

TECHNICAL VALIDATION

HPE Ezmeral Software

Software Foundation for AI and Analytics Across
Hybrid Environments

By Tony Palmer, Practice Director, Technical Validation Services
Enterprise Strategy Group

June 2023

Contents

Introduction	3
Background	3
HPE Ezmeral Software	4
Enterprise Strategy Group Technical Validation	5
HPE Ezmeral Unified Analytics Software.....	5
Enterprise Strategy Group Testing	5
HPE Ezmeral Data Fabric Software	11
Enterprise Strategy Group Testing	11
Conclusion	15

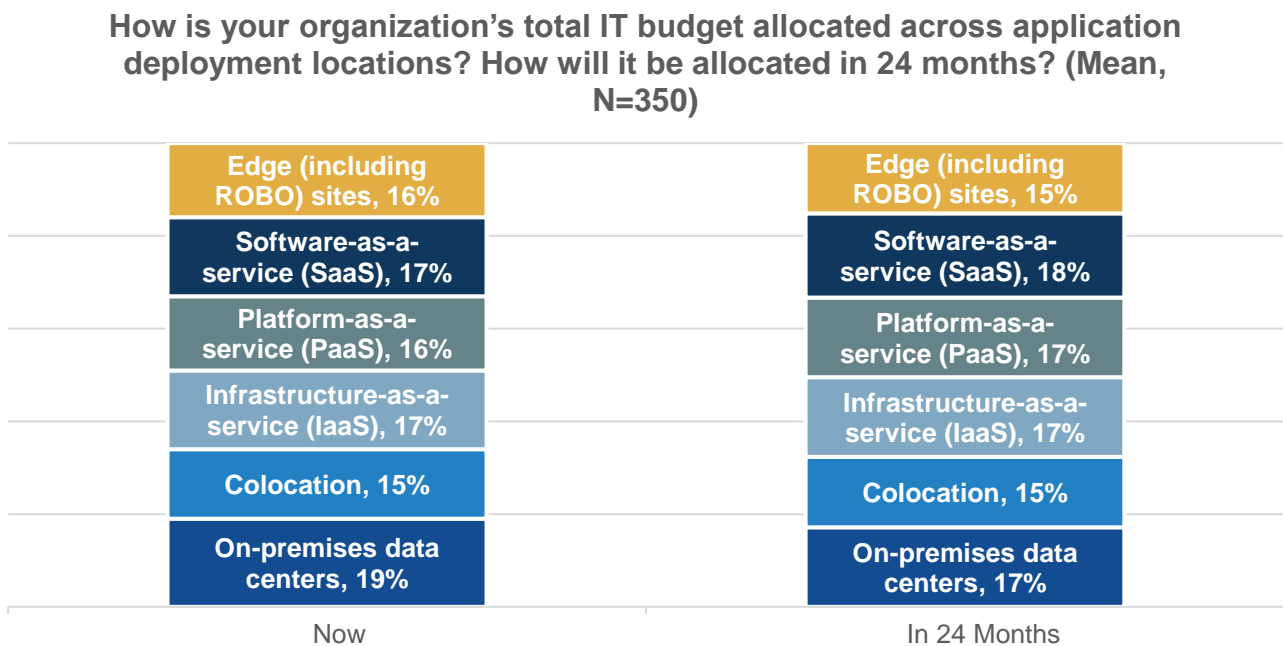
Introduction

This Technical Validation by TechTarget's Enterprise Strategy Group documents our evaluation of HPE Ezmeral Software. This report focuses on how HPE Ezmeral Software gives customers the software foundations they need to develop and deploy analytics applications seamlessly across hybrid-cloud environments.

Background

Organizations are increasingly distributed. Enterprise Strategy Group research found that the majority of enterprises surveyed leveraged two or more data centers (61%), two or more colocation providers (68%), three or more infrastructure-as-a-service (IaaS) providers (85%), and three or more platform-as-a-service (PaaS) providers (84%).¹ These organizations are also leveraging edge locations, allocating nearly the same budget to edge—including remote office/branch office (ROBO)—as to on-premises data centers, colocation facilities, and as-a-service options (see Figure 1). Workloads need to combine data from dissimilar sources and formats across these locations, adding complexity and cost to the access, management, and securing of data. In fact, 86% of organizations surveyed by Enterprise Strategy Group confirmed that they use seven or more integrated tools across all states of their data lifecycle, from initial data collection to visualization of data for decision-makers.²

Figure 1. IT Budget Allocated Across Many Locations



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

This creates multiple challenges. Organizations' limited data visibility across hybrid locations makes managing, accessing, and extracting value from their data more difficult and time-consuming, resulting in negative impacts on both time to production and time to innovation. Organizations want to use their data in multiple ways for numerous purposes, including leveraging AI/ML for critical decision-making, optimizing operations, developing new revenue streams, and enhancing the customer experience, to name just a few and open source technology is foundational

¹ Source: Enterprise Strategy Group Research Report, *Multi-cloud Application Deployment and Delivery Decision Making*, to be published May/June 2023.

² Source: Enterprise Strategy Group Research Report, *Data Platforms: The Evolution to Achieving Data-driven Empowerment*, to be published May/June 2023.

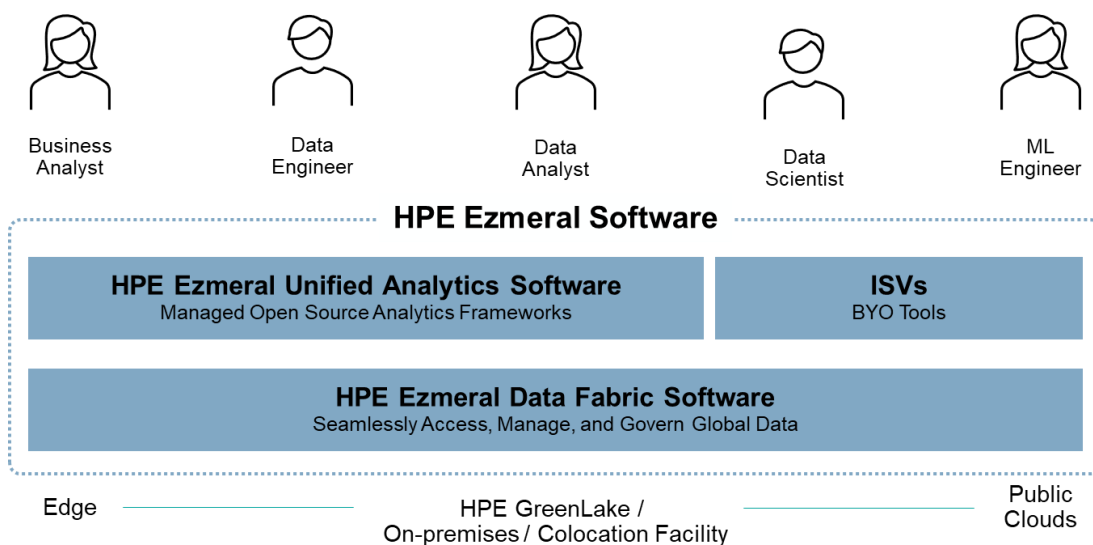
to data strategies. Our research reveals that more than half of organizations (54%) prefer or use open source technology for data analytics whenever they can, but open source support (48%) and integration (44%) are the most significant challenges reported to Enterprise Strategy Group.³ In the same survey, investing in technology that enables better collaboration (45%) and making data more accessible (44%) are the top two strategies organizations report using to improve usage of data.

