

### Overview

## HPE Ezmeral Data Fabric

### Market Overview

The power of analytics, artificial intelligence (AI), and machine learning (ML) are helping organizations become data-driven to attract, retain, and grow their customers case and drive innovation. To be successful, consumers of these technologies need direct access to large volumes of different data types, systems, and solutions. But isolated systems are restricting their access to data, which results in slower business innovation.

#### The challenge of data integrity

Data is stored across different locations and in different formats. It lives in public clouds, at the edge, and in co-location or on-premises data centers. Each of these locations has specific access methods and processes making it difficult for data teams to deliver unified and normalized data fast enough for real-time insights.

These data islands are not going away, so industry analysts agree that a unified data layer is required for increased accuracy, confidence, and trust in the processing outcomes. <sup>1</sup>What's needed is a new approach. One that delivers unified data faster by securely working around data islands.

That new approach is known as a data fabric.

**Notes:** <sup>1</sup>451 Research, 2022; "Predictions 2022: Data and Analytic Strategies Build Trust and Accelerate Decision Making," Gartner, 2021; "Enterprise Data Fabric Enables DataOps," Forrester, 2021

### Product Overview

HPE Ezmeral Data Fabric simplifies data access patterns, acquisition processing, and surfacing of different data types across multiple use cases, workloads, and organizational teams. The combination of infrastructure independence and broad API support allows this solution to unify data from new and existing analytic systems and solutions into a single logical data backbone without complex extract, transform, and load (ETL) processes or copying of data.

Unlike traditional storage, HPE Ezmeral Data Fabric is not limited to just storing data. Instead, it provides access to, indexes and processes data from multiple sources, types, and physical locations. Data is stored as a file or object. Next it processes and stores metadata and data and stores them into an integrated NoSQL database. Metadata can be used for quick queries, such as fraud identification, while the data can be directly accessed for complex model training.

In essence, HPE Ezmeral Data Fabric increases productivity by relieving analytic and data science teams from the tedious process of negotiating access rights to data silos, then copying and normalizing the data before it can be analyzed. It frees IT from manual data management processes enabling them to focus on innovation. By integrating into existing authentication and authorization systems, it reduces the cost and risk associated with building, deploying, securing, managing and integrating unique users, apps, and workloads into the organization's security and data access framework.

### Business Outcomes

**Drives adaptive intelligence** with autonomous capabilities that simplifies advanced analytics and data intelligence across the enterprise. Industry standard API support delivers more granular security and access rights while making it easier to integrate new technologies for developers. It can support multiple use cases and provide cost savings by enabling data reuse across multiple projects.

**Provides a 360-degree view of data** by unifying files, objects, NoSQL databases, real-time and batch streams into a logical data backbone that provides the foundational layer required for analytics, AI, and ML. This enables faster identification or trends and anomalies and empowers data teams to search and identify structured and unstructured datasets for each project. This solution is the foundational data layer for the HPE Ezmeral software portfolio, HPE GreenLake for Data Fabric, and HPE GreenLake for Private Cloud Enterprise.

---

## Overview

**Consistent data access** through a single access point for users or apps. Using a global namespace, a simple URL enterprise into each applications reduces the number of trouble tickets and dependence on specialized skills required to provide each end users, app and partner solution access to the organization's security and data access framework. From the user's point of view, data is consumed as if it was stored locally.

**Hybrid security made easy** with a single security management system that integrates into existing authentication and authorization systems. Without HPE Ezmeral Data Fabric, files, objects, databases, streams, apps, and users are manually integrated into existing IAM systems leading to higher costs, expensive skillsets, and increased risk for the organization.

**Continuous intelligence and situation awareness** with an ecosystem of certified open-source tools that layer directly on top of the data fabric to process JSON and column-oriented real-time and batch streams. Built-in high availability replicates both streams and messaging providing high performance up to a billion messages per second at millisecond-level delivery times.

