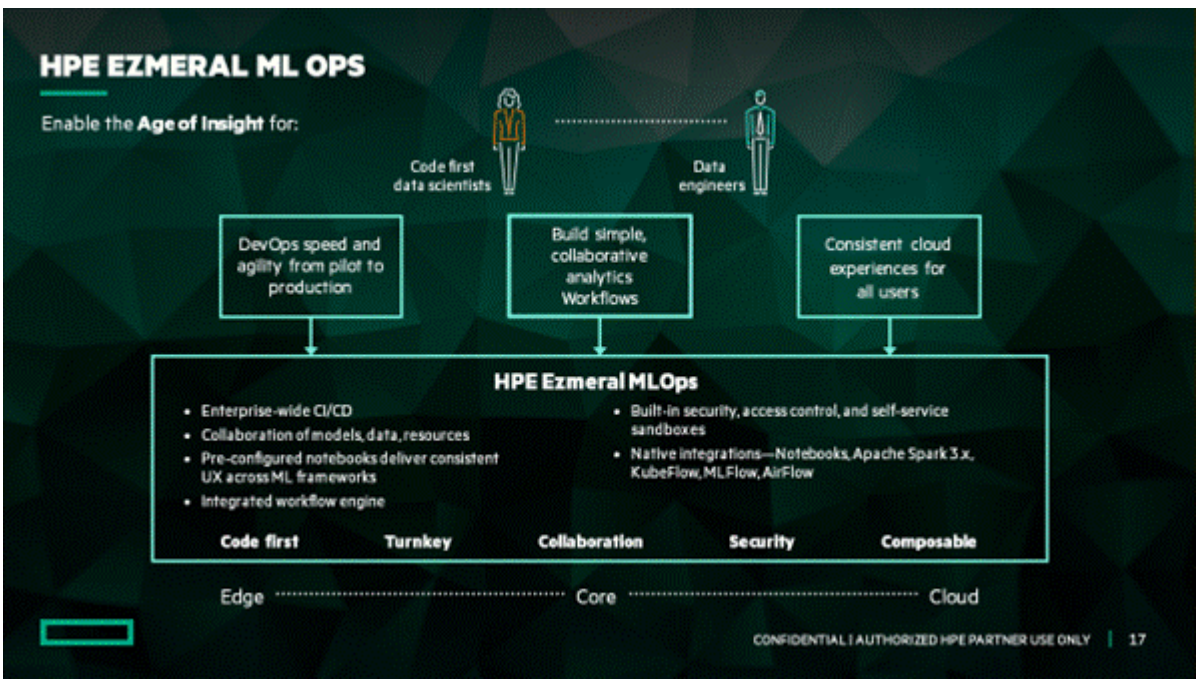
HPE Ezmeral Machine Learning (ML) Ops

In most organizations today, machine learning projects do not have the standard processes and rigor that we now associate with software development. The continuous integration/continuous deployment (CI/CD) workflows have become ubiquitous for all software development today. On the machine learning front, data scientists still spend a significant amount of time and effort when moving projects from development to production. Model version control is still manual, making it hard to update models in production. Code sharing is manual; data copied onto local storage leading to variability of results between environments. There is also the lack of standardization on tools and frameworks, which makes it tedious and time-consuming to ensure accuracy of predictions across all environments.