Taking all these challenges into consideration, what is needed is an open, consistent software stack to simplify managing data and analytics applications across hybrid cloud environments in support of data-first strategies. Software should allow organizations to focus on their distributed businesses by leveraging data globally, accelerating application development, and enabling analytics-driven processes and insights with a consistent cloud experience from edge to cloud, with confidence.

HPE Ezmeral Software

HPE Ezmeral Software is designed to give customers the software foundations to develop and deploy analytics applications across hybrid cloud environments quickly, confidently, and easily. HPE Ezmeral Unified Analytics Software is hybrid by design, so it can be deployed across business edges, in the cloud, or on-premises; it has an open and extensible framework to make it easy to extend with custom applications, open source applications, or third-party independent software vendor (ISV) solutions; provides a consistent user experience across environments; provides a single user experience across system management, data storage, and analytical workloads; includes enterprise-grade security (with zero trust authentication, single sign-on, app-level isolation, user, application, and data access controls and observability to ensure apps are always secure and functioning at their peak levels); and provides connectivity to distributed data sources regardless of where the data is located. HPE Ezmeral Data Fabric Software also provides a global data plane and simplifies data management for hybrid data deployments across on-premises, edge, multiple clouds, and HPE GreenLake.

Figure 2. HPE Ezmeral Software



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

³ Source: Enterprise Strategy Group Research Report, [Cloud Analytics Trends](#), March 2022.

HPE Ezmeral Software leverages a fully managed, curated set of open source tools to address the needs of the various teams that work to extract insight and value from data, including business analytics, data engineering, data analytics, data science, and machine learning engineering. An extensible catalog of custom open source apps, third-party ISV solutions, and external data sources is provided with open interfaces. HPE calls Ezmeral Software hybrid-by-design, which means that multiple users and applications can securely access a single logical source of data that is distributed across multiple locations. The user and the data they need can be miles apart but access is as simple as clicking on an entry in the user's directory tree.

HPE Ezmeral Software is delivered via a consumption-based SaaS offering to provide a consistent user experience with predictable cost of ownership and includes transparent tracking of spending across environments, allowing organizations to optimize TCO.

Enterprise Strategy Group Technical Validation

Enterprise Strategy Group evaluated how HPE Ezmeral Software gives customers a managed open source software foundation to develop and deploy analytics applications seamlessly across hybrid cloud environments.

HPE Ezmeral Unified Analytics Software

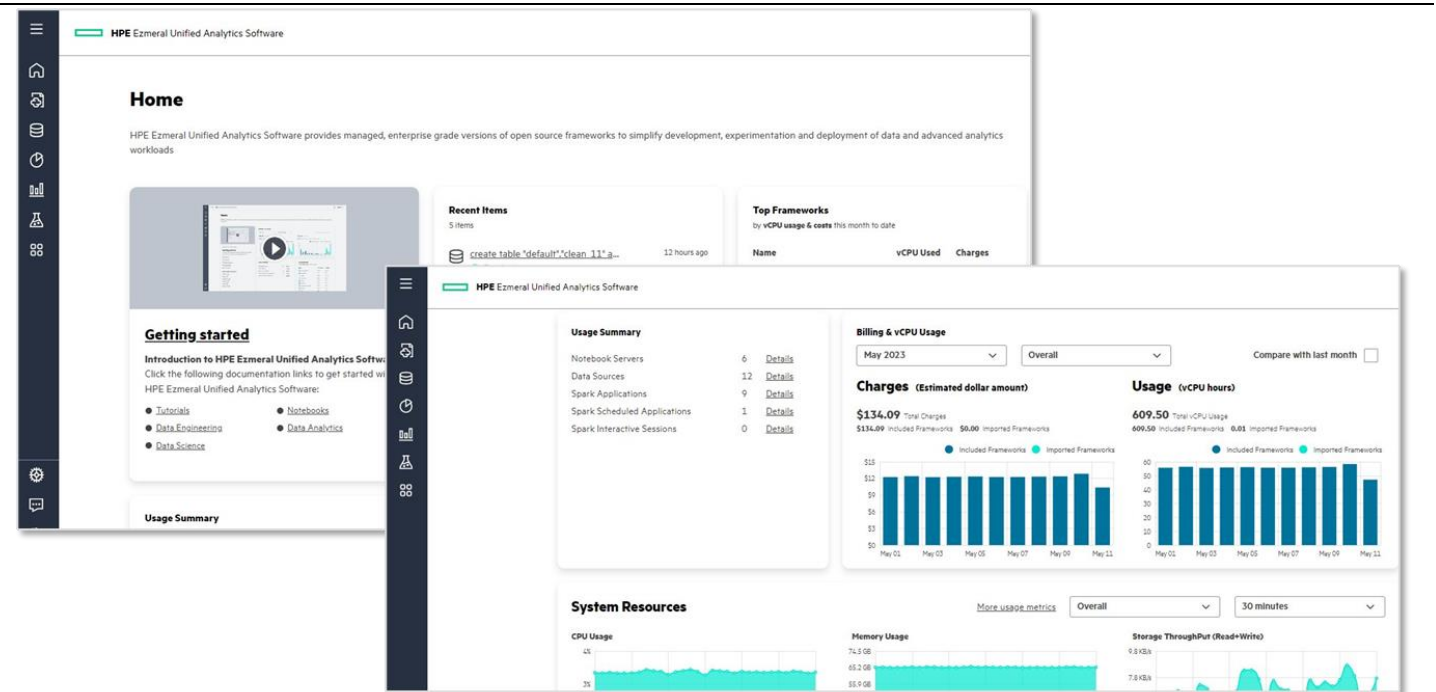
HPE's curation of all software gives users end-to-end tools to build, train, and deploy data and applications across hybrid environments with speed and simplicity. Self-service analytics tools boost productivity by minimizing or even eliminating bottlenecks and distractions, enabling teams to focus on their real work. Unified Analytics is the process of analyzing and interpreting data from any number of disparate sources in an enterprise environment.

Enterprise Strategy Group examined how HPE Ezmeral Unified Analytics Software provides a fully managed solution that includes the most popular open-source frameworks that is designed to address these challenges. HPE Ezmeral Unified Analytics Software brings together data science, data engineering, and data analytics capabilities into an end-to-end analytics solution to enable organizations to improve productivity as their teams develop and deploy high-value applications quickly and cost-efficiently.

Enterprise Strategy Group Testing

We started at the HPE Unified Analytics Software Dashboard. Organizations can access and manage everything they need from here (see Figure 3). The dashboard provides a simple experience that segments workflows based on the types of actions that a data engineer, data analyst, or data scientist would typically perform.