---

### File and Object Data Service

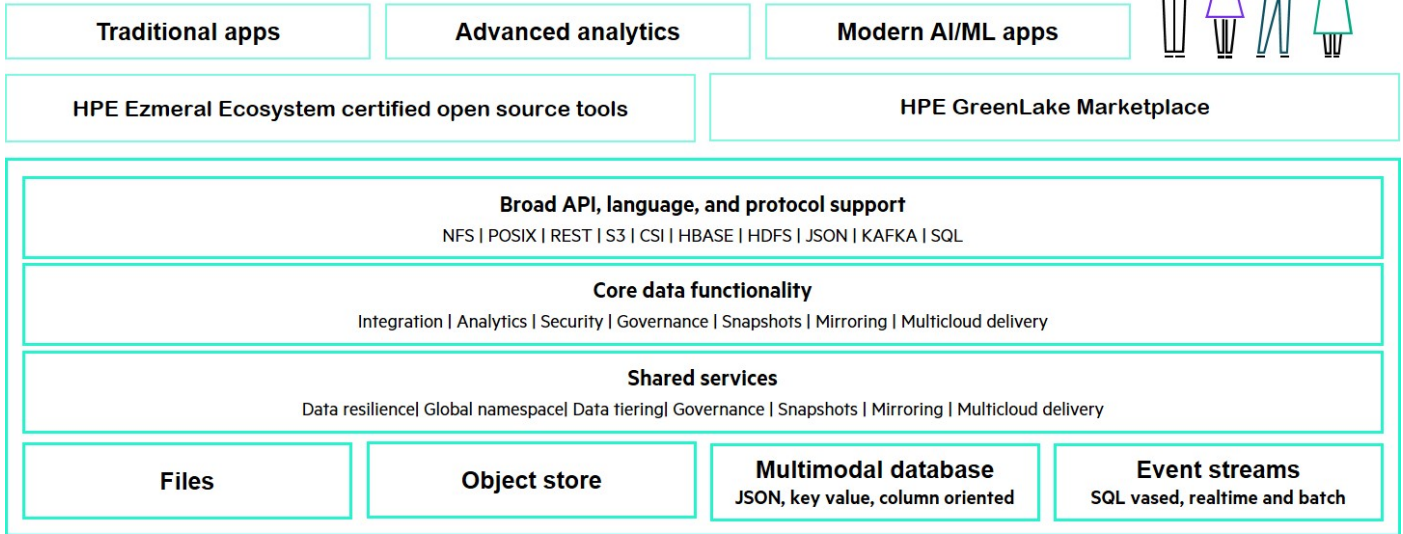
Customers who have existing file systems or have or anticipate scalability problems or have a need for object store will benefit from the file and object store capabilities of HPE Ezmeral Data Fabric. This is perfect for customers who are using existing solutions and experiencing higher costs and complexity because of increasing data volumes, new sites (i.e., edge, cloud and IoT) and new application requirements. HPE Ezmeral Data Fabric offers greater organizational agility, performance at a lower cost and reduced management complexity.

With the introduction of release 7.1 of Ezmeral Data Fabric in October 2022, the core persistence capabilities of the platform now include: files, objects, NoSQL databases and event store.

---



# Overview



## HPE Ezmeral Overview



## Standard Features

### Feature Set for HPE Ezmeral Data Fabric

#### Storage

- Distributed datastore for files and objects
- Support for Data Tiering: Hot (default), Warm (Erasure coded) and Cold (off-load to external stores)
- Support for NVMe, SSD and HDD

#### Access

- POSIX client (unlimited client instances)
- Platinum POSIX client (unlimited client instances)
- NFSv4 and NFSv3
- HDFS API and httpfs
- Supports S3 API
- Container Client including PACC, CSI (unlimited client instances)

#### Data Protection

- Built-in 3-way replication
- Resiliency and self-healing
- Auto-balancing
- Multi-site volume mirroring

#### Performance and Optimization

- Data Topologies
- Compression
- Multiple file server instances per node
- Data tiering into erasure-coded storage for capacity optimization or to third-party “cold” storage to lower costs

#### Administration and Security

- Unified security including authentication, authorization, wire and data-at-rest encryption.
- Administration using CLI, REST, GUI
- Rolling upgrades
- Volume Mirroring for Disaster Recovery
- UI Installer
- Quotas
- Snapshots
- Multi-tenancy (not for Object Store)
- Global Namespace

---

### HPE Ezmeral Data Fabric 7.1 Enhancements & New Features

- There are no new features added to the Ezmeral Data Fabric.
- New enhancements are made in the add-on options as shown below.