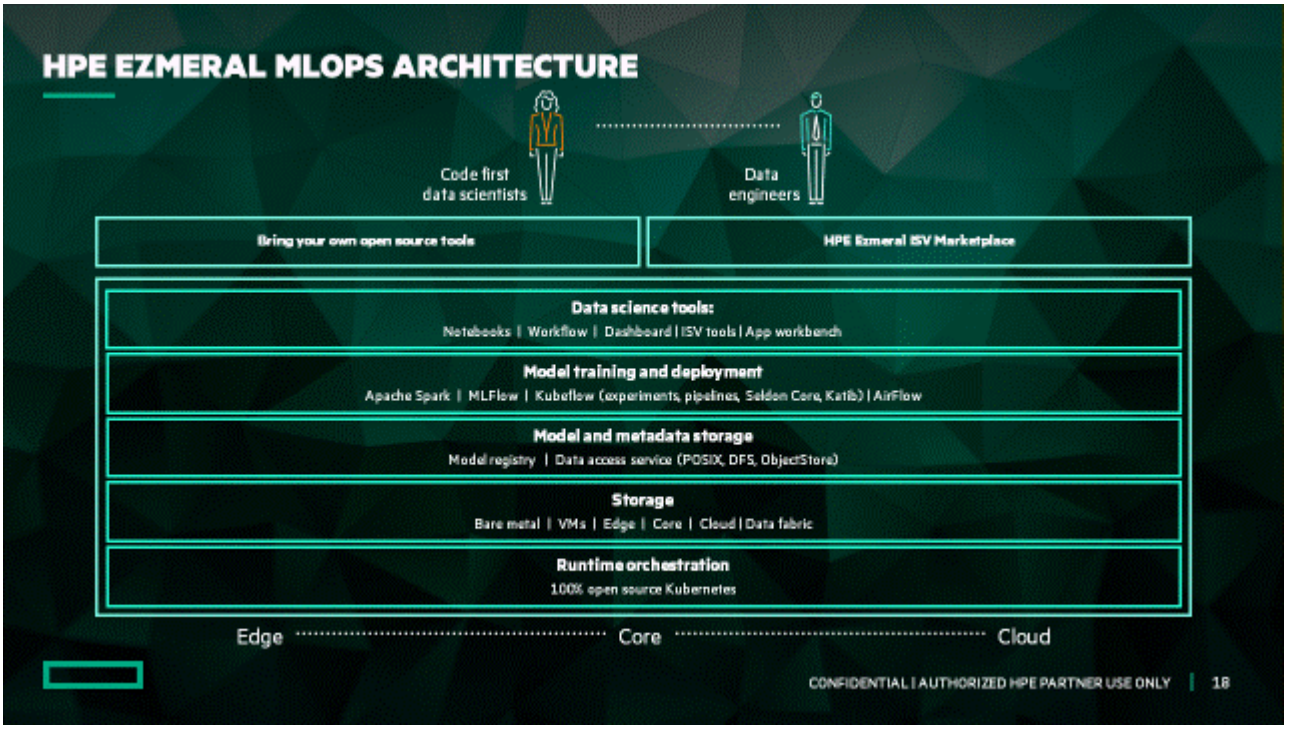
HPE Ezmeral Machine Learning Ops software (HPE Ezmeral ML Ops) leverages the functionality of HPE Ezmeral Runtime Enterprise as a secure, scalable enterprise platform and adds DevOps-like speed and agility to enterprise machine learning. With HPE Ezmeral ML Ops, enterprises can implement CI/CD workflows and standardize their ML pipelines. The HPE Ezmeral ML Ops software platform supports every stage of the machine learning lifecycle, from development, to training, deployment and monitoring - supporting sandbox experimentation with your choice of ML/DL frameworks, integrations with model and code repositories, to deploying and tracking models in production.

HPE Ezmeral ML Ops gives data scientists and developers the ability to quickly and easily build and train machine learning models. HPE Ezmeral ML Ops allows data scientists to manage and track models built on any platform and deploy them into a scalable and secure production environment. Using HPE Ezmeral ML Ops, data scientists can spin-up containerized environments for distributed data processing, Machine Learning (ML), or Deep Learning (DL) in minutes rather than weeks. HPE Ezmeral ML Ops provides data science teams the flexibility to run their ML/DL workloads either on-premises, in multiple public clouds, or a hybrid model and respond to dynamic business requirements in a variety of use cases.