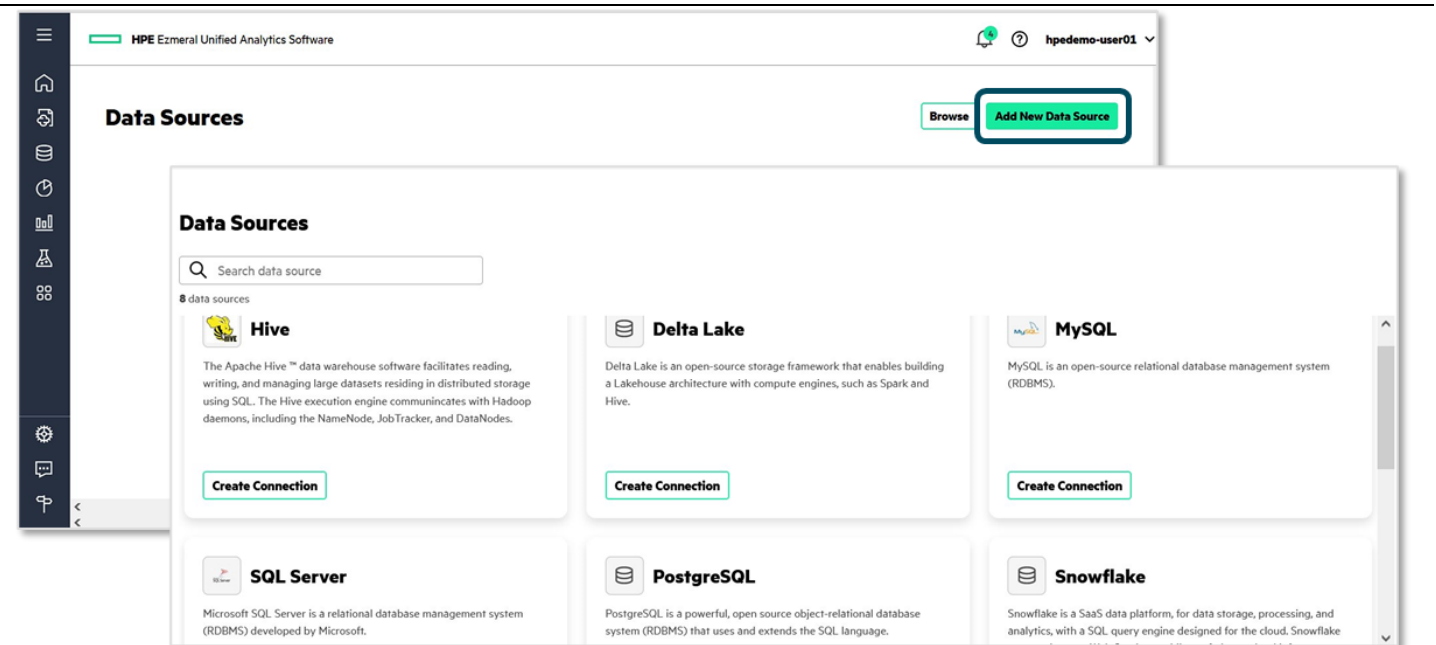
Figure 3. The HPE Ezmeral Unified Analytics Software Dashboard



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

For example, data engineers build data pipelines to ingest, cleanse, and prepare data for downstream personas to leverage. HPE provides a streamlined workflow to add access to different data sources. HPE provides an uncomplicated way to do this out of the box, with multiple different connectors to access a variety of data sources (see Figure 4).

Figure 4. Adding Access to Data Sources

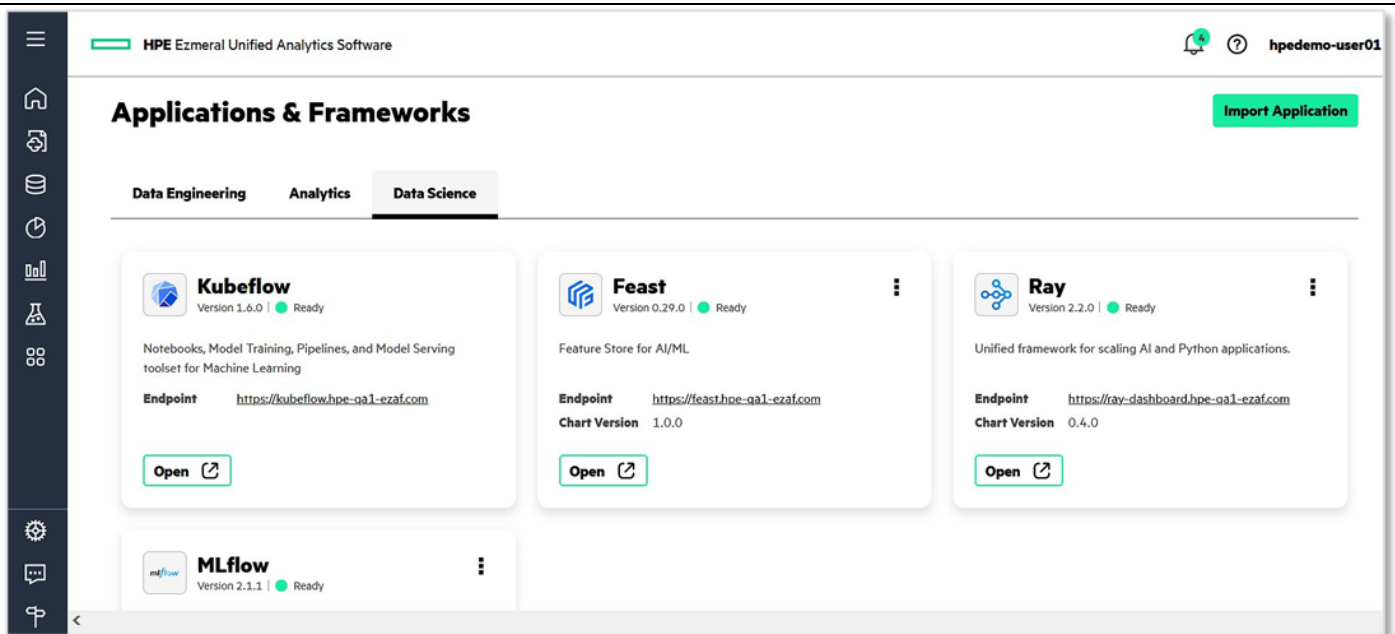


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

This gives organizations a simple option—native to the platform—to establish connectivity with different existing data sources. Once connectivity is established, organizations can browse through the different data sources to determine what exists in each of them.

Next, we clicked Applications and Frameworks, which took us to HPE’s curated collection of popular open source data engineering, data analytics, and data science frameworks to provide the right tool for the job at hand. Figure 5 shows the data science tab with a selection of the tools available.

