

Overview

HPE Machine Learning Development System

The HPE Machine Learning Development System is a standardized, validated, and pre-configured solution that reduces IT complexity and gives you out-of-the-box performance, allowing you to focus time and resources on model-development.

More companies are incorporating machine learning (ML) and deep learning (DL) into their products and services. And they're doing so at an accelerated rate, driving the need for artificial intelligence (AI) model development and training to become a core competency. Your enterprise has invested in AI and is ready now for the next evolutionary leap. The goal is for AI to become one of your frontiers of innovation, so that you can win with AI — not merely use it. But resource constraints all too often dictate the success or failure of an AI initiative. Whether it's due to the complexity and cost of model development and training at scale, or the enormous operational difficulties of deploying and managing AI infrastructure, or any of a handful of other common “last mile” challenges of operationalizing AI, your team may be thwarted in pursuing their best ideas. Machine Learning Engineers (MLEs), for example, often end up focusing on infrastructure more than high-quality models, and lock-ins with cloud systems limit the freedom to pivot to better solutions.

The HPE Machine Learning Development System helps you overcome these challenges, so you can bring your most impactful AI applications to life. You get everything needed to scale AI models easily from idea to impact — a model training and development software platform, high-performance computing, infrastructure software, networking components, accelerators, start-up services, and solution support — all preconfigured, fully installed, and performant out of the box.

The HPE Machine Learning Development System accelerates the pace at which your enterprise can improve model quality, scale into production, and achieve desired business objectives.

Differentiation of ML Dev System

The unique value proposition for the HPE Machine Learning Development System includes these attributes:

- **Complete**
Get everything — hardware and software, high-performance computing, cluster management, networking, accelerators, model training and development platform, and installation and support services — in one system purpose-built for scaling AI workloads. Single provider eliminates integration and support issues. Frees team to focus on driving outcomes.
- **Built for scale**
Give your team the resources to improve model quality and scale into production without hassles or delays. Free your data scientists and MLEs from writing infrastructure code. And save IT from wasting time in researching and configuring infrastructure solutions for AI model development and training. Seamlessly scale experiments from 10 to 100-s GPUs powered by distributed training, high-performance computing, and cluster management on a validated, purpose-built infrastructure.
- **Flexible**
Flexibility today with standard and custom options — modifiable prebuilt solutions with your preference, and a foundation for heterogeneity using accelerators. A variety of financing options available through HPE Financial Services.
- **Trusted**
Sterling reputation and proven success: HPE in datacenter and advisory services

System Components

HPE Machine Learning Development System is built with the following major components:

- **Compute System**
 - with HPE Cray XD670 servers with NVIDIA H100 GPUs or HPE Apollo XL675d servers with NVIDIA A100 GPUs
- **Management System**
 - with HPE ProLiant DL325 Gen 11 servers
- **Interconnect**
 - with HPE NVIDIA Infiniband HDR or NDR switches and adapters
 - with HPE Aruba Networking CX 6300M switches and adapters

Overview

- **Cluster Building**
 - with prebuilt configurations with 2, 4, 6, 8 nodes with racking, cabling, and power systems.
- **Infrastructure Software**
 - with RHEL, SUSE Rancher, HPE Performance Cluster Manager
- **Machine Learning Software**
 - with HPE Machine Learning Development Environment Software
- (optional) **Storage System**
 - with HPE ClusterStor or WEKA

Target Use Cases

- **Computer Vision**
 - Object Detection, Image Classification, Image Segmentation, Video Analytics, Aerial Image Analysis, Objectionable Content Detection.
 - **Natural Language Processing**
 - Text categorization, language modeling, summarization, machine translation, entity extraction, event detection, sentiment analysis, question-answering, conversational assistants.
 - **Event Stream Prediction**
 - Combined sensors: LIDAR, Camera, traditional sensors.
 - Security: network traffic analysis.
 - **Semi-structured data analysis**
 - Time series prediction, gene sequence prediction, etc.
-