With HPE Ezmeral ML Ops software, Hewlett Packard Enterprise is making it easier for organizations to deliver a flexible and secure multi-tenant architecture, with the agility, flexibility, speed and performance needed to address evolving workload and application requirements. You can deploy HPE Ezmeral ML Ops using pre-tested and optimized HPE Apollo building blocks on-premises, as well as in hybrid IT architectures and in a multi-cloud model.



Standard Features



The HPE Ezmeral ML Ops Platform Architecture

HPE Ezmeral ML Ops contains the following new feature enhancements with the version 5.4 release:

Product	Feature
HPE Ezmeral ML Ops	1. KubeFlow 1.4 & Kale add-on 2. Enhanced UX: improved UX for data scientists 3. Ezmllib: Packaged libraries for simplified coding experience 4. Improved error handling in Notebook magics for KubeFlow

- **The HPE Ezmeral ML Ops architecture consists of the following components delivered in the previous 5.x releases: Apache Spark 3.2 integration** - We've expanded our capabilities and support of the most popular open-source data analytics and engineering offering. With Apache Spark operator natively integrated, we now allow our clients to better leverage GPUs and directly access data from S3. With Livy integration, we enable users to simplify job submission from external locations like a Juypter notebook or terminal. Users can also leverage the Airflow workflow engine to manage data processing and data science pipelines.
- **Simplified ML Ops collaboration for multi-tenant environments** - With out-of-the-box MLFlow integration, we now have a global shared common metadata store with experimentation and model management capabilities. Global shared repository enables sharing of model metadata and related artifacts across tenants.
- **Simplified App Modernization** - The new version of the App Workbench makes it easier to bring your own Kubernetes app. Using an enhanced GUI, you can create your own custom images with just a few clicks. Additionally, our new HPE Ezmeral Ecosystem program with 16 new HPE Ezmeral validated ISVs to streamline the addition of new applications.
- **Simplified policy management** - Our centralized policy management allows clients to adhere to corporate standard policies with the OPA (Open Policy Agent) rule engine and apply the policies across tenants and clusters. We've also added policy drift management, which automatically detects and reports of out of compliance clusters.
- **Runtime Security out-of-the-box with HPE Ezmeral Runtime Enterprise** - Falco integration

Standard Features

provides anti-virus like security management for runtime containers.

HPE Ezmeral ML Ops delivers the following new AI/ML functionality in addition to the HPE Ezmeral Runtime Enterprise functionality

Leverages the power of containers to create complex machine learning and deep learning stacks that include distributed TensorFlow, Apache Spark, H2O, and Python ML and DL toolkits.

- Spins-up distributed, scalable, machine learning, and deep learning training environments in minutes rather than months - on-premises, public cloud, or in a hybrid model.
- Enables a choice of programming languages and open-source tools to support even the most complex ML pipelines. For example, start with data pre-processing in Spark with Scala, followed by model development with TensorFlow on GPUs, and finally model deployment on CPUs with TensorFlow runtime.
- Implements CI/CD processes for your ML projects with a model registry. Model registry stores models and versions created within HPE Ezmeral ML Ops, as well as those created using different tools/platforms.
- Improves the reliability and reproducibility of machine learning projects a shared source control repository (GitHub & BitBucket).
- Enables the deployment of models in production with secure, scalable, highly available endpoint deployments with out-of-the-box auto-scaling, and load balancing.
- Allows data scientists to focus on the core task of building ML models to improve business outcomes rather than managing infrastructure.
- Enables Out-of-the-box application images to rapidly deploy containerized environments - sandbox, distributed training, or serving (inferencing) - with popular ML and DL tools, interfaces and languages. Such as Python, R-Studio, TensorFlow, Spark, and more.
- Enables the creation of custom application images with any combination of tools, library packages, and frameworks to suit your needs.

