Overview

HPE Machine Learning Development Environment Software

Built upon the open-source **Determined** model development platform, **HPE Machine Learning Development Environment Software** allows model developers and researchers to focus on building better models faster, by reducing complexity, and removing the need to write boilerplate code associated with managing ML infrastructure.

It easily integrates with popular ML **frameworks** & **tools**, and supports cloud or on-prem **infrastructure** environments, with a consistent user experience (UX).

Our platform also makes it easy for IT and MLOps teams to setup and share AI infrastructure to improve collaboration & productivity for ML teams, and reduce costs.

With HPE Machine Learning Development Environment Software, you can easily:

- Train models faster
- Build better models
- Manage and share your AI infrastructure
- Track and reproduce your work
- Integrate into your enterprise
- Bring your own cloud
- Evaluate and customize large language models (LLMs)

What's New

- Generative AI studio is a standard feature now available with HPE Machine Learning Development Environment. Generative AI studio is the easiest, and most secure way to build and scale generative AI applications on-premises using foundational models.
- Run ML and HPC jobs alongside each other on the same cluster, with support for workload managers like **Slurm® or PBS®**, and secure container runtimes like Singularity/Apptainer, Podman, or NVIDIA® Enroot.
- Use Role-Based Access Controls (**RBAC**) to authorize development and MLOps teams to securely collaborate and share ML resources and artifacts.
- Train models on NVIDIA or AMD® GPUs without any code changes, with foundational support for accelerator heterogeneity.
- Use **DeepSpeed®** for 3D-Parallel (Data-, Model-, and Pipeline-parallel) distributed training, to speed up training of large models like **GPT-NeoX**.
- Use PyTorch® Distributed Data Parallel (DDP) for flexibility and choice of distributed training strategies.
- Version, annotate, and organize trained models in the Model Registry so that MLOps teams can effectively collaborate with model developers to manage your models' lifecycle.
- Define your own logic to coordinate across multiple trials within an experiment.
- Implement your own **custom hyperparameter search** algorithms, ensembling, active learning, neural architecture search, and reinforcement learning.
- New SKUs
  - R9H29AAE - HPE ML Dev Env SW 3yr Sub E-RTU
  - R9Y51AAE - HPE ML Dev Env SW 4yr Sub E-RTU
  - R9Y52AAE - HPE ML Dev Env SW 5yr Sub E-RTU
  - S2E65AAE - HPE ML Dev Env SW 1-8GPU 1yr Mngd E-RTU
  - S2E66AAE - HPE ML Dev Env SW 8+ GPU 1yr Mngd E-RTU

Standard Features

Train Models Faster

Train models faster, and at-scale, without changing your model code, or writing cumbersome boilerplate to manage your ML infrastructure.

We take care of ML infrastructure provisioning, networking, data loading, checkpointing, fault-tolerance and high-availability - i.e., we make distributed model training easy, fast, and cost-efficient.

HPE Machine Learning Development Environment Software ships cutting-edge **distributed training** strategies and techniques:

- **DeepSpeed**: for 3D-Parallel (Data-, Model-, and Pipeline-parallel) distributed training, to speedup training of large models like **GPT-NeoX**, and beyond!
- **Horovod**: for easy-to-use data-parallel distributed training.
- PyTorch Distributed Data Parallel (**DDP**): for flexibility and choice of distributed training strategies.

Build Better Models

Find better model configurations efficiently with a production-grade implementation from the creators of the cutting-edge Asynchronous Successive Halving (**ASHA**) Hyperband for **hyperparameter tuning**.

Define your own logic to coordinate across multiple trials within an experiment.

Implement your own custom hyperparameter search algorithms, ensembling, active learning, neural architecture search, and reinforcement learning methods.

Efficiently Manage and Share Your AI Infrastructure

Efficiently manage and share your on-prem or cloud GPUs & accelerators with Machine Learning workflow-aware smart scheduling and resource management to improve productivity and collaboration for your ML development and operations teams.

Consistent User Experience for deployments ranging from laptop to a supercomputer scale, and everything in-between:

- Baremetal
- Virtual Machine (incl. cloud-native and on-prem Infrastructure-as-a-Service solutions)
- Kubernetes®
- Slurm
- PBS

Run ML and HPC jobs alongside each other on the same cluster, with support for workload managers like Slurm or PBS, and secure container runtimes like Singularity/Apptainer, Podman, or NVIDIA® Enroot.

