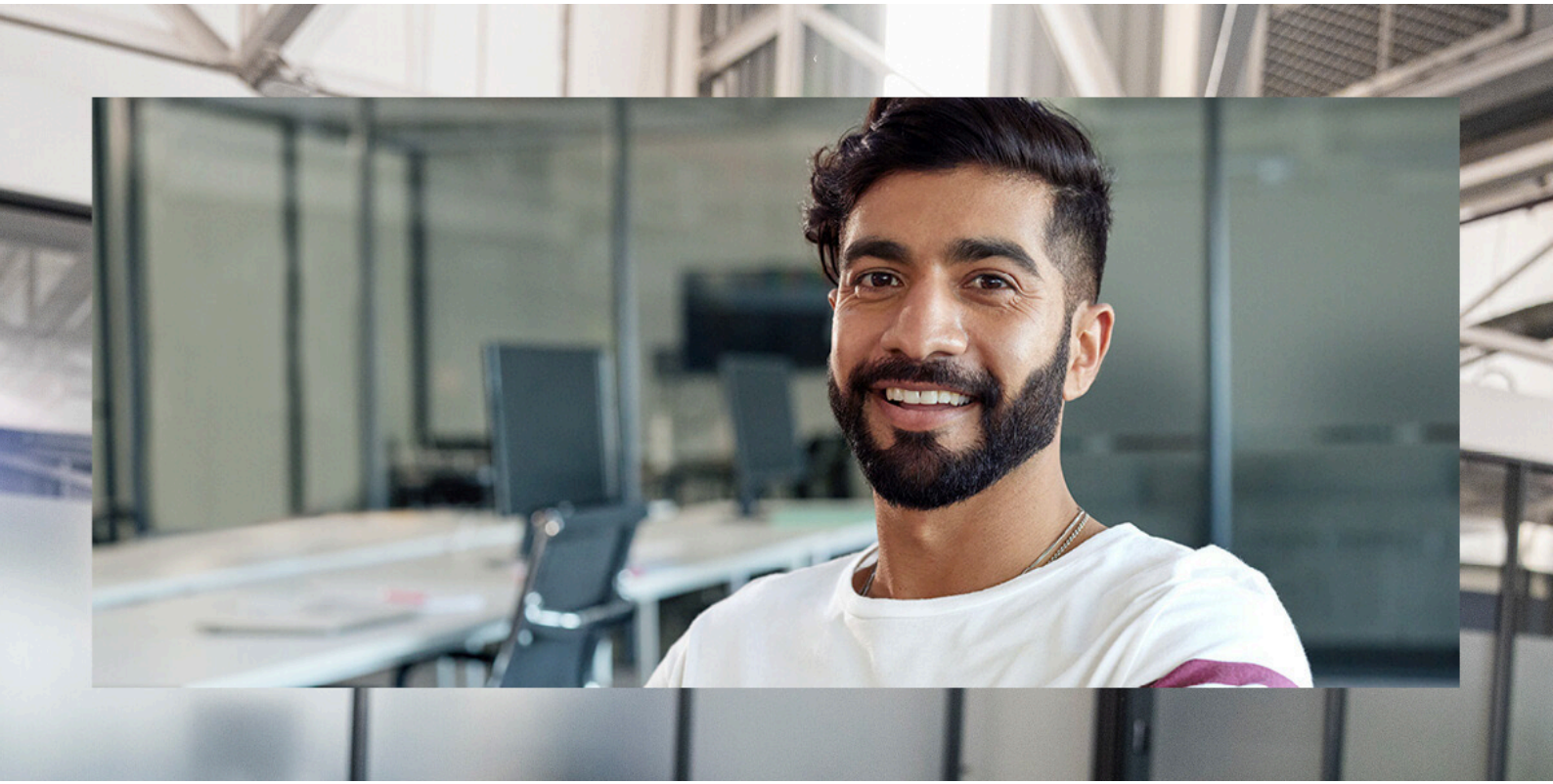
HPE GreenLake

Technical white paper

HPE Ezmeral Data Fabric Software

Global data plane



Contents

Introduction

The ability to see a logical view of all the globally distributed data that a company possesses, regardless of where it is physically, is an essential enabler to extract the full value from that data.