For more information about the HPE Apollo Systems for HPE Ezmeral Runtime Enterprise, refer to the QuickSpecs at:
https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00062186enw

For more information about the HPE Ezmeral Runtime Enterprise, refer to the QuickSpecs at:
https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00056658enw

---

Features and Benefits

HPE Ezmeral ML Ops addresses the entire machine learning pipeline from the initial construction of models to data preparation, model training and validation, deployment and monitoring, all while offering flexibility, scale and security of data and models. HPE Ezmeral ML Ops delivers the following new features and benefits:

Model Building - Pre-packaged, self-service sandbox environments

With HPE Ezmeral ML Ops, Data Scientists can now quickly spin-up environments with their preferred data science tools-such as TensorFlow, Apache Spark, Keras, PyTorch and more-to explore a variety of enterprise data sources and simultaneously experiment with multiple machine learning or deep learning frameworks to pick the best fit model for the business problems they are addressing.

Model Training - Scalable training environments with secure access to big data

Data science teams can create their own containerized environments - single node or distributed multi-node clusters - for development & test or production workloads on-demand within minutes.

Patented innovations provide highly performant training environments that can securely access shared enterprise data sources in on-premises or cloud-based storage. Independently scale compute and storage on an as-needed basis to improve efficiency and reduce costs.

Standard Features

Model Deployment - Flexible, scalable, endpoint deployment
HPE Ezmeral ML Ops deploys the model's native runtime image (e.g. Python, R, H2O, etc) into a containerized endpoint. Integration with the model registry allows data scientists to track model versions and seamlessly update models when needed.
HPE Ezmeral ML Ops provides out-of-the-box model serving environments to create HTTP endpoints with tokens.

The HPE Ezmeral Runtime Enterprise gateway provides high availability and load balancing for model endpoints. Each deployment can have multiple runtime engines for handling scoring logic. HPE Ezmeral ML Ops' auto scaling dynamically scales nodes for scoring engines.

Model Monitoring - End-to-end visibility across the ML pipeline
Complete visibility into runtime resource usage such as GPU, CPU, and memory utilization. Track, measure, and report model performance, save and inspect inputs and outputs for each scoring request. 3rd party integrations track accuracy and interpretability

---

Collaboration - Enable CI/CD workflows with code, model, and project repositories
HPE Ezmeral ML Ops provides out of the box integration with GitHub for organizations that choose to use GitHub as their project/code repository and provides source control, eases collaboration, and enables lineage tracking for improved auditability. This enables storing multiple models (multiple versions with metadata) for various runtime engines in the model registry. Run A/B testing or Canary testing to validate the model before large-scale deployment. HPE Ezmeral ML Ops provides an NFS-based project repository that eases collaboration and provides lineage tracking which improves model auditability

---

Security and Control - Secure multi-tenancy with integration to enterprise authentication mechanisms
HPE Ezmeral ML Ops software provides multi-tenancy and data isolation to ensure logical separation between each project, group, or department within the organization. The platform integrates with enterprise security and authentication mechanisms such as LDAP, Active Directory and Kerberos. Different project teams, groups, or departments across the enterprise can share the same infrastructure and access the same data sources for their AI / ML and Big Data analytics workloads

---

Hybrid Deployment - On-premises, public cloud, or hybrid
Run the HPE Ezmeral ML Ops software on-premises on any infrastructure (including in multiple data centers), on multiple public clouds (Amazon® Web Services, Google® Cloud Platform, or Microsoft® Azure), or in a hybrid model, providing effective utilization of resources and lower operating costs. It is important with public cloud and hybrid deployments that special attention be given to network requirements. HPE Ezmeral ML Ops supports all the major cloud providers.

In addition to the new features and business benefits delivered through the HPE Ezmeral ML Ops software, the underlying functionality delivered previously via HPE Ezmeral Runtime Enterprise will continue to be part of the overall HPE Ezmeral Runtime Enterprise framework and integrated with the new HPE Ezmeral Machine Learning Ops software. Those sustaining benefits are outlined below.

