

# Hybrid Multicloud MLOps with Domino Data Lab and NetApp

NetApp Solutions

NetApp March 19, 2024

# **Table of Contents**

Ηу	brid Multicloud MLOps with Domino Data Lab and NetApp	. 1
	Hybrid Multicloud MLOps with Domino Data Lab and NetApp	. 1
	Technology Overview	. 1
	Architecture	. 4
	Initial Setup.	. 5
	Expose Existing NetApp Volumes to Domino	. 6
	Access the same Data Across Different Environments	. 9
	Where to Find Additional Information	15

# Hybrid Multicloud MLOps with Domino Data Lab and NetApp

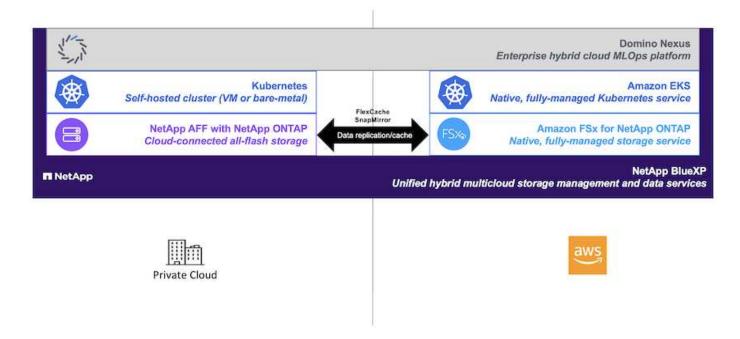
# Hybrid Multicloud MLOps with Domino Data Lab and NetApp

Mike Oglesby, NetApp

Organizations all over the world are currently adopting AI to transform their businesses and processes. Because of this, AI-ready compute infrastructure is often in short supply. Enterprises are adopting hybrid multicloud MLOps architectures in order to take advantage of available compute environments across different regions, data centers, and clouds - balancing cost, availability, and performance.

Domino Nexus, from Domino Data Lab, is a unified MLOps control plane that lets you run data science and machine learning workloads across any compute cluster — in any cloud, region, or on-premises. It unifies data science silos across the enterprise, so you have one place to build, deploy, and monitor models. Likewise, NetApp's hybrid cloud data management capabilities enable you to bring your data to your jobs and workspaces, no matter where they are running. When you pair Domino Nexus with NetApp, you have the flexibility to schedule workloads across environments without having to worry about data availability. In other words, you have the ability to send your workloads and your data to the appropriate compute environment, enabling you to accelerate your Al deployments while navigating regulations around data privacy and sovereignty.

This solution demonstrates the deployment of a unified MLOps control plane incorporating an on-premises Kubernetes cluster and an Elastic Kubernetes Service (EKS) cluster running in Amazon Web Services (AWS).



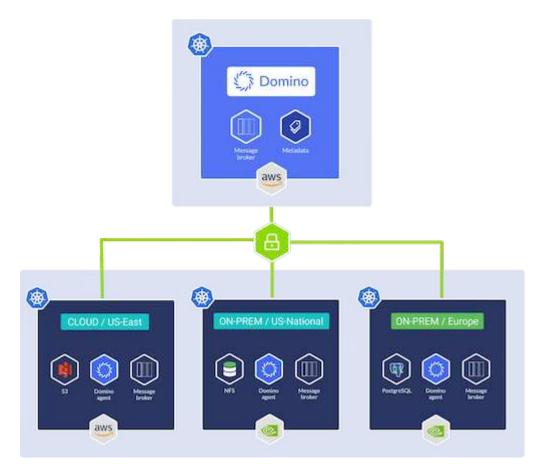
# **Technology Overview**

#### **Domino Data Lab**

Domino Data Lab powers model-driven businesses with its leading Enterprise AI platform trusted by over 20% of the Fortune 100. Domino accelerates the development and deployment of data science work while increasing collaboration and governance. With Domino, enterprises worldwide can develop better medicines, grow more productive crops, build better cars, and much more. Founded in 2013, Domino is backed by Coatue Management, Great Hill Partners, Highland Capital, Sequoia Capital and other leading investors.

Domino lets enterprises and their data scientists build, deploy and manage AI on a unified, end-to-end platform — fast, responsibly and cost-effectively. Teams can access all of the data, tools, compute, models, and projects they need across any environment, so they can collaborate, reuse past work, track models in production to improve accuracy, standardize with best practices, and make AI responsible and governed.