The global data plane feature of HPE Ezmeral Data Fabric Software provides a view into files objects, databases, and event streams that can be in separate clusters across multiple edge, on-premises, colocation, and cloud environments with single security management and audit plane. It enables the development of new location-independent applications to move smoothly across bare metal, private cloud, public cloud, or at the edge. The global data plane from HPE Ezmeral Data Fabric Software aggregates multiple data type information into a unified structure and simplifies the conceptual design of large systems and allows multiple applications to work together on the same data sets.

End users get a unified view of multiple data types allowing them access to these resources without having to be aware of the physical data location. This allows for easy data management, minimal overhead, and distributed scale since the data can now be spread across, on-premises, multiple clouds, and across multiple geographic regions (see Figure 1).

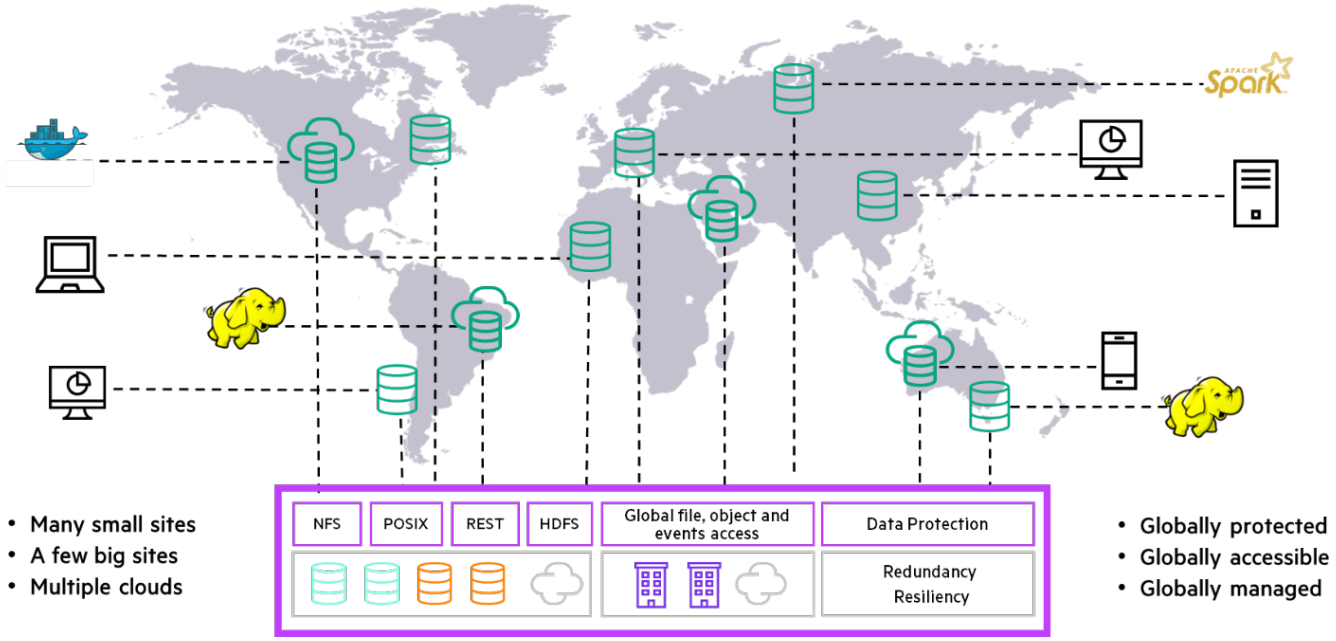


Figure 1. HPE Ezmeral Data Fabric global data plane deployment

A global data plane is comprised of multiple global namespaces, each global namespace consists of one or more data fabrics. A data fabric can belong to multiple global namespaces within the global data plane.