---

### HPE Ezmeral Data Fabric 7.0 Add-on options:

#### HPE Ezmeral Data Fabric Analytics for Apache Spark

- Upgraded Spark version to 3.3.0 in Ezmeral Ecosystem Pack (EEP 9.0.0)
- With Spark 3.3.0 and Hadoop 3.3.4, the platform now supports GPU based acceleration for Spark based applications.



## Standard Features

### HPE Ezmeral Data Fabric Analytics Standard Package

Includes Hadoop, Drill, Spark, Apache Ranger and Apache NiFi (see Spark details above)

Hadoop 3.3.4: Is an add-on option for HPE Ezmeral Data Fabric with Ezmeral Ecosystem Pack (EEP 9.0.0)

- Hadoop ecosystem components are included as optional add-on. This includes support for various Hadoop components like: Hive, HBase, Hue, Livy, Airflow etc.

Apache Drill 1.20.2.0 Is an add-on option for HPE Ezmeral Data Fabric in Ezmeral Ecosystem Pack (EEP)

- To support SQL queries
- Apache Drill supports high-performance analysis on the semi-structured and rapidly evolving data coming from modern Big Data applications, while still providing the familiarity and ecosystem of ANSI SQL, the industry-standard query language. Drill provides plug-and-play integration with existing Apache Hive and Apache HBase deployments.

As part of Ezmeral Data Fabric 7.1 and Ezmeral Ecosystem Pack 9.0.0, two new components are introduced in the Analytics Standard Package:

- Apache Ranger 2.3.0
  - Apache Ranger support as part of Ezmeral Ecosystem Pack 9.0 is to help Cloudera customers migrate with compatible functionality on the HPE Ezmeral Data Fabric platform. Apache Ranger will provide centralized security administration to manage all security related tasks in a central UI or using REST APIs. Apache Ranger will provide support for different authorization methods - Role based access control, attribute-based access control etc.
- Apache NiFi 1.16.3
  - Apache NiFi supports powerful and scalable directed graphs of data routing, transformation, and system mediation logic. Apache NiFi is an integrated data logistics platform for automating the movement of data between disparate systems. It provides real-time control that makes it easy to manage the movement of data between any source and any destination.

The following components are deprecated in this release of Ezmeral Ecosystem Pack

- Flume
- Oozie

For more details see: [https://docs-datafabric.mip.storage.hpccorp.net/71/c\\_ecosystem\\_intro.html](https://docs-datafabric.mip.storage.hpccorp.net/71/c_ecosystem_intro.html)

### HPE Ezmeral Data Fabric Analytics Premium Package

Includes Hadoop, Drill, Spark (see details above), Apache Ranger, Apache NiFi, Streams and No SQL Databases.

Event Streams: Is an add-on option for HPE Ezmeral Data Fabric File and Object

- New feature added to support Kafka wire protocol, so that there is no need to compile Kafka API applications to publish and consume messages from Data Fabric Streams.
- Kafka client 2.6.1 and Kafka Eco 6.0 support in Ezmeral Ecosystem Pack (EEP)
- With Kafka client 2.6.1 and Kafka Eco 6.0, the customers can get the following benefits:
  - Multiple vulnerabilities fixed
  - Scala2.13 support added
  - MirrorMaker2.0 added
  - Stable spring-kafka framework requires Kafka 2.5-2.8 (spring-kafka compatible with Kafka 2.1.1 are EOL) kafka-eventstreams is used instead of mapr-streams to access the streams on the server side



## Standard Features

NoSQL Databases: Is an add-on option for HPE Ezmeral Data Fabric

With the optional add-on option, you get the NoSQL databases which includes the following:

- JSON Document Database: JSON document database is supported on the platform and one can use OJAI API to interact with the data.
- Columnar (Key-value) Database: Columnar Database is supported on the platform and you can use HBase API to interact with this data.