Self-Service Environments
Users can get up and running quickly with HPE Ezmeral Runtime Enterprise ElasticPlane functionality. New containerized environments are provisioned on-demand with just a few mouse clicks-whether they're transient for development and testing, or long-running for a production workload. Data scientists and analysts can now quickly respond to dynamic business requirements for a variety of use cases ranging from deep learning with AI Frameworks like TensorFlow to analytical SQL on Hadoop.

Containers and Bare-Metal Performance
HPE Ezmeral Runtime Enterprise has developed patented I/O optimization innovations to deliver the agility

Standard Features

benefits of Docker containers for AI and Big Data, while ensuring performance comparable to that of bare-metal servers. HPE Ezmeral Runtime Enterprise's IOBoost functionality provides application-aware caching and elastic resource management that adapts dynamically to changing application requirements, delivering the best possible performance.

Flexibility for Tools of Choice

HPE Ezmeral Runtime Enterprise software offers pre-integrated container images, with many of the most common AI and Big Data tools. You can easily spin up Hortonworks/Cloudera clusters, create standalone Spark and Kafka environments, deploy machine learning with TensorFlow or H2O, while installing your choice of data science tools like Jupyter or Zeppelin notebooks. You can even add your own preferred tools and quickly upgrade to new versions-providing the ultimate in agility and configurability.

Compute and Storage Separation

HPE Ezmeral Runtime Enterprise disconnects analytical processing from data storage, giving you the ability to independently scale compute and storage based on the needs of your workloads. With HPE Ezmeral Runtime Enterprise, you have unparalleled flexibility to mix and match infrastructure. A unique set of "Select" servers, including density optimized, HPE Apollo compute and storage nodes, Synergy compute nodes, EdgeLine and ProLiant based compute nodes, offer building blocks that are configured specifically for HPE Ezmeral Runtime Enterprise. These predefined configurations can be found in the OCA HPE Ezmeral Runtime Enterprise Wizard. When installing HPE Ezmeral Runtime Enterprise on the HPE Apollo r2600 (XL170r and XL190r), HPE Apollo 4200 and HPE Apollo 6500 (XL270d), you get the benefit of one of the most scalable and powerful platforms on the market. HPE's Elastic Platform for Analytics effectively enables the separation of compute and storage, making it the ideal infrastructure for HPE Ezmeral Runtime Enterprise software.

Data Access from Any Storage

With HPE Ezmeral Runtime Enterprise's DataTap capability, you can access data from any shared storage system (including HDFS as well as NFS) or cloud storage (e.g. Amazon S3). This means you don't need to make multiple copies of data or move data before running your analysis. Sensitive data can stay in your secure storage system with enterprise-grade data governance, without the cost and risks of creating and maintaining multiple copies or moving large-scale data.

Flexible App Store

The HPE Ezmeral Runtime Enterprise software platform includes an "App Store" for common Analytics / Machine Learning / Deep Learning tools, distributed computing frameworks, and data science notebooks. Open-source distributions for Spark, Kafka, Hadoop, and other frameworks as well as representative Machine Learning and Analytics applications are provided as pre-configured Docker images in the App Store and available via one-click deployment. A "bring your own app" model is also supported allowing our customers to create their own Docker images using HPE Ezmeral Runtime Enterprise templates.

Several applications come pre-configured in the HPE Ezmeral Runtime Enterprise App Store and can be installed via one-click deployment to accelerate time to business value. These sample Docker images provide a particular version and starting configuration for popular distributions, applications, and tools. With our App Workbench, you can populate your App Store with images for other versions and additional tools.