Standard Features

Below table lists the core and optional components of ML Dev System

Core Component	Description
Accelerated Compute <ul style="list-style-type: none"> Quantities of 2 – 120 	HPE Cray XD670 Gen 11 – 8 GPU per node <ul style="list-style-type: none"> NVIDIA H100 80GB Tensor Core GPU with NVLink Intel Xeon CPU 4th Gen Or HPE Apollo 6500 Gen 10 Plus – 8 GPU per node <ul style="list-style-type: none"> NVIDIA A100 80GB Tensor Core GPU with NVLink AMD Milan CPU
Management <ul style="list-style-type: none"> Quantities of 3 or 6 	HPE DL325 Gen 10 Plus v2 Server (1U)
Fabric	Mellanox InfiniBand HDR, HPE Aruba Networking CX 6300 1GbE Switch
Training Platform	HPE Machine Learning Development Environment
Cluster Manager	HPE Performance Cluster Manager
Operating System	Red Hat Enterprise Linux/SLES
Container Engine	Docker
Deployment services and Solution Support	HPE Services / Product Success team
Optional Components	Description
Storage	HPE ClusterStor
Services	HPE Advisory and Professional Services

Component Architecture

HPE Machine Learning Development Environment

HPE Machine Learning Development Environment software addresses the challenges of developing and deploying complex infrastructure and training models at scale. Key benefits of HPE Machine learning Development Environment are as follows:

- Train models faster using state-of-the art distributed training.
- Find better models with advanced hyper parameter optimization HPE Machine Learning Development Environment is a cloud or on-prem based solutions that helps machine learning (ML) engineers and IT systems, and platform engineers focus on innovation and accelerate their time to production by removing the complexity and cost associated with ML model development. Our platform reduces the time to value for model developers by removing the need to write infrastructure code and makes it easy for IT administrators to set up, manage, secure, and share AI compute clusters. HPE Machine Learning Development Environment integrates with HPCM to monitor and manage both infrastructure and model metrics in a single interface.
- Maximize GPU resources with smart on-prem and cloud scheduling.
- Track and reproduce work with built-in experiment tracking.

Additional Details: <https://www.hpe.com/us/en/compute/hpc/cray-ai-development.html>



Standard Features

Compute Node – Apollo 6500 Gen10+

HPE Machine Learning Development System uses HPE Cray XD670 Servers with 8x NVIDIA® H100 Tensor Core SXM5 GPUs connected via NVLink as scalable compute nodes. HPE Cray XD670 Systems accelerate performance over previous MLDS versions by incorporating the latest and highest performing NVIDIA GPUs with NVLink or AMD Instinct MI100 with 2nd Gen Infinity Fabric Link to take on the most complex HPC and AI workloads. NVIDIA® H100 Tensor Core SXM5 GPUs are proven to be one of the fastest GPUs for scalable systems hosting HPC and AI training applications as shown in MLCommons MLPerf™ v3.0.

This purpose-built platform provides enhanced performance with premier GPUs, fast GPU interconnects, high-bandwidth fabric, and configurable GPU topology, providing rock-solid reliability, availability, and serviceability (RAS).

QuickSpecs: <https://www.hpe.com/psnow/doc/a50004292enw>

Management Node – ProLiant DL325

HPE Machine Learning Development System uses HPE ProLiant DL325 Gen 10 Plus v2 servers, to provide management functions. Powered by the 3rd generation AMD® EPYC® 7003 Series processor, the HPE ProLiant DL325 Gen10 Plus v2 offers greater processing power. On top of that, the chassis is smaller (1U) compared to the previous generation providing better compatibility to your infrastructure. Tri-mode RAID controller support provides flexibility to support across SAS/SATA/NVMe types of storage options.

QuickSpecs: https://www.hpe.com/psnow/doc/a00073548enw.html?jumpid=in_lit-psnow-red

Cluster Manager – HPE Performance Cluster Manager

HPE Machine Learning System leverages HPE Performance Cluster Manager for High Performance Computing cluster management. HPE Performance Cluster Manager delivers an integrated system management solution for Linux based High Performance Computing (HPC) clusters. HPE Performance Cluster Manager provides complete provisioning, management, and monitoring for clusters scaling to 100,000 nodes. The software enables fast system setup from bare-metal, comprehensive hardware monitoring and management, image management, software updates and power management.