Seamlessly use spot or preemptible instances to manage cloud costs.

Train models on NVIDIA or AMD® GPUs without any code changes, with foundational support for accelerator heterogeneity.

Track and Reproduce Your Work

Easily track and reproduce your work with experiment tracking that works out-of-the-box: covering model code, configuration, hyperparameters, metrics, and checkpoints.

Version, annotate, and organize trained models with our built-in Model Registry, enabling MLOps teams can effectively collaborate with model developers to manage your models' lifecycle.

Standard Features

Integrate Into Your Enterprise

Authenticate users using enterprise Single Sign-On (SSO) services provided by Active Directory®, Okta®, PingID®, etc., with support for OpenID Connect (OIDC) and SAML.

Integrate with user provisioning systems to automate the onboarding and offboarding of your teams, with support for SCIM.

Use Role-Based Access Controls (RBAC) to authorize development and MLOps teams to securely collaborate and share ML resources and artifacts.

Bring Your Own Cloud

Using HPE Machine Learning Development Environment Software as a managed service, deploy the MLDE core platform within your cloud of choice, whether that is AWS or GCP.

Supported environments

- AWS using EC2
- GCP using GKE

Supported Hardware

HPE Machine Learning Development Environment Software can be **deployed** on hardware equipped with NVIDIA or AMD GPUs, on a variety of on-prem or cloud environments.

Hardware

- The master node should be configured with at least four Intel Broadwell or later CPU cores, 8 GB of RAM, and 200 GB of free disk space. The master node does not need GPUs.
- Each GPU-equipped compute node should be configured with at least two Intel Broadwell or later CPU cores, 4 GB of RAM, and 50 GB of free disk space.
  - NVIDIA GPUs from Hardware Generation Volta (Compute Capability version 7) or newer are supported - e.g., V100, A100, H100, H200, or newer. See Hardware Generation https://docs.nvidia.com/deploy/cuda-compatibility/index.html#frequently-asked-questions
  - AMD GPUs from Compute DNA ("CDNA") version 2 or higher are supported - e.g., MI210, MI250, MI250X, MI300X, MI300A, or newer. see **https://www.amd.com/en/technologies/cdna.html**

Notes: Most of the disk space required by the master is for the experiment metadata database. If PostgreSQL is set up on a different machine, the disk space requirements for the master are minimal.

Supported Operating Systems and Platforms

- Red Hat® Enterprise Linux (RHEL®); SUSE® Linux Enterprise Server (SLES®); Ubuntu®

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/completecare

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

HPE Machine Learning Development Environment Software Products

HPE Machine Learning Development Environment Software is licensed on a per-GPU basis for the duration of the stated subscription term.

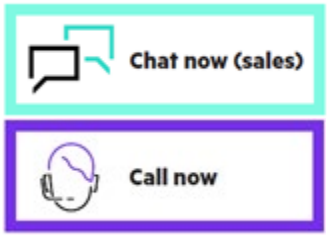
Description	SKU
HPE ML Dev Env SW 1-19 GPU 1yr Sub E-RTU	R8W23AAE
HPE ML Dev Env SW 20-99 GPU 1yr Sub E-RTU	R8W26AAE
HPE ML Dev Env SW 100+ GPU 1yr Sub E-RTU	R8W27AAE
HPE ML Dev Env SW 3yr Sub E-RTU	R9H29AAE
HPE ML Dev Env SW 4yr Sub E-RTU	R9Y51AAE
HPE ML Dev Env SW 5yr Sub E-RTU	R9Y52AAE
HPE ML Dev Env SW 1-8GPU 1yr Mngd E-RTU	S2E65AAE
HPE ML Dev Env SW 8+ GPU 1yr Mngd E-RTU	S2E66AAE

Summary of Changes

Date	Version History	Action	Description of Change
03-Sep-2024	Version 5	Changed	Standard Features section was updated.
03-Jun-2024	Version 4	Changed	Overview and Configuration Information sections were updated.
18-Dec-2023	Version 3	Changed	HPE Services Rebranding
05-Sep-2023	Version 2	Changed	Overview, Standard Features and Configuration Information sections were updated.
06-Mar-2023	Version 1	New	New QuickSpecs

Copyright

Make the right purchase decision.
Contact our presales specialists.



Get updates

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

a50006978enw - 17088 - WorldWide - V5 - 03-September-2024