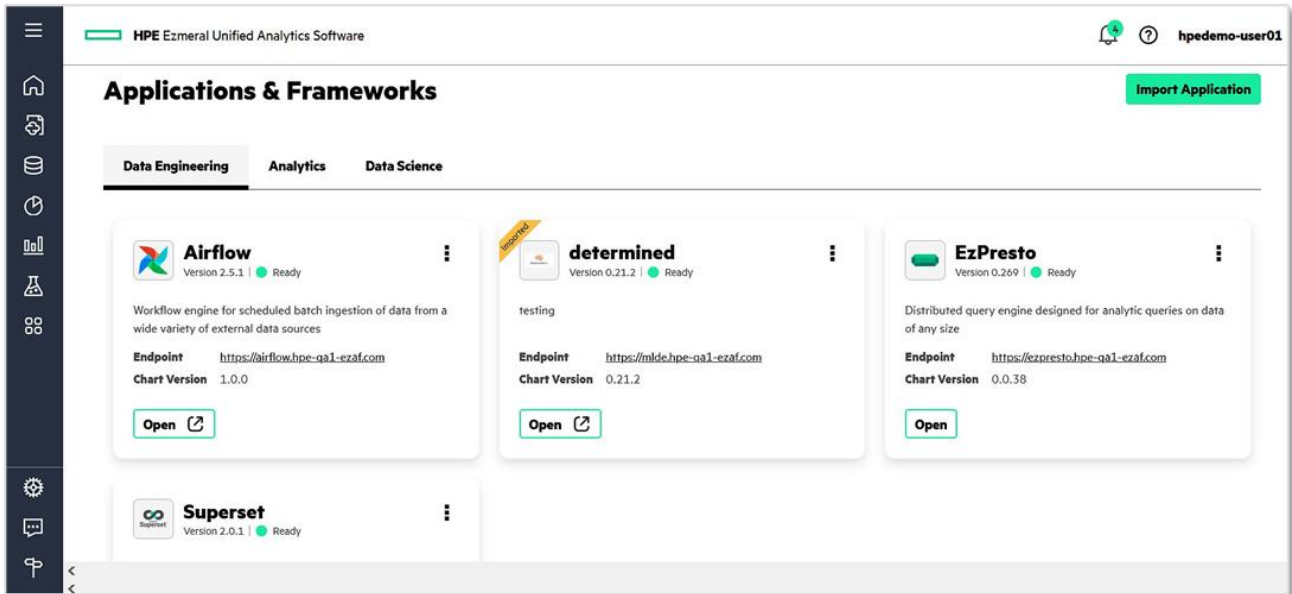
Figure 5. Applications and Frameworks-Data Science



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Data Engineering tools are shown in Figure 6. When organizations connect HPE Ezmeral Unified Analytics Software to various data sources they create a data catalog. Within the data catalog, organizations can cache datasets to disk or memory. This is important with exceptionally large datasets containing billions of rows or hundreds of terabytes of static or near-static data sources. This enables organizations to pull the data that they need once, store it locally, and avoid the compute and egress costs of pulling the data down multiple times.

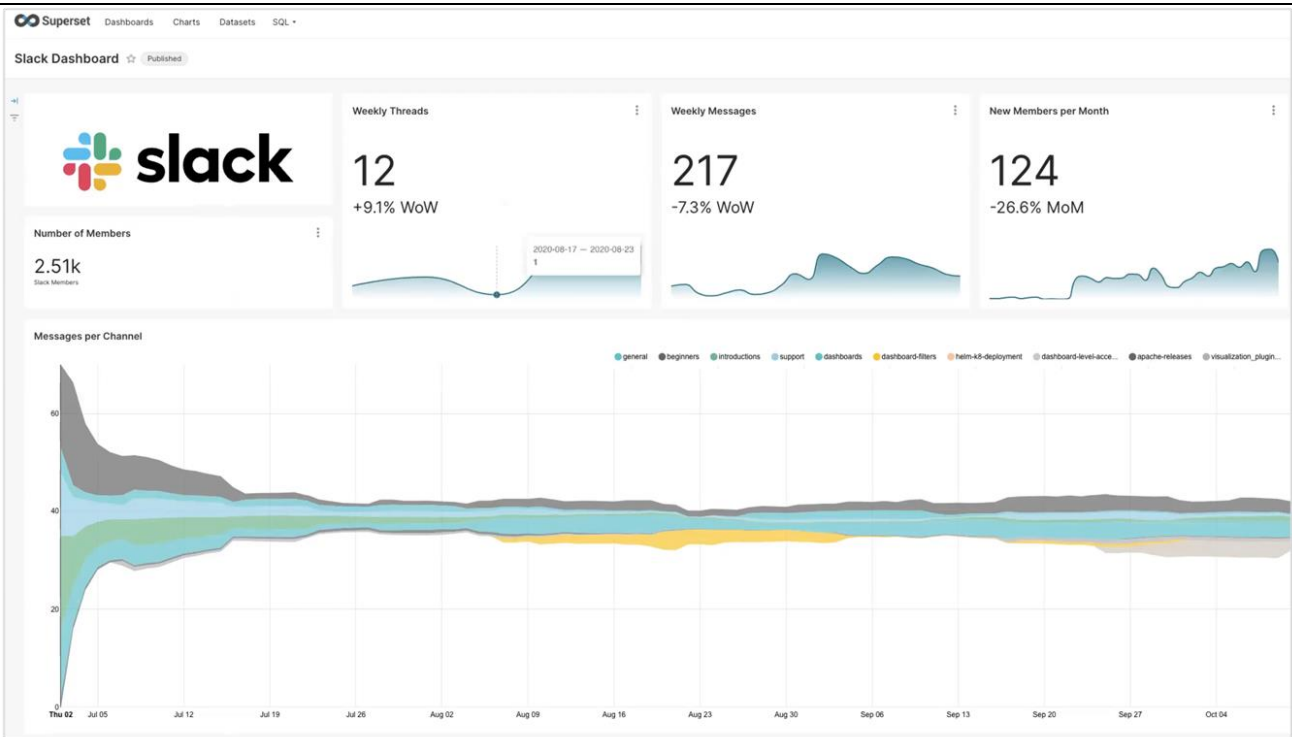
Figure 6. Applications and Frameworks-Data Engineering



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

They can use Superset to visualize the data in the catalog (see Figure 7). Once an organization creates any type of chart in Superset, they can specify query conditions on a selected data set to visualize the results in the chart. Superset works with EzPresto, the HPE Ezmeral Unified Analytics Software accelerated SQL query engine, to process the query and display results in the chart. The chart can be easily added to a dashboard that visualizes analytical workloads.