For more information about the HPE Apollo Systems for HPE Ezmeral Runtime Enterprise, refer to the QuickSpecs at:
https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00062186enw

For more information about the HPE Ezmeral Runtime Enterprise, refer to the QuickSpecs at:
https://h20195.www2.hpe.com/v2/getdocument.aspx?docname=a00056658enw

Standard Features

HPE Ezmeral Runtime Enterprise and HPE Ezmeral ML Ops Capabilities Matrix

	HPE Ezmeral Runtime Enterprise Essentials	HPE Ezmeral Runtime Enterprise	HPE Ezmeral MLOps
ECP Management Plane	Yes	Yes	Yes
Managed Gateway	Yes	Yes	Yes
Authentication Proxy	Yes	Yes	Yes
Air-gap Support	Yes	Yes	Yes
Logging (Elastic Search)	Yes	Yes	Yes
Metrics (Elastic Search)	Yes	Yes	Yes
Alerts (Nagios)	Yes	Yes	Yes
AD/LDAP/SAML/OIDC	Yes	Yes	Yes
Host Management	Yes	Yes	Yes
Kubernetes multi version	Yes	Yes	Yes
Kubernetes Dashboard	Yes	Yes	Yes
Kubernetes Upgrade	Yes	Yes	Yes
Container Network Interface (CNI)	Yes	Yes	Yes
Service Mesh (Istio)	Yes	Yes	Yes
Tenant Management	Yes	Yes	Yes
Data Tap (DTAP)	Yes	Yes	Yes
Tenant share (FS Mount), tenant storage	Yes	Yes	Yes
Operating Systems (OS)	Yes	Yes	Yes
RHEL OS	Yes	Yes	Yes
SLES OS	Yes	Yes	Yes
GPU support for Kubernetes hosts	Yes	Yes	Yes
Storage		Yes	Yes
Persistent Storage (HPE Ezmeral Data Fabric with CSI)		Yes	Yes
Global Storage		Yes	Yes
Web Terminal		Yes	Yes
Policy Management		Yes	Yes
Runtime Security (Falco)		Yes	Yes
User K8s Cluster Backup and Recovery		Yes	Yes
Stateful Applications with KubeDirector		Yes	Yes
Import External Cluster (for additional fee)		Yes	Yes
Ezmeral Analytics for Spark (for additional fee with HPE Ezmeral Runtime Enterprise)		Yes	Yes

Standard Features

Spark operator (Driver and Executor)		Yes	Yes
History server		Yes	Yes
Livy server		Yes	Yes

	HPE Ezmeral Runtime Enterprise Essentials	HPE Ezmeral Runtime Enterprise	HPE Ezmeral MLOps
Spark job submission and monitoring UI		Yes	Yes
Jupyter Notebook (KubeDirector app)		Yes	Yes
Airflow Workflow Engine		Yes	Yes
Spark Thrift Server		Yes	Yes
Machine Learning Capabilities			Yes
KubeFlow			Yes
Airflow			Yes
MLFlow			Yes
KubeDirector Apps			Yes
– Notebook			Yes
– Training			Yes
– Deployment			Yes
– Tensorflow App			Yes

Service and Support

Consume IT on your terms

HPE GreenLake 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 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

Managed services to run your IT operations

HPE GreenLake Management Services provides 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 Ezmeral Platform Services

HPE Ezmeral Deployment & Integration Services coordinates the app images planning, design, development, configuration, and validation of the customers new to HPE Ezmeral Platform solutions. During the pandemic, services are being delivered remotely with access to the customer's environment or collaboratively with the customer's personnel, to help the customer quickly leverage the features and benefits of the HPE Ezmeral Platform. After purchase, an HPE Ezmeral services specialist is assigned to work with the customer to help deploy and integrate their new HPE Ezmeral Software with the Customer's IT infrastructure. In addition, the HPE Ezmeral services specialist also works with the customer to provide valuable knowledge transfer that can help them quickly take ownership of the new HPE Ezmeral solution.

Ezmeral Training

HPE data and analytics training delivers the skills and expertise for success with Big Data, containers, AI/ML and deep learning-including HPE Ezmeral and HPE Data Fabric. Live virtual instructor-led and eLearning courses are available.