HPE Ezmeral Data Fabric Software

HPE Ezmeral Data Fabric Software provides customers with a simplified, consistent data plane that federates data across multiple sources, multiple formats (files, objects, tables, and streams), and across hybrid deployments. It is a single solution that spans edge-to-cloud deployments and combines most popular analytic data formats into a single data plane. HPE Ezmeral Data Fabric Software simplifies data management with built-in security, data management and lineage and governance across hybrid data deployments allowing customers to spend more time focused on data insights instead of discovering, integrating, and normalizing hybrid data across edge to cloud.

- Global visibility
- Multiformat support
- Simplified data management
- Security
- Optimized for analytics
- Predictable economics with consumption-based pricing

Use cases

There are multiple use cases for the global data plane in a modern data-first enterprise.

Cloud bursting

Enterprises frequently need to host certain data in very specific locations due to reasons of governance and compliance. They tend to hold this data in on-prem data centers with tight security and access controls and in specific geographical locations. However, research and development, beta testing of new products, and such do not necessarily need to adhere to such strict limitations. For these applications, being able to quickly and easily instantiate and tear down resources, such as storage, compute, GPU, are of prime importance. In these instances, being able to use public cloud resources is an efficient and cost-effective way to operate. Public cloud providers, typically operating on a consumption-based model, allow enterprises to monitor, manage, and control costs. However, when a new application has been developed and is ready to go into production, for the reasons mentioned previously, enterprises may want to deploy that application in their own data centers. This use of public cloud infrastructure for short term use is known as cloud bursting and is easily implemented within the HPE Ezmeral Data Fabric. Turning up a new instance of HPE Ezmeral Data Fabric Software in AWS or Azure can be done in minutes, and when the resource is no longer required, it can be torn down instantly.

Data highway

Large enterprises will invariably have multiple data types and sources situated in geographically dispersed locations. Having direct access to global data by a native data plane that spans all hybrid locations and combines files, objects, table, and streaming data into a single trusted data source is a powerful advantage. It provides at-a-glance visibility and direct access to data and once set up, automated policies can be created to consistently apply security, data sovereignty, and management configurations across all environments. This data highway can be realized with the global namespace feature within HPE Ezmeral Data Fabric Software. The global namespace stitches multiple data fabrics together allowing direct access to the data as well as data sharing and team collaboration. Additionally, configuration, geofencing, and management can also be established through the single sign-on and data management UI that is core to HPE Ezmeral Data Fabric Software.

Stretch clusters

Some customers prefer to have a single data fabric rather than a group of data fabrics configured hierarchically in a global namespace. For these customers, the concept of stretch clusters, which is a data fabric with nodes in different physical locations, is a perfect solution. Data Fabric clusters consisting of multiple nodes are typically installed in the same location, and often in the same rack. This is important for many reasons, including resiliency and redundancy. The data fabric is also very sensitive to latency, and deployments must be planned carefully to avoid jitter and delay between nodes in a cluster. However, with low-latency direct connections, such as dark fiber, between the nodes in a data fabric, the nodes can be dispersed across multiple geographic regions.

Create the global data plane

The global data plane is defined when data fabric clusters are deployed. It is important to plan the global data plane architecture before beginning the configuration of the clusters.

When creating a global data plane, the following factors should be considered:

- Location of the data fabric in relation to your application's access needs
- Data fabric naming convention for each location
- Volume names as it relates to each location for your application needs
- Any business continuity and disaster recovery requirements should be accounted for including additional capacity for cross-cluster replication of data

By creating a plan for your data, you can apply security controls more easily for user and application needs, as well as protect your data.