- **Open and Flexible:** Access the broadest ecosystem of open source and commercial tools, and infrastructure, for the best innovations and no vendor lock-in.
- **System of Record:** Central hub for Al operations and knowledge across the enterprise, enabling best practices, cross-functional collaboration, faster innovation, and efficiency.
- **Integrated:** Integrated workflows and automation built for enterprise processes, controls, and governance satisfy your compliance and regulatory needs.
- **Hybrid Multicloud:** Run Al workloads close to your data anywhere on-premises, hybrid, any cloud or multi-cloud for lower cost, optimal performance and compliance.



#### **Domino Nexus**

Domino Nexus is a single pane of glass that lets you run data science and machine learning workloads across any compute cluster — in any cloud, region, or on-premises. It unifies data science silos across the enterprise,

so you have one place to build, deploy, and monitor models.

## NetApp BlueXP

NetApp BlueXP unifies all of NetApp's storage and data services into a single tool that lets you build, protect, and govern your hybrid multicloud data estate. It delivers a unified experience for storage and data services across on-premises and cloud environments, and enables operational simplicity through the power of AlOps, with the flexible consumption parameters and integrated protection required for today's cloud-led world.

## **NetApp ONTAP**

ONTAP 9, the latest generation of storage management software from NetApp, enables businesses to modernize infrastructure and transition to a cloud-ready data center. Leveraging industry-leading data management capabilities, ONTAP enables the management and protection of data with a single set of tools, regardless of where that data resides. You can also move data freely to wherever it is needed: the edge, the core, or the cloud. ONTAP 9 includes numerous features that simplify data management, accelerate, and protect critical data, and enable next generation infrastructure capabilities across hybrid cloud architectures.

#### Simplify data management

Data management is crucial to enterprise IT operations and data scientists so that appropriate resources are used for AI applications and training AI/ML datasets. The following additional information about NetApp technologies is out of scope for this validation but might be relevant depending on your deployment.

ONTAP data management software includes the following features to streamline and simplify operations and reduce your total cost of operation:

- Inline data compaction and expanded deduplication. Data compaction reduces wasted space inside storage blocks, and deduplication significantly increases effective capacity. This applies to data stored locally and data tiered to the cloud.
- Minimum, maximum, and adaptive quality of service (AQoS). Granular quality of service (QoS) controls help maintain performance levels for critical applications in highly shared environments.
- NetApp FabricPool. Provides automatic tiering of cold data to public and private cloud storage options, including Amazon Web Services (AWS), Azure, and NetApp StorageGRID storage solution. For more information about FabricPool, see TR-4598: FabricPool best practices.

#### Accelerate and protect data

ONTAP delivers superior levels of performance and data protection and extends these capabilities in the following ways:

- Performance and lower latency. ONTAP offers the highest possible throughput at the lowest possible latency.
- Data protection. ONTAP provides built-in data protection capabilities with common management across all platforms.
- NetApp Volume Encryption (NVE). ONTAP offers native volume-level encryption with both onboard and External Key Management support.
- Multitenancy and multifactor authentication. ONTAP enables sharing of infrastructure resources with the highest levels of security.

#### **Future-proof infrastructure**

ONTAP helps meet demanding and constantly changing business needs with the following features:

- Seamless scaling and nondisruptive operations. ONTAP supports the nondisruptive addition of capacity to
  existing controllers and to scale-out clusters. Customers can upgrade to the latest technologies, such as
  NVMe and 32Gb FC, without costly data migrations or outages.
- Cloud connection. ONTAP is the most cloud-connected storage management software, with options for software-defined storage and cloud-native instances in all public clouds.
- Integration with emerging applications. ONTAP offers enterprise-grade data services for next generation
  platforms and applications, such as autonomous vehicles, smart cities, and Industry 4.0, by using the same
  infrastructure that supports existing enterprise apps.

## Amazon FSx for NetApp ONTAP

Amazon FSx for NetApp ONTAP is a first-party, fully managed AWS service that provides highly reliable, scalable, high-performing, and feature-rich file storage built on NetApp's popular ONTAP file system. FSx for ONTAP combines the familiar features, performance, capabilities, and API operations of NetApp file systems with the agility, scalability, and simplicity of a fully managed AWS service.