HPE Performance Cluster Manager features

- View metrics and alerts via GUI, CLI, Ganglia, Naglios, Kibana, or Grafana.
- Customize system telemetry and alerts to best suite user needs.
- Setup automatic reactions to events to prevent failures.

Health checks helps customers run applications reliably and at peak performance.

QuickSpecs: https://www.hpe.com/psnow/doc/a00044858enw?jumpid=in_lit-psnow-red



Service and Support

HPE Services

No matter where you are in your digital transformation journey, you can count on HPE Services to deliver the expertise you need when, where and how you need it. From planning to deployment, ongoing operations and beyond, our experts can help you realize your digital ambitions.

<https://www.hpe.com/services>

Consulting Services

No matter where you are in your journey to hybrid cloud, experts can help you map out your next steps. From determining what workloads should live where, to handling governance and compliance, to managing costs, our experts can help you optimize your operations.

<https://www.hpe.com/services/consulting>

HPE Managed Services

HPE runs your IT operations, providing services that monitor, operate, and optimize your infrastructure and applications, delivered consistently and globally to give you unified control and let you focus on innovation.

[HPE Managed Services | HPE](#)

Operational services

Optimize your entire IT environment and drive innovation. Manage day-to-day IT operational tasks while freeing up valuable time and resources. Meet service-level targets and business objectives with features designed to drive better business outcomes.

<https://www.hpe.com/services/operational>

HPE Complete Care Service

HPE Complete Care Service is a modular, edge-to-cloud IT environment service designed to help optimize your entire IT environment and achieve agreed upon IT outcomes and business goals through a personalized experience. All delivered by an assigned team of HPE Services experts. HPE Complete Care Service provides:

- A complete coverage approach -- edge to cloud
- An assigned HPE team
- Modular and fully personalized engagement
- Enhanced Incident Management experience with priority access
- Digitally enabled and AI driven customer experience

<https://www.hpe.com/services/complecare>

HPE Tech Care Service

HPE Tech Care Service is the operational support service experience for HPE products. The service goes beyond traditional support by providing access to product specific experts, an AI driven digital experience, and general technical guidance to not only reduce risk but constantly search for ways to do things better. HPE Tech Care Service delivers a customer-centric, AI driven, and digitally enabled customer experience to move your business forward. HPE Tech Care Service is available in three response levels. Basic, which provides 9x5 business hour availability and a 2-hour response time. Essential which provides a 15-minute response time 24x7 for most enterprise level customers, and Critical which includes a 6-hour repair commitment where available and outage management response for severity 1 incidents.

<https://www.hpe.com/services/techcare>



Service and Support

HPE Lifecycle Services

HPE Lifecycle Services provide a variety of options to help maintain your HPE systems and solutions at all stages of the product lifecycle. A few popular examples include:

- Lifecycle Install and Startup Services: Various levels for physical installation and power on, remote access setup, installation and startup, and enhanced installation services with the operating system.
- HPE Firmware Update Analysis Service: Recommendations for firmware revision levels for selected HPE products, taking into account the relevant revision dependencies within your IT environment.
- HPE Firmware Update Implementation Service: Implementation of firmware updates for selected HPE server, storage, and solution products, taking into account the relevant revision dependencies within your IT environment.
- Implementation assistance services: Highly trained technical service specialists to assist you with a variety of activities, ranging from design, implementation, and platform deployment to consolidation, migration, project management, and onsite technical forums.
- HPE Service Credits: Access to prepaid services for flexibility to choose from a variety of specialized service activities, including assessments, performance maintenance reviews, firmware management, professional services, and operational best practices.

Notes: To review the list of Lifecycle Services available for your product go to:

<https://www.hpe.com/services/lifecycle>

For a list of the most frequently purchased services using service credits, see the [HPE Service Credits Menu](#)

Other Related Services from HPE Services:

HPE Education Services

Training and certification designed for IT and business professionals across all industries. Broad catalogue of course offerings to expand skills and proficiencies in topics ranging from cloud and cybersecurity to AI and DevOps. Create learning paths to expand proficiency in a specific subject. Schedule training in a way that works best for your business with flexible continuous learning options.