- **Data and Analytics Courses**
- Data and Analytics Training from HPE brochure

Configuration Information

Ordering Information

HPE Ezmeral ML Ops Subscription Options

Ezmeral SKUs are available in HPE's new BRIM subscription ordering system as well as the traditional S4 ordering system. Please work with the **Ezmeral Deal Desk** to determine which is the best system for placing your order.

BRIM SKU

Description	SKU
HPE Ezmeral Runtime Enterprise Essentials E-LTU	R9J34AAE
HPE Ezmeral Runtime Enterprise E-LTU	R9J35AAE
HPE Ezmeral Runtime Enterprise Analytics for Apache Spark E-LTU	R9J38AAE
HPE Ezmeral Runtime Enterprise Imported Cluster Manager E-LTU	R9J39AAE

S4 SKU

HPE Ezmeral Machine Learning Ops 1-year 24x7 E-LTU	S0M18AAE
HPE Ezmeral Machine Learning Ops 2-year 24x7 E-LTU	S0M19AAE
HPE Ezmeral Machine Learning Ops 3-year 24x7 E-LTU	S0M20AAE
HPE Ezmeral Machine Learning Ops 4-year 24x7 E-LTU	S0M21AAE
HPE Ezmeral Machine Learning Ops 5-year 24x7 E-LTU	S0M22AAE

Notes:

- Each subscription SKU includes a term license to use on one core with up to 2 terabytes of HPE Ezmeral Data Fabric File and Object Store per core, plus the associated 24x7 HPE Pointnext Tech Care support.
- HPE Ezmeral Runtime Enterprise and HPE Ezmeral ML Ops software is licensed by the number of unique cores available to the kernel in the OS on which the HPE Ezmeral Runtime software is directly installed, regardless of the number of threads in each core.
- All servers, virtual and physical, require licensing with the exception of HPE Ezmeral Runtime Enterprise Gateway servers.