## HPE Ezmeral Data Fabric File and Object Store Capability Matrix

Capabilities	HPE Ezmeral Data Fabric File & Object Store
Distributed File and Object store	Yes
Unified security including authentication, authorization, policy-based security, external KMIPkeystore, and encryption (on-wire & at-rest)	Yes
Support compression, auto-balancing, resilient through replication	Yes
Highly scalable, global namespace, quotas, auditing, rolling upgrades	Yes
Data topologies, data tier support: 'hot' for fast data; 'warm' for Erasure coded and 'cold' to offload data to 3rd party S3-based external stores.	Yes
Container Client including PACC, Flex Volume Plugin, CSI (unlimited client instances)	Yes
Snapshots and mirroring of data across clusters (used for data safety, access from multiple geo locations and disaster recovery)	Yes
Operating systems supported: RHEL, SUSE, Ubuntu and Oracle Enterprise Linux	Yes
<b>Management using CLI, REST and GUI</b>	Yes
<b>Support for NVMe, SSD &amp; HDD</b>	Yes
<b>Support for NFSv4 &amp; NFSv3</b>	Yes
<b>Supports the following for File store only:</b>	Yes
<ul style="list-style-type: none"> <li>• Multi-tenancy</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• Container client including PACC, Flex Volume Plugin, CSI</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• Data Topologies</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• Platinum POSIX client</li> </ul>	Yes
<b>Supports the following for Object store only:</b>	Yes
<ul style="list-style-type: none"> <li>• S3 API, Buckets, Objects, IAM accounts,</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• Policies for accounts, buckets, objects</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• Write Once Read Many (WORM) times support</li> </ul>	Yes
<ul style="list-style-type: none"> <li>• S3 Select</li> </ul>	Yes
<b>HPE Ezmeral Data Fabric Installer</b>	Yes
<b>HPE Ezmeral Data Fabric Monitoring</b>	Yes
<b>Multiple file server instances per node</b>	Yes