<https://www.hpe.com/services/training>

Defective Media Retention

An option available with HPE Complete Care Service and HPE Tech Care Service and applies only to Disk or eligible SSD/Flash Drives replaced by HPE due to malfunction.

Consult your HPE Sales Representative or Authorized Channel Partner of choice for any additional questions and services options.

Parts and Materials

HPE will provide HPE-supported replacement parts and materials necessary to maintain the covered hardware product in operating condition, including parts and materials for available and recommended engineering improvements.

Parts and components that have reached their maximum supported lifetime and/or the maximum usage limitations as set forth in the manufacturer's operating manual, product quick-specs, or the technical product data sheet will not be provided, repaired, or replaced as part of these services.

How to Purchase Services

Services are sold by Hewlett Packard Enterprise and Hewlett Packard Enterprise Authorized Service Partners:

- Services for customers purchasing from HPE or an enterprise reseller are quoted using HPE order configuration tools.
- Customers purchasing from a commercial reseller can find services at <https://ssc.hpe.com/portal/site/ssc/>



Service and Support

AI Powered and Digitally Enabled Support Experience

Achieve faster time to resolution with access to product-specific resources and expertise through a digital and data driven customer experience.

Sign into the HPE Support Center experience, featuring streamlined self-serve case creation and management capabilities with inline knowledge recommendations. You will also find personalized task alerts and powerful troubleshooting support through an intelligent virtual agent with seamless transition when needed to a live support agent.

<https://support.hpe.com/hpesc/public/home/signin>

Consume IT On Your Terms

HPE GreenLake edge-to-cloud platform brings the cloud experience directly to your apps and data wherever they are—the edge, colocations, or your data center. It delivers cloud services for on-premises IT infrastructure specifically tailored to your most demanding workloads. With a pay-per-use, scalable, point-and-click self-service experience that is managed for you, HPE GreenLake edge-to-cloud platform accelerates digital transformation in a distributed, edge-to-cloud world.

- Get faster time to market
- Save on TCO, align costs to business
- Scale quickly, meet unpredictable demand
- Simplify IT operations across your data centers and clouds

To learn more about HPE Services, please contact your Hewlett Packard Enterprise sales representative or Hewlett Packard Enterprise Authorized Channel Partner. Contact information for a representative in your area can be found at "Contact HPE"

<https://www.hpe.com/us/en/contact-hpe.html>

For more information

<http://www.hpe.com/services>



Configuration Information

Product SKUs and Ordering Experience

HPE Machine Learning Development Environment SKUs are software only.

HPE Machine Learning Development System is a complete turnkey solution including multiple hardware, software, and services SKUs.

Both standard and custom solutions have a 'starting' SKU. After selecting this SKU, the OCA wizard guides Solution Architects to build the full solution using other SKUs.

The starting SKU for standard solution is a hardware SKU that provides a pre-configured Apollo 6500.

The starting SKU for custom solution is a software SKU that provides Machine Learning Development Environment.

HPE Machine Learning Development System SKUs

Steps to Choose

Compute - Apollo 6500 Gen10+

- Minimum 4
- Maximum 120

Description

Software

HPE Performance Cluster Manager 1 Node 3yr 24x7 Support Perpetual E-LTU

Q9V60AAE

Notes: 3 years

Red Hat Enterprise Linux for HPC Compute Node 3yr Subscription E-LTU

R1P41AAE

Notes: 3 years

SUSE Manager Lifecycle Management 1-2 Sockets or 1-2 VM 1-year 24x7 E-LTU

R8V86AAE

Notes: 3 years (optional)

Management Stack

HPE ProLiant DL325 Gen10 Plus v2 8SFF Configure-to-order Server

P38471-B21

HPE Aruba Networking CX 6300M 48p 1GbE 4p SFP56 Power-to-Port 2 Fan Trays 1 PSU Bundle

JL762A

Red Hat Enterprise Linux for HPC Compute Node 3yr Subscription E-LTU

R1P41AAE

Compute Fabric

Mellanox InfiniBand HDR 40-port QSFP56 Managed Back to Front Airflow Switch

P06249-B21

Services

Tech Care Support

HPE 3Y Tech Care Essential Service

HU4A6A3

HPE 3Y Tech Care Essential with Defective Media Retention Service

HU4A7A3

HPE 3Y Tech Care Essential with Comprehensive Defective Material Retention Service