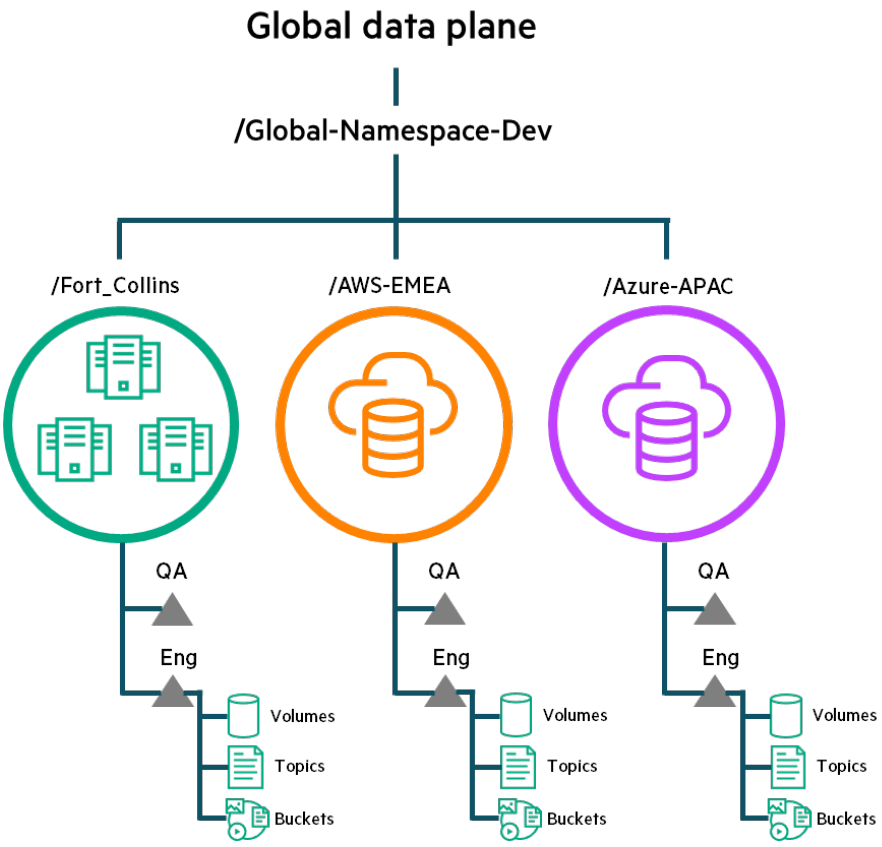


Figure 2. An example of a global namespace

Creating and configuring a global data plane depends on the deployment model used. HPE Ezmeral Data Fabric can be deployed in a number of different modes:

- Traditional, customer managed deployment based on customer resources
- As a fully managed subscription service

Before beginning the configuration of the clusters, it is important to plan the global data plane architecture. For example, see Figure 2.

HPE Ezmeral Data Fabric Software

In a fully managed deployment, a global namespace can be created from within the user-interface (UI).

To create a new namespace, log in to the UI and navigate to the fabric visualization view by clicking the icon  . Click the namespace icon in the create menu bar, and in the pop-up menu, enter the top level of the namespace you wish to create.