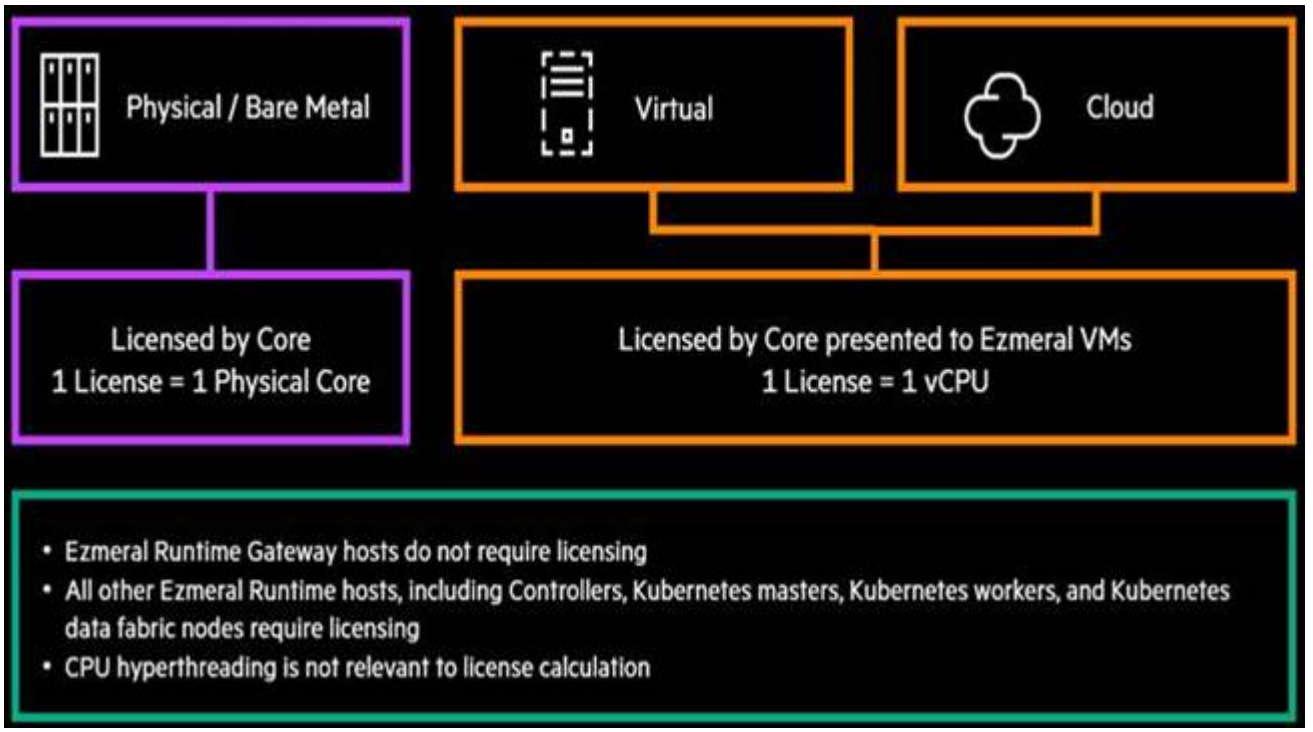
HPE Pointnext Tech Care

HPE Pointnext Tech Care is the operational service experience for HPE products. Tech Care 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 Pointnext Tech Care has been reimagined from the ground up to support a customer-centric, AI driven, and digitally enabled customer experience to move your business forward.

HPE Ezmeral products are covered by HPE Pointnext Tech Care Essentials, providing 24x7 global support coverage.

Configuration Information

HPE Ezmeral Runtime Enterprise Customer Scenario



HPE Ezmeral Runtime Enterprise Imported Cluster Manager

HPE Ezmeral Enterprise Imported Cluster Manager offers the ability for HPE Ezmeral Runtime Enterprise to manage external Kubernetes clusters running on Amazon EKS, Google Anthos (GKE), Microsoft Azure Kubernetes Service (AKS), VMware Tanzu PKS, and Cloud Native Computing Foundation (CNCF) K8s, in addition to clusters deployed on-premises.

HPE Ezmeral Runtime Enterprise Clusters spun up externally are registered and managed by HPE Ezmeral Runtime Enterprise. These are add-on SKUs with the pre-requisite of having HPE Ezmeral Runtime Enterprise or HPE Ezmeral ML Ops installed.

Total Server CPU Cores	HPE Ezmeral ML Ops Minimum Subs	HPE Ezmeral ML Ops Maximum Subs (Comprehends additional 30% for Virtual CPU's)
16	16	20
20	20	24
24	24	32
28	28	36
32	32	40
36	36	44
40	40	52
44	44	56
48	48	62

Notes:

- The following end user information is required at time of order to receive the electronic license:
  - End user organization name
  - End user organization address
  - End user contact name
  - End user email address
  - Reseller and distributor organization name and address (if applicable)

Summary of Changes

Date	Version History	Action	Description of Change
06-Sep-2022	Version 8	Changed	Standard Features section was updated
15-Aug-2022	Version 7	Changed	Service and Support and Configuration Information sections were updated.
01-Aug-2022	Version 6	Changed	Standard Features and Configuration Information sections were updated.
02-May-2022	Version 5	Changed	Updated all references from HPE Ezmeral Container Platform to HPE Ezmeral Runtime Enterprise. Updated architecture diagram. Updated messaging to align with updated HPE Ezmeral ML Ops Solution Brief.
15-Sep-2021	Version 4	Changed	Overview and Configuration Information sections were updated.
06-Apr-2021	Version 3	Changed	Overview, Standard Features and Service and Support section were updated.
03-Aug-2020	Version 2	Changed	QuickSpecs content and layout were updated.
10-Sep-2019	Version 1	New	New QuickSpecs

Copyright

a00067742enw - 16428 - Worldwide - V8 - 06-September-2022