gen10 Plus/Shasta compute: 1+1 hot-swappable power supplies	Hot-swap, redundant power supplies, support up to 300 KW per cabinet
Interconnect	HPE Slingshot switch	HPE Slingshot switch, switches integrated in cabinet
Cooling	Air-cooled	Direct liquid-cooling, up to ASHRAE W4 water
Storage controller	System dependent	N/A
Warranty (parts, labor, on-site support)	3/3/3	3/3/3



SOFTWARE ECOSYSTEM FOR HPE APOLLO SYSTEMS

HPE offers customers a comprehensive software portfolio for HPC and converged workloads so you can choose the right mix of software to fit your exact needs—from choice of operating systems, system management software, workload and orchestration solutions to application development tools and more.

Our software has a proven track record of helping customers maximize the use of their HPC systems, including HPE Apollo clusters, so they can get results faster. Besides tools we created, we also offer leading third-party software and validate popular HPC open source solutions.

Interactive HPC software portfolio guide is available [here](#).

System Management Software for HPE Apollo Systems

HPE Performance Cluster Manager

Delivers a fully integrated system management solution for standard Linux-based HPE clusters. The software enables fast system setup from bare-metal, comprehensive hardware monitoring and management (aggregating iLO information), image management and software updates, as well as power management.

The HPE Performance Cluster Manager reduces the time and resources spent administering HPC systems, lowering total cost of ownership. Learn more [here](#).

HPE OneView

In enterprise and cross-discipline data center environments deploying HPE Apollo Systems, HPE OneView is a single, integrated management platform which also supports

HPE Synergy, HPE BladeSystem, c-Class, ProLiant server platforms, and HPE 3PAR storage systems. With best-in-class infrastructure lifecycle management, HPE OneView allows IT to manage their entire infrastructure lifecycle more efficiently through a single interface developed for the way you think and work. With greater visibility and control of infrastructure, HPE OneView helps IT become more efficient, agile and productive, saving time and money. HPE OneView discovery and monitoring is available on all HPE Apollo Systems. More information [here](#).

Application Development Tools

We offer two homegrown application development and acceleration solutions alongside third-party tools:

HPE Message Passing Interface (MPI)

The MPI development environment is designed to enable the development and optimization of the HPC applications on InfiniBand-based HPE HPC clusters. The software suite helps customers tune and accelerate compute-intensive applications running on standard Linux clusters. HPE Message Passing Interface (MPI) can even boost performance of applications developed with other MPI implementations at runtime without the need to recompile.

HPE Cray Programming Environment

Is a set of tightly coupled compilers, libraries, debuggers performance analysis and other tools designed to help programmers shorten their software development cycle and improve performance of HPC applications on HPE Apollo 80 and 2000 systems.



Family guide

For more information on HPE Pointnext Services, go to: hpe.com/pointnext

HPE Factory Express provides customization and deployment services along with your storage and server purchases. You can customize hardware to your exact specifications in the factory—helping speed deployment. hpe.com/info/factoryexpress

Optimize your IT investment strategy with new ways to acquire, pay for and use technology, in lockstep with your business and transformation goals. hpe.com/solutions/hpefinancialservices

HPE POINTNEXT SERVICES

HPE Pointnext Services help you confidently reduce risk, realize agility and stability, and make the most of your HPE Apollo System and high-performance compute investment. You can choose from a flexible selection of service levels to meet your requirements. We help you reap the benefits of your HPE Apollo solution as you successfully plan your solution, deploy and operate it with minimal disruption to your current environment. Connect to HPE to get advisory, operational and training help while preventing problems and solve issues faster. Our support technology lets you tap into the knowledge of millions of devices and thousands of experts to stay informed and in control, anywhere, anytime. Our advisory experts help you to architect and integrate your AI/ML solution securely. Choose from a flexible selection of services:

- **AI & data advisory services**—Plan, architect, pilot and design advanced solutions, deploy AI, machine learning, deep learning workloads, and integrate it with partner technologies.
- **Installation and Startup**—rapidly get up and running smoothly with knowledge transfer.
- **HPE Datacenter Care**—Delivered by an assigned team of HPE Pointnext Services experts, this service offers a personalized, proactive support approach with priority access to reactive support.
- **HPE Cray Advanced Support**—HPE's most comprehensive support solution is tailored to meet specific requirements of HPE Cray high performance computing.
- **HPE Foundation Care**—Provides hardware and software support with a simplified choice of coverage windows and response times. The service includes collaborative call management for assistance with leading x86 operating system software.
- **HPE Education Services**—Propel your IT staff to exceptional levels of performance by working with HPE Education Services, a leader in technical training. Our consultants can also help you prepare and support your staff throughout technology change.
- **HPE Integration and Performance Services**—Offer access to expertise for every step of the way—from strategy to design, as well as deployment and operations.

LEARN MORE AT

hpe.com/info/hpc

Make the right purchase decision.
Contact our presales specialists.



Chat



Email



Call



Get updates


**Hewlett Packard
Enterprise**

© Copyright 2015–2016, 2020 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

AMD is a trademark of Advanced Micro Devices, Inc. Intel, Intel Xeon, Intel Xeon Platinum, and Intel Xeon Gold are trademarks of Intel Corporation in the U.S. and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. NVIDIA, GPUdirect, NVLink, and Tesla are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All third-party marks are property of their respective owners.

4AA5-8958ENW, September 2020, Rev. 4