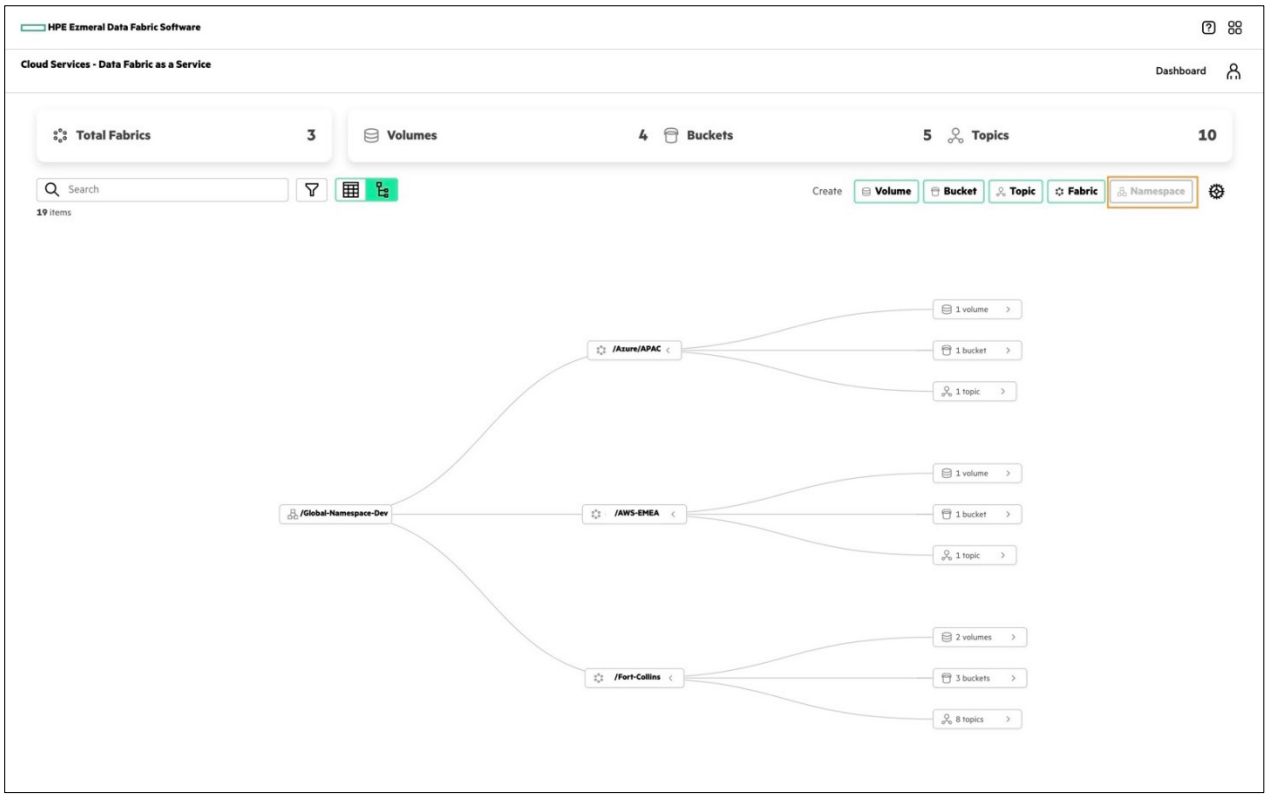


Figure 3. Configure a global data plane

Figure 3 shows a screenshot from an HPE Ezmeral Data Fabric UI presenting a graphical representation of three data fabric clusters that are connected in a hybrid cloud deployment. Two are deployed in public clouds, one in Europe named "/AWS/EMEA", and one in Asia named "/Azure/APAC", and the third is an on-prem deployment in a colocation space in North America, named "/Fort_Collins".

Customer managed

In the customer-managed version of the HPE Ezmeral Data Fabric Software, the global data plane is created when the data fabric clusters are configured. After all the clusters have been created, cross-cluster security must be configured. This is done using the

`configure-crosscluster.sh utility`.

Examples for configuring cross-cluster features are available at docs.datafabric.hpe.com/72/ReferenceGuide/Examples.html.

Note

Cross-cluster connections cannot be performed from the management control system (MCS).

Summary

The global data plane of HPE Ezmeral Data Fabric aggregates files, database, object, and event information into a unified structure regardless of whether it is on-prem, in a public or private cloud, or an edge deployment. This allows developers, data engineers, and data analysts, and data scientists to develop, deploy, and run applications, along with workloads that can directly access those files, objects, databases, and event streams, as if they were local. It simplifies the design of large integrated systems and allows multiple applications to work together on the same data sets.

Learn more at
HPE.com/datafabric

Explore **HPE GreenLake**

Chat now (sales)

a00118158ENW, Rev. 4

Hewlett Packard Enterprise