## Standard Features

### HPE Ezmeral Data Fabric for Analytics Capabilities Matrix

	HPE Ezmeral Data Fabric Analytics for Apache Spark	HPE Ezmeral Data Fabric Analytics Standard Package	HPE Ezmeral Data Fabric Analytics Premium Package
<b>Apache Spark Ecosystem</b>	Yes	Yes	Yes
<ul style="list-style-type: none"> <li>Apache Spark Core, and YARN</li> </ul>	Yes	Yes	Yes
<ul style="list-style-type: none"> <li>GraphX</li> </ul>	Yes	Yes	Yes
<ul style="list-style-type: none"> <li>HPE Ezmeral Data Fabric DB Binary and OJAI Connectors and for Apache Spark</li> </ul>	Yes	Yes	Yes
<ul style="list-style-type: none"> <li>HPE Ezmeral Data Fabric Streams Integration</li> </ul>	Yes	Yes	Yes
<ul style="list-style-type: none"> <li>SparkR, Spark Standalone and Spark on YARN support</li> </ul>	Yes	Yes	Yes
<b>GPU based acceleration</b>	Yes	Yes	Yes
<b>SQL ecosystem with Drill</b>		Yes	Yes
<ul style="list-style-type: none"> <li>Drill Explorer, Monitoring, query and admin UI, Drill standalone and Drill-on-YARN support</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>File formats (Text,JSON,Parquet), Impersonation</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>JDBC/ODBC drivers, Multiple data type support</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>Queries on File, on Hive tables and views</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>Queries on HPE Ezmeral Database tables and secondary indexes</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>Schema-less ANSI-compliant distributed SQL query engine, SQLLine</li> </ul>		Yes	Yes
<b>Hadoop Ecosystem</b>		Yes	Yes
<ul style="list-style-type: none"> <li>Apache, Hbase, Hive, Hue, Yarn</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>Apache MapReduce v2 (MapReduce v1 is not supported)</li> </ul>		Yes	Yes
<ul style="list-style-type: none"> <li>HPE Ezmeral Data Fabric DB OJAI Connector for Apache Hive</li> </ul>		Yes	Yes
<b>Airflow</b>		Yes	Yes
<b>Apache Ranger</b>		Yes	Yes
<b>Apache NiFi</b>		Yes	Yes
<b>HPE Ezmeral Database</b>			Yes
<ul style="list-style-type: none"> <li>Key-Value Database using HBase API</li> </ul>			Yes
<ul style="list-style-type: none"> <li>JSON document database using OJAI API</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Automatic compactions, Change Data Capture (CDC)</li> </ul>			Yes
<ul style="list-style-type: none"> <li>HPE Ezmeral Data Fabric DB Data Access Gateway</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Multi-master table replication</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Multiple file servers instances per node</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Multi-tenancy, Policy-based security, Resiliency and self-healing, secondary indexes, strong consistency</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Table replication that can be used for disaster recovery and multi-site access</li> </ul>			Yes
<ul style="list-style-type: none"> <li>Unified Security including authentication, authorization, encryption (wire and data-at-rest) and auditing</li> </ul>			Yes



## Standard Features

	HPE Ezmeral Data Fabric Analytics for Apache Spark	HPE Ezmeral Data Fabric Analytics Standard Package	HPE Ezmeral Data Fabric Analytics Premium Package
<b>HPE Ezmeral Data Fabric Event Data Streams</b>			Yes
<ul style="list-style-type: none"> <li>• Kafka wire-level protocol support</li> </ul>			Yes
<ul style="list-style-type: none"> <li>• Automatic partition balancing</li> </ul>			Yes
<ul style="list-style-type: none"> <li>• Distributed publish-subscribe messaging infrastructure</li> </ul>			Yes
<ul style="list-style-type: none"> <li>• Kafka API support, Connect REST Proxy, Schema Registry, Streams</li> </ul>			Yes
<ul style="list-style-type: none"> <li>• KSQL, Multi-site stream replication, multi-tenancy</li> </ul>			Yes
<ul style="list-style-type: none"> <li>• Unified Security including authentication, authorization, encryption (wire and data-at-rest) and auditing</li> </ul>			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. Many 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.

### HPE 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**

---

HPE Support 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 HPE Support Services at <https://ssc.hpe.com/portal/site/ssc/>
- 



## Configuration Information

### Ordering Information

HPE Ezmeral Data Fabric File and Object Store subscriptions include an electronic license to use on one terabyte of storage, Customer's license capacity under management is defined in the Additional Licensing Authorizations (ALA) for **HPE Ezmeral Software**.

#### Notes:

- Each subscription SKU includes a term license to use on one core plus the associated 24x7 HPE Pointnext Tech Care support.
- HPE Ezmeral Data Fabric is licensed by the number of unique cores available to the kernel in the OS on which the HPE Ezmeral software is directly installed, regardless of the number of threads in each core.

All HPE Ezmeral Data Fabric subscriptions include support with 24x7 coverage.