Figure 7. Data Visualization

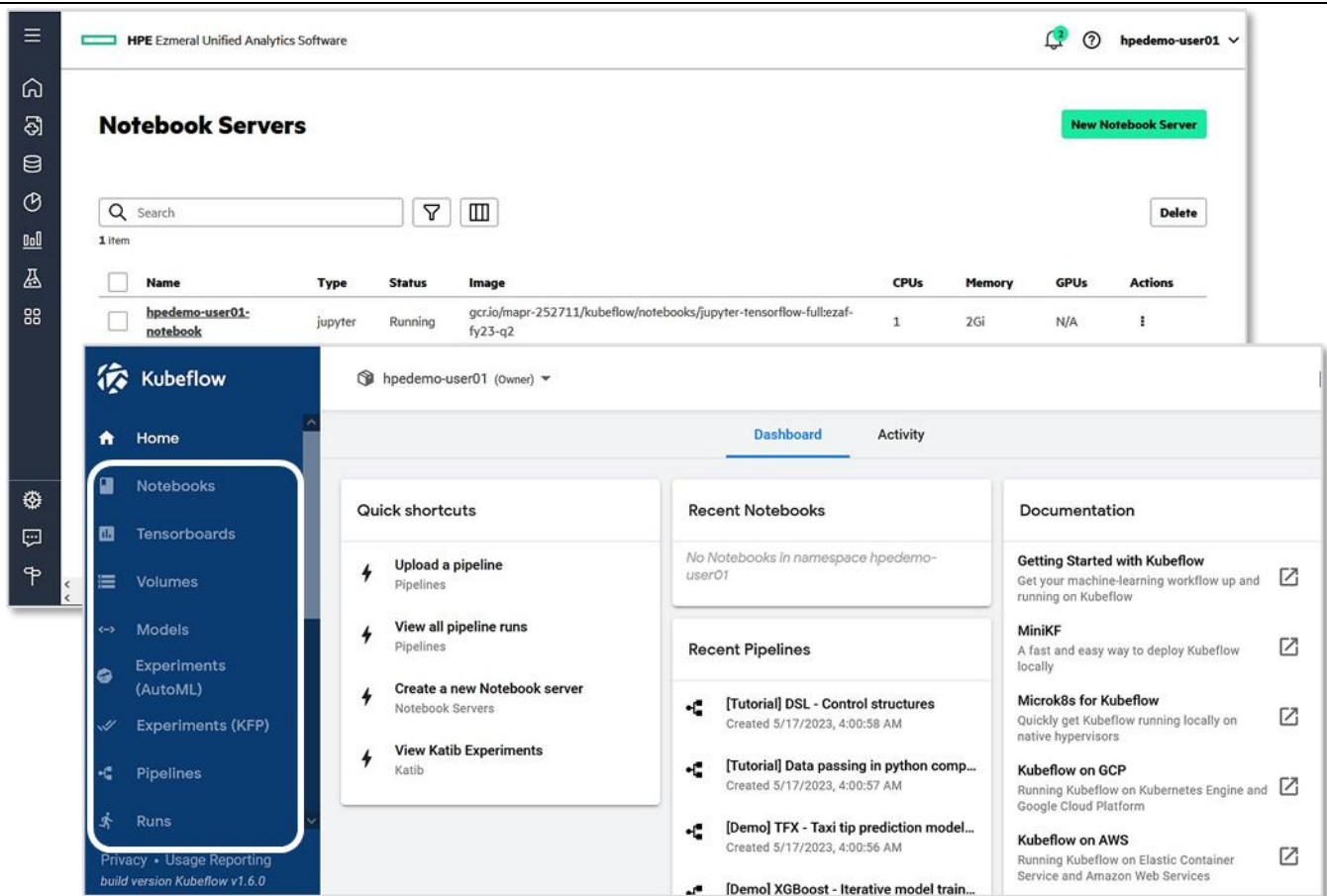


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

In addition to the integrated experience of the catalog of open-source frameworks and data sources, HPE Ezmeral Unified Analytics Software provides security and observability frameworks across all of an organization’s distributed data as well. It’s important to note that the entire product is open and extensible, so if a customer has a preferred open source package that is not yet in the catalog, they can easily add it themselves, and the package inherits the security brackets and observability frameworks that HPE has built into the platform.

Finally, we looked at elements of the end-to-end data science workflow, including notebooks and pipelines (see Figure 8), experiment tracking, and the model registry. We opened a Jupyter notebook from within the platform, which launched Kubeflow and opened the notebook with one click, automatically logging the user into the notebook using their initial login credentials. Within Kubeflow we were able to examine individual pipelines, runs, ML model structure, tags, parameters, and the metrics being tracked against.

Figure 8. Data Science Workflow



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

When a user is logged into the platform, their identity is carried throughout the platform and into all applications managed by the platform, providing a single sign-on experience with role-based access controls. HPE Ezmeral Unified Analytics Software incorporates zero trust security, which means that every service-to-service interaction is authenticated and encrypted using mTLS by default.

Why This Matters

Modern enterprises operate in distributed hybrid/multi-cloud environments with applications and data distributed across data centers, colocation facilities, IaaS providers, PaaS providers, and edge locations.⁴ Enterprise Strategy Group research reveals that gaining timely insight from this distributed data is a challenge for most organizations, with nearly half (45%) reporting that it takes them weeks to gain insight.⁵ These enterprises need the right analytics tools to be able to efficiently consume their data regardless of where it is located across hybrid/multi cloud environments and to extract useful insights from that data.

Enterprise Strategy Group found that HPE Ezmeral Unified Analytics Software addresses these challenges. HPE Ezmeral is extremely simple to navigate and use for data engineers, analysts, and scientists. From connecting to different data sources, cataloging the data, and using the data to running an ad-hoc query, building a machine learning model, and using their toolsets of choice, HPE Ezmeral Unified Analytics Software improves every stage of the data analytics lifecycle.

HPE uses an open source framework to help eliminate these challenges for distributed data but also to eliminate vendor lock in for platforms and tooling. In addition, HPE applies the principles of continuous improvement and innovation to reduce the cost constraints of proprietary zero trust enterprise security. A net effect is a boost in productivity. HPE provides all the tools an organization needs to make users more productive, accelerate development to quickly build and deploy applications across hybrid environments, and get real-time insights faster.

⁴ Source: Enterprise Strategy Group Research Report, *Multi-cloud Application Deployment and Delivery Decision Making*, to be published May/June 2023.

⁵ Source: Enterprise Strategy Group Research Report, [Cloud Analytics Trends](#), March 2022.