## **NetApp Astra Trident**

Astra Trident enables consumption and management of storage resources across all popular NetApp storage platforms, in the public cloud or on premises, including ONTAP (AFF, FAS, Select, Cloud, Amazon FSx for NetApp ONTAP), Element software (NetApp HCI, SolidFire), Azure NetApp Files service, and Cloud Volumes Service on Google Cloud. Astra Trident is a Container Storage Interface (CSI) compliant dynamic storage orchestrator that natively integrates with Kubernetes.

#### **Kubernetes**

Kubernetes is an open source, distributed, container orchestration platform that was originally designed by Google and is now maintained by the Cloud Native Computing Foundation (CNCF). Kubernetes enables the automation of deployment, management, and scaling functions for containerized applications, and is the dominant container orchestration platform in enterprise environments.

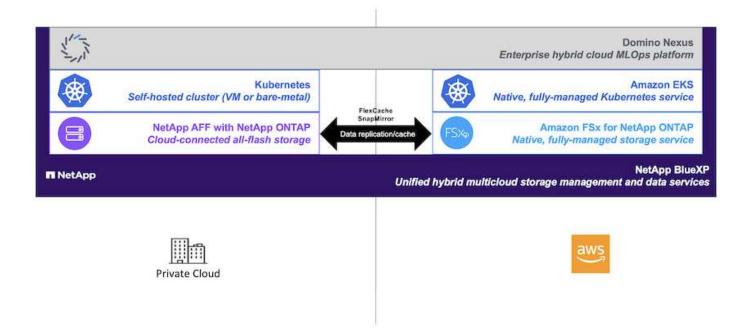
## Amazon Elastic Kubernetes Service (EKS)

Amazon Elastic Kubernetes Service (Amazon EKS) is a managed Kubernetes service in the AWS cloud. Amazon EKS automatically manages the availability and scalability of the Kubernetes control plane nodes responsible for scheduling containers, managing application availability, storing cluster data, and other key tasks. With Amazon EKS, you can take advantage of all the performance, scale, reliability, and availability of AWS infrastructure, as well as integrations with AWS networking and security services.

## **Architecture**

This solution combines Domino Nexus' hybrid multicloud workload scheduling capabilities with NetApp data services to create a unified hybrid cloud MLOps platform. See the following table for details.

Component	Name	Environment
MLOps Control Plane	Domino Enterprise Al Platform with Domino Nexus	AWS
MLOps Platform Compute Environments	Domino Nexus Data Planes	AWS, On-premises data center
On-premises Compute Platform	Kubernetes with NetApp Astra Trident	On-premises data center
Cloud Compute Platform	Amazon Elastic Kubernetes Service (EKS) with NetApp Astra Trident	AWS
On-premises Data Platform	NetApp storage appliance powered by NetApp ONTAP	On-premises data center
Cloud Data Platform	Amazon FSx for NetApp ONTAP	AWS



# **Initial Setup**

This section describes the initial setup tasks that need to be performed in order to utilize Domino Nexus with NetApp data services in a hybrid environment incorporating an onpremises data center and AWS.

## **Prerequisites**

Before you perform the steps that are outlined in this section, we assume that you have already performed the following tasks:

- You have already deployed and configured your on-premises NetApp ONTAP storage platform. For more information, refer to the NetApp product documentation.
- You have already provisioned an Amazon FSx for NetApp ONTAP instance in AWS. For more information, refer to the Amazon FSx for NetApp ONTAP product page.

- You have already provisioned a Kubernetes cluster in your on-premises data center. For more information, refer to the Domino admin guide.
- You have already provisioned an Amazon EKS cluster in AWS. For more information, refer to the Domino admin guide.
- You have installed NetApp Astra Trident in your on-premises Kubernetes cluster. Additionally, you have configured this Trident instance to use your on-premises NetApp ONTAP storage platform when provisioning and managing storage resources. For more information, refer to the NetApp Astra Trident documentation.
- You have installed NetApp Astra Trident in your Amazon EKS cluster. Additionally, you have configured this Trident instance to use your Amazon FSx for NetApp ONTAP instance when provisioning and managing storage resources. For more information, refer to the NetApp Astra Trident documentation.
- You must have bi-directional network connectivity between your on-premises data center and your Virtual Private Cloud (VPC) in AWS. For more details on the various options for implementing this, refer to the Amazon Virtual Private Network (VPN) documentation.