#### **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.

#### Notes:

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

## HPE Ezmeral Data Fabric 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

#### Description

	<b>SKU</b>
HPE Ezmeral Data Fabric File and Object Store E-LTU	R9J44AAE
HPE Ezmeral Data Fabric Analytics for Apache Spark E-LTU	R9J45AAE
HPE Ezmeral Data Fabric Analytics Standard Package E-LTU	R9J46AAE
HPE Ezmeral Data Fabric Analytics Premium Package E-LTU	R9J47AAE

### S4

#### Description

	<b>SKU</b>
HPE Ezmeral Data Fabric File and Object Store 1-year 24x7 E-LTU	SOM33AAE
HPE Ezmeral File and Object Store 2-year 24x7 E-LTU	SOM52AAE
HPE Ezmeral File and Object Store 3-year 24x7 E-LTU	SOM34AAE
HPE Ezmeral Data Fabric File and Object Store 4-year 24x7 E-LTU	SOM35AAE
HPE Ezmeral Data Fabric File and Object Store 5-year 24x7 E-LTU	SOM36AAE
HPE Ezmeral Data Fabric Analytics for Apache Spark 1-year 24x7 E-LTU	SOM37AAE
HPE Ezmeral Data Fabric Analytics for Apache Spark 2-year 24x7 E-LTU	SOM38AAE



## Configuration Information

### S4

#### Description

	<b>SKU</b>
HPE Ezmeral Data Fabric Analytics for Apache Spark 3-year 24x7 E-LTU	SOM39AAE
HPE Ezmeral Data Fabric Analytics for Apache Spark 4-year 24x7 E-LTU	SOM40AAE
HPE Ezmeral Data Fabric Analytics for Apache Spark 5-year 24x7 E-LTU	SOM41AAE
HPE Ezmeral Data Fabric Analytics Standard Package 1-year 24x7 E-LTU	SOM42AAE
HPE Ezmeral Data Fabric Analytics Standard Package 2-year 24x7 E-LTU	SOM43AAE
HPE Ezmeral Data Fabric Analytics Standard Package 3-year 24x7 E-LTU	SOM44AAE
HPE Ezmeral Data Fabric Analytics Standard Package 4-year 24x7 E-LTU	SOM45AAE
HPE Ezmeral Data Fabric Analytics Standard Package 5-year 24x7 E-LTU	SOM46AAE
HPE Ezmeral Data Fabric Analytics Premium Package 1-year 24x7 E-LTU	SOM47AAE
HPE Ezmeral Data Fabric Analytics Premium Package 2-year 24x7 E-LTU	SOM48AAE
HPE Ezmeral Data Fabric Analytics Premium Package 3-year 24x7 E-LTU	SOM49AAE
HPE Ezmeral Data Fabric Analytics Premium Package 4-year 24x7 E-LTU	SOM50AAE
HPE Ezmeral Data Fabric Analytics Premium Package 5-year 24x7 E-LTU	SOM51AAE

#### Learn More:

- HPE Ezmeral Learn On Demand: <https://learn.ezmeral.software.hpe.com/>
- HPE Ezmeral Data Fabric: [hpe.com/DataFabric](https://hpe.com/DataFabric)
- Support & documentation: [hpe.com/support/hpesc](https://hpe.com/support/hpesc)
- User documentation: <https://docs.datafabric.hpe.com>



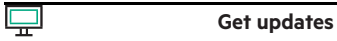
## Summary of Changes

<b>Date</b>	<b>Version History</b>	<b>Action</b>	<b>Description of Change</b>
07-Nov-2022	Version 8	Changed	Overview, Standard Features, Service and Support and Configuration Information sections were updated.
15-Aug-2022	Version 7	Changed	Service and Support and Configuration Information sections were updated.
01-Aug-2022	Version 6	Changed	Service and Support and Configuration Information sections were updated.
02-May-2022	Version 5	Changed	Overview and Configuration Information sections were updated.
15-Nov-2021	Version 4	Changed	Service and Support section was updated
06-Apr-2021	Version 3	Changed	Overview and Configuration Information sections were updated.
01-Mar-2021	Version 2	Changed	Overview, Standard Features, Service and Support and Configuration Information sections were updated.
01-Feb-2021	Version 1	New	New QuickSpecs



## Copyright

**Make the right purchase decision.  
Contact our presales specialists.**



---

© Copyright 2022 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.

a50002556enw - 16711 - Worldwide - V8 - 07-November-2022