HPE Ezmeral Data Fabric Software

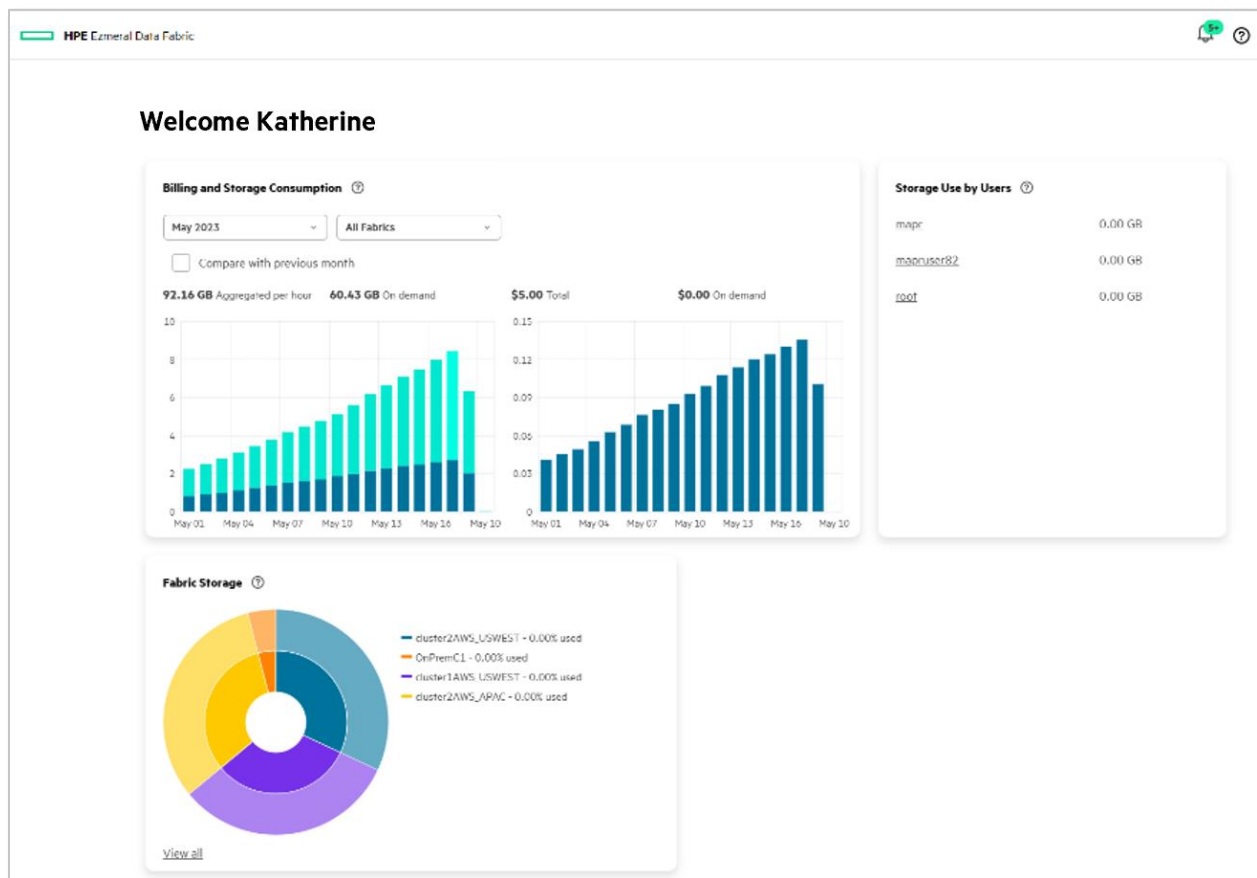
Enterprise Strategy Group explored how HPE Ezmeral Data Fabric Software enables organizations to seamlessly access, manage, and govern all their multi-format data across hybrid/multi-cloud environments in diverse formats.

Enterprise Strategy Group Testing

HPE Ezmeral Data Fabric Software is available as a service (SaaS) or it can be deployed on the customer’s own infrastructure, either connected or in an air-gapped environment. When deployed as SaaS, everything about the fabric, the architecture, the design, the installation, the configuration, and the setup is all managed by HPE so that an up-and-running fabric is delivered to the customer.

Upon logging in, the user can connect to the dashboard (see Figure 7). In this example, the service is running in HPE GreenLake, but the user experience is similar whether in an air-gapped or customer-managed environment.

Figure 9. Data Fabric Dashboard

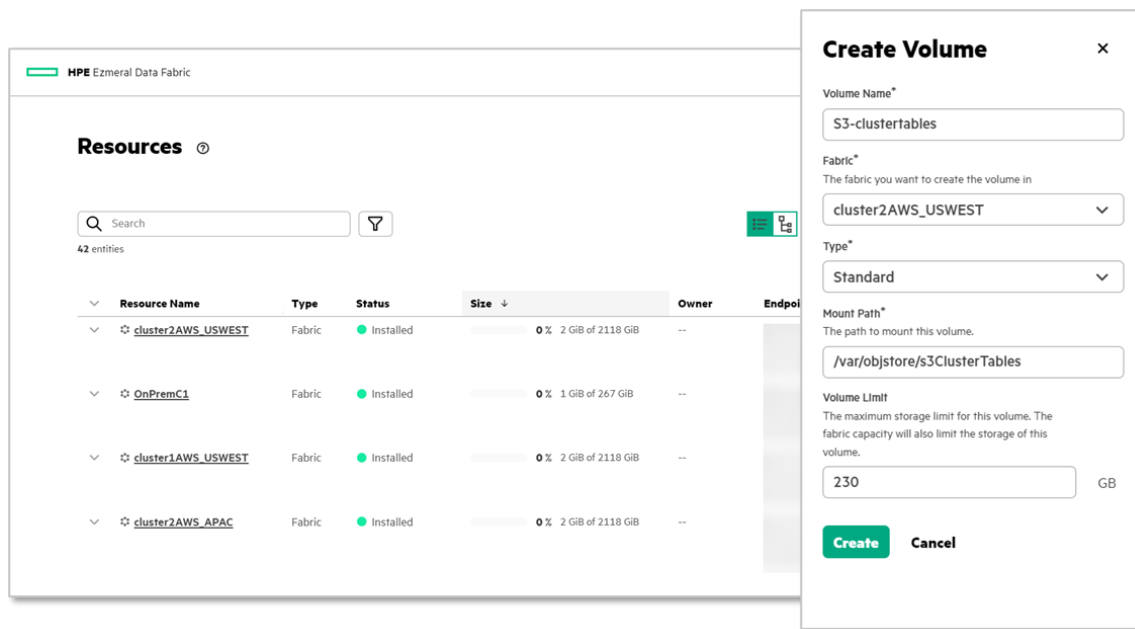


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

In this environment, we were logged in as a developer persona. There are three fabrics this user has access to: one in AWS, one in Microsoft Azure, and one on-premises, in an HPE colocation facility (see Figure 10).

This environment has several buckets, volumes, and Kafka topics already built, but in this exercise, we wanted to create a new volume. We clicked on the create *Volume* button and entered just six parameters: a name, the fabric to create the volume on, the mount path, the volume type (standard or mirrored), the volume storage limit, and percentage full notification threshold.