HU4A8A3

Complete Care Support

HPE 3Y Complete Care Addon Essential Service

HU4D5A3

HPE 3Y Complete Care Addon Essential with Defective Media Retention Service

HU4D6A3

HPE 3Y Complete Care Addon Essential with Comprehensive Defective Material Retention Service

HU4D7A3

Factory Express

HPE Factory Express Standard Unit of SVC

H4F41A1

HPE Factory Express Level 4 SVC

HA454A1

HPE FE Cluster Hig Den-Node HW Intg SVC

AC069A

HPE Startup Compute 1 Day SVC

HPE Technical Installation Startup SVC

HA124A1



Configuration Information

HPE Startup Compute 1 Day SVC

HPE Technical Installation Startup SVC

HA124A1

HPCM e-learning course (optional)

HPE Training Credit Servers Hybrid IT Service

HF385E/A1

Storage

HPE Parallel File System (optional) R7R35A (HDD), R7R36A (SSD)

Advisory and Professional Services

Additional HPE Services for customer ML/DL requirements (optional)

OCA Ordering Process (mandatory and optional components)

The HPE Machine Learning Development System OCA wizard allows users to build the entire solution.

Users can find the new HPE Machine Learning Development System using any of the following methods:

- Entering through the search box -> HPE Machine Learning
- Entering through the catalog -> Enterprise Software-> HPE Machine Learning

Within the OCA wizard, the user will have different tabs to navigate in a sequential order. Users will have two options for standard and custom offerings.

- Standard SKUs (R9K16A, R9K17A, R9K18A, R9K19A - minimum 4, maximum 120)
- Select one of the four standard SKUS which are based on Apollo XL675d Gen10 Plus. The standard offerings are pre-configured and contain A100 GPUs, AMD Milan processors, storage, and memory. The quantity of standard SKUs to be selected is based on the number of Apollo 6500 Gen 10 plus nodes desired. HPE Machine Learning Development Environment, HPCM and Red Hat Linux (compute/management nodes) software are included.
- Custom SKU (R9K20A – HPE Machine Learning Environment SW - minimum 32, maximum 960)
- Select the custom SKU which is just the software. This is followed up choosing the accelerated compute (a configurable XL675d Gen10 Plus). The quantity of custom SKU is based on the number of GPUs needed. HPE Machine Learning Environment and HPCM (compute/management nodes) software are included. Red Hat Linux or SUSE can be added through the menu view.
- Management Selection

Select the "Management Node/Fabric" tab and choose three to six DL325 Gen10 Plus v2 for the management node (The three nodes are used for: 1) login, 2) HPCM master, 3) Machine Learning Development Environment Master).

Fabric Selection

Select the "Management Node/Fabric" to define the InfiniBand Fabric. The OCA wizard calculates the minimum set of switches and cables. However, it is up to the user to make necessary adjustments on cabling to the specific layout needs of the end customer. To aid the customers, OCA has a link to an infrastructure matrix which is based on a 42U rack layout.

Storage Selection (optional)

The storage tab allows the user to add Parallel File System storage. For this release, storage must be configured in a separate order to be integrated on site.

Factory Express

Factory express services will automatically be added and can be viewed in the BOM. OCA users will need to manually fill the Customer Intent section to be able to complete a configuration.

Customer Intent

The Customer Intent information will be used by the integration center later in the process.



Configuration Information

Cabling

Cabling can be configured through the wizard. It is recommended that a solution architect reviews the cabling setup based on customer requirements. Additionally, included in the wizard is a cabling matrix for further reference. The cable matrix provides reference to cable types and lengths for compute, management, and network configurations.

Components View

Once users are in the Components view, users need to validate that each of the Compute Nodes, switches and management nodes are associated to the rack they want. Users can drag and drop components among different racks to fill the racks according to the end customer needs, matching the information filled in the Customer Intent form.