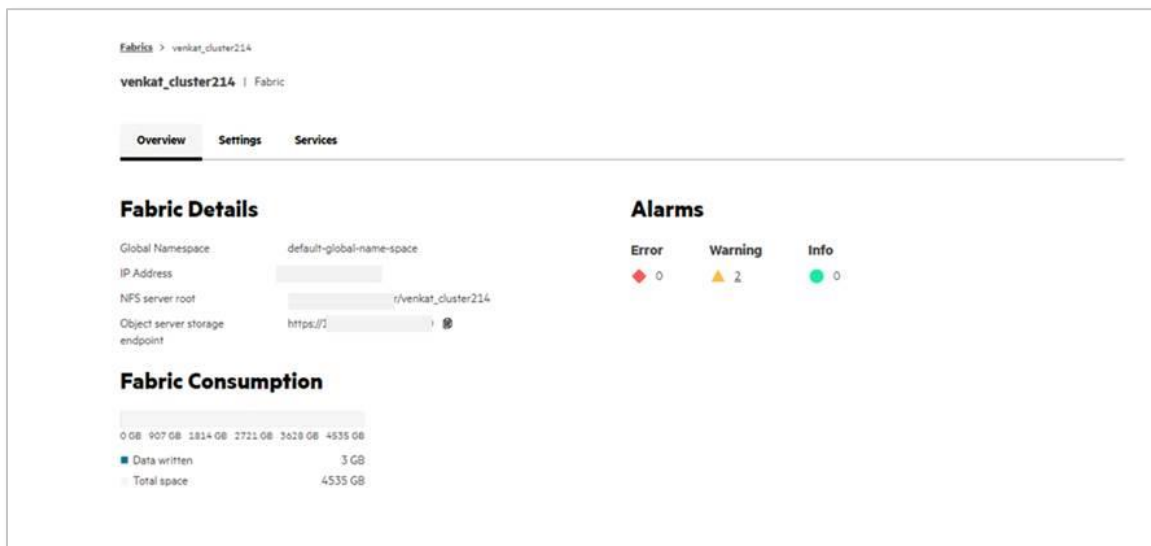
Figure 10. Data Fabrics and Resources—Creating a Volume



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

HPE Ezmeral Data Fabric Software works in the background to create, configure, then attach the volume to the appropriate data fabric within minutes. All the information needed to access the volume is available in the user interface, including security keys. We clicked on a Fabric to access the details page (see Figure 11).

Figure 11. Fabric Details

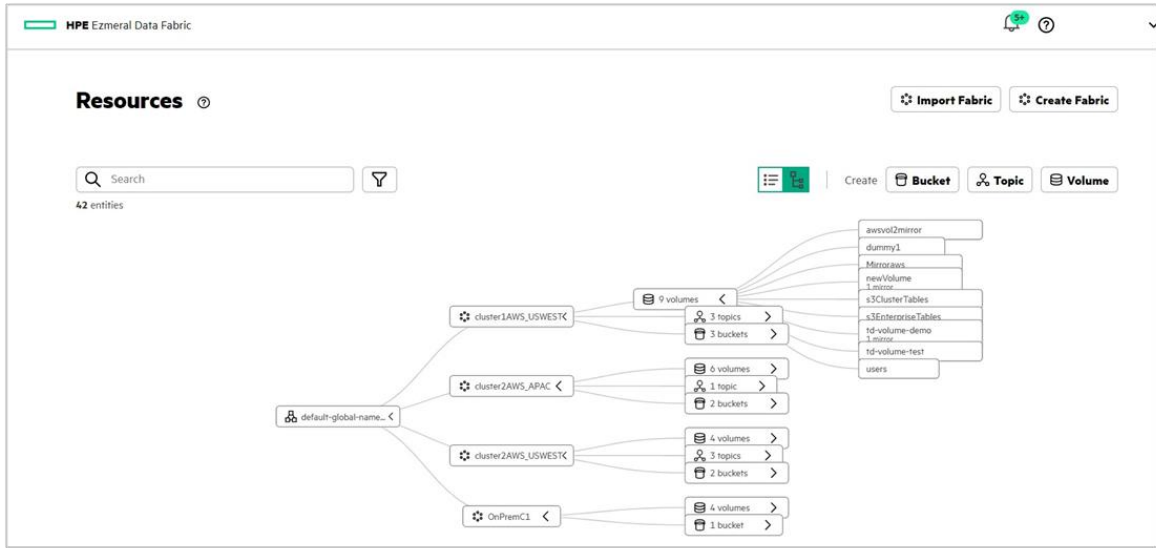


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Sharing storage with other developers is as simple as clicking share and entering the email addresses of the collaborators.

It's important to note that it's extremely easy for organizations to connect to external data fabrics and present them as part of a single global view that HPE Ezmeral Unified Analytics Software can use for projects. Enterprise Strategy Group looked at the HPE Ezmeral Data Fabric Software global namespace. As seen in Figure 12, all three fabrics in our test environment are in a single namespace. What this means for users is that from the root of any of the fabrics, AWS, Azure, or on-premises, a developer can access and use any file, object, or topic across all of the fabrics.

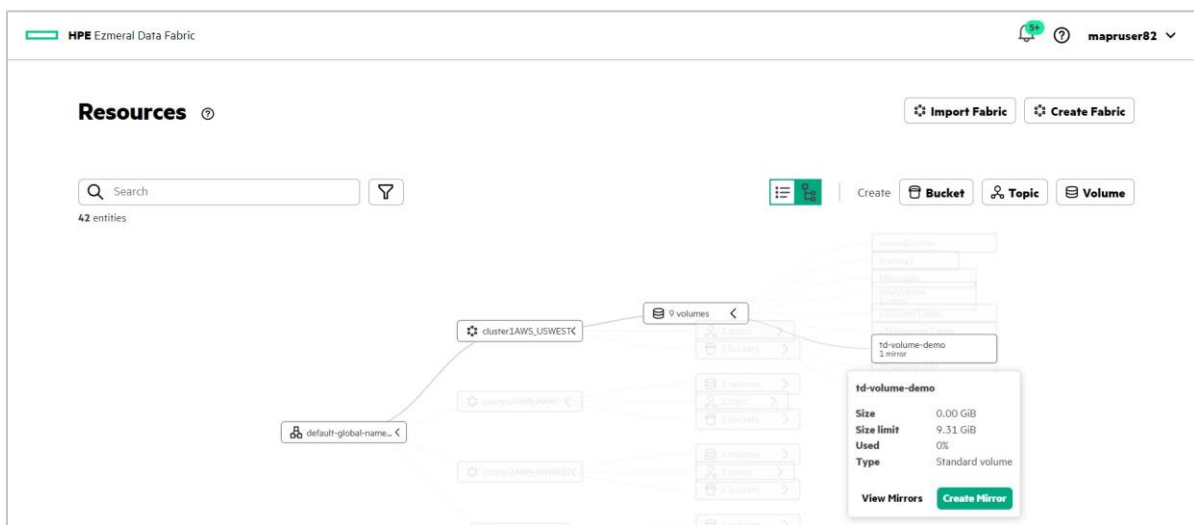
Figure 12. Global Namespace



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

With geographically distributed data, mounting a remote volume directly can be inefficient due to latency, so HPE Ezmeral Data Fabric Software provides the option to securely mirror volumes across data fabrics.