HPE Power Advisor will be available prior to configuring OCA and available to help with power requirements for your HPE Machine Learning Development system configuration. <https://poweradvisorexternal.it.hpe.com/?Page=Index>



Technical Specifications

Standard and Custom Solution Options

Feature	Standard	Custom
Number of compute nodes	4-120	4-120
CPU	2x AMD EPYC 7543 (64 cores @ 2.8-3.7GHz) OR 2x AMD EPYC 7763 (128 cores @ 2.45-3.5GHz)	Any AMD Milan that is Apollo 6500 Gen 10 Plus Compatible
Memory	2TB Or 4TB	Any Apollo 6500 Gen 10 Plus Compatible
Scratch storage	15 TB NVMe OR 30 TB NVMe	Any Apollo 6500 Gen 10 Plus Compatible
GPU	NVIDIA HGX A100 System 80GB Tensor Core GPU with NVLink	NVIDIA HGX A100 System 80GB Tensor Core GPU with NVLink
Number of storage nodes	4-128	4-128
Number of management nodes	3	6 for High Availability
Fabrics for compute	Mellanox InfiniBand HDR; Full bandwidth network topology	Mellanox InfiniBand HDR; Customizable topology
OS	RHEL	RHEL, SUSE, Ubuntu(roadmap)
Service and Solution Support	Core Services + Optional HPE A&PS	Core Services + Optional HPE A&PS



Technical Specifications

For illustrative purposes, here are two standard example configurations

Configuration	“Small”	“Medium”
Compute node	4x HPE Apollo 6500 Gen 10 Plus	20x HPE Apollo 6500 Gen 10 Plus
CPU	2x AMD EPYC 7543	2x AMD EPYC 7763
GPU	8x NVIDIA A100 80GB with NVLink	8x NVIDIA A100 80GB with NVLink
Storage node	4x HPE PFSS nodes	8x HPE PFSS nodes
Management node	3x HPE DL325 Gen 10 Plus v2	3x HPE DL325 Gen 10 Plus v2
Fabrics for compute	1x Mellanox InfiniBand HDR switch	8x Mellanox InfiniBand HDR switch
Management Fabric	2x HPE Aruba Networking CX 6300M Gbe switch	2x HPE Aruba Networking CX 6300M Gbe switch
Training Platform	4X HPE Machine Learning Development Environment Standard SKU	20X HPE Machine Learning Development Environment Standard SKU
Cluster manager	7 x HPCM 3-year license	23 x HPCM 3-year license
OS	7 x RHEL 3-year license	23 x RHEL 3-year license
Number of racks	1X	5X
Number of cables	19X	173X
Theoretical FLOPs	9.6 PFLOPs (AI/fp16); 632 TFLOPs (fp64)	48 PFLOPs (AI/fp16); 3.2PFLOPs (fp64)
Estimated power consumption	24 kW	120 kW
Support	Essential Care	Essential Care
Startup Service	HPE startup 1-day workshop	HPE startup 1-day workshop
Deployment Service	Factory Express (Level 4)	Factory Express (Level 4)



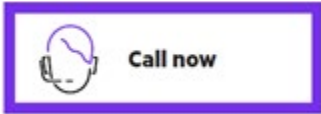
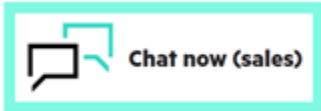
Summary of Changes

Date	Version History	Action	Description of Change
06-May-2024	Version 5	Changed	Standard Features section was updated. Obsolete SKU was removed.
19-Feb-2024	Version 4	Changed	Networking product names were updated.
08-Jan-2024	Version 3	Changed	Overview, Standard Features, Service and Support sections were updated. Obsolete SKUs were removed. HPE Services Rebranding
06-Jun-2022	Version 2	Changed	Added Number of Racks, Cables, FLOPs, Power consumption columns. Technical Specifications section was updated.
27-Apr-2022	Version 1	New	New QuickSpecs



Copyright

Make the right purchase decision.
Contact our presales specialists.



© Copyright 2024 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™ and EPYC™ are registered trademarks of Advanced Micro Devices, Inc. in the U.S. and other countries. Microsoft®, Windows®, and Windows Server® are U.S. registered trademarks of the Microsoft group of companies.

For hard drives, 1GB = 1 billion bytes. Actual formatted capacity is less

a50004279enw - 16883 - Worldwide - V5 - 06-May-2024