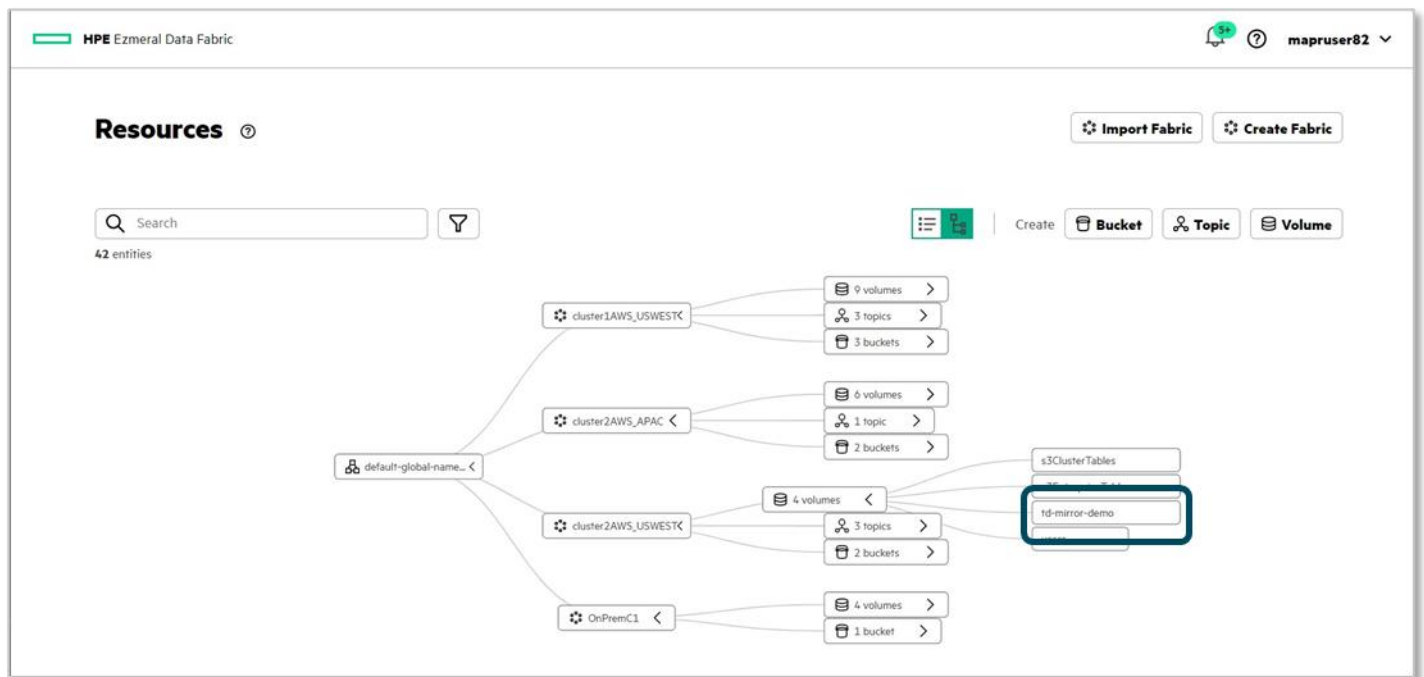
Figure 13. Creating a Mirror



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Once a user has determined that a particular volume created and updated in one location needs to be used in another location, they can configure a mirror and specify a refresh frequency. This action creates a policy that will update both environments through incremental mirroring processes.

Figure 14. Mirror Created on Another Cluster



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

This has numerous use cases, including moving edge cluster data to a core cluster for protection, cloud bursting to absorb workload spikes, and training models.

Why This Matters

Data, applications, and workloads are constantly on the move between edge locations, data centers, and multiple public cloud environments. At least half of IT professionals surveyed said their organizations move data between data centers and public cloud services (53%) and between multiple cloud services (50%) regularly or all the time.⁶ This movement complicates every stage of the data lifecycle and adds time, effort, and complexity to extracting valuable insights. Organizations clearly need a different approach if they want to leverage their data to remain competitive and reduce time to innovation.

Enterprise Strategy Group validated that HPE Ezmeral Data Fabric Software allows users to access and share data regardless of format or location. This simplifies management, access control, security, and governance across hybrid locations such as on-premises, public cloud, HPE GreenLake, and edge deployments.

⁶ Source: Enterprise Strategy Group Research Report, [Application Infrastructure Modernization Trends Across Distributed Cloud Environments](#), March 2022.

Conclusion

The modern enterprise is more distributed than ever, with multiple data centers, colocation facilities, IaaS providers, and PaaS providers all hosting applications and data in disparate formats.⁷ Both users and applications need direct access to data distributed across multiple locations and formats. Without it, companies experience increased complexity and risk to the organization in the form of increased costs from retaining duplicate data sets, software licenses across multiple products, and storage to retain data on each location. Normalizing data across multiple locations results in a delay in processing which means that insights may be stale and left unused. Enterprise Strategy Group research reveals that, in general, 73% of organizations report that it takes them from days to months to gain insight from their data.⁸

HPE Ezmeral Software was designed to address hybrid data challenges by providing seamless access to data with an open source solution delivered as a service. HPE Ezmeral helps organizations monitor applications, provides notifications and alerts, and has a mechanism to collect and ship logs to optimize the delivery of timely insights. Fine-grained access and data placement controls help organizations meet demanding regulatory, compliance, sovereignty, and performance needs. HPE Ezmeral Software provides customers with a consistent user experience that provides both the flexibility to place data and workloads where TCO can be optimized, and the ability to track and manage budgets.

Enterprise Strategy Group validated that HPE Ezmeral Unified Analytics Software is extremely simple to navigate and use for multiple data-centric personas. HPE Ezmeral Unified Analytics Software simplifies how users connect to different data sources, catalogue and use the data, run ad-hoc queries, and build machine learning models using open source software tools curated by HPE or their toolsets of choice.

Enterprise Strategy Group validated that HPE Ezmeral Data Fabric Software provides a federated, unified namespace that enables developers to access and share file, object, table, and stream data regardless of location in a consistent, global namespace. This unified data provides the foundational data layer required for trusted insights.

With HPE Ezmeral Software, organizations can deploy analytics-powered applications anywhere in a hybrid environment to operationalize analytics faster and solve complex business problems at scale. If your organization is looking for an end-to-end platform to optimize data productivity from edge to cloud with predictable economics, fine-grained controls, and flexible data placement, Enterprise Strategy Group recommends you take a serious look at HPE Ezmeral Software.

⁷ Source: Enterprise Strategy Group Research Report, *Multi-cloud Application Deployment and Delivery Decision Making*, to be published May/June 2023.

⁸ Source: Enterprise Strategy Group Research Report, [Cloud Analytics Trends](#), March 2022.

©TechTarget, Inc. or its subsidiaries. All rights reserved. TechTarget, and the TechTarget logo, are trademarks or registered trademarks of TechTarget, Inc. and are registered in jurisdictions worldwide. Other product and service names and logos, including for BrightTALK, Xtelligent, and the Enterprise Strategy Group might be trademarks of TechTarget or its subsidiaries. All other trademarks, logos and brand names are the property of their respective owners.


Information contained in this publication has been obtained by sources TechTarget considers to be reliable but is not warranted by TechTarget. This publication may contain opinions of TechTarget, which are subject to change. This publication may include forecasts, projections, and other predictive statements that represent TechTarget's assumptions and expectations in light of currently available information. These forecasts are based on industry trends and involve variables and uncertainties. Consequently, TechTarget makes no warranty as to the accuracy of specific forecasts, projections or predictive statements contained herein.

Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of TechTarget, is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact Client Relations at cr@esg-global.com.

About Enterprise Strategy Group

TechTarget's Enterprise Strategy Group provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

 contact@esg-global.com

 www.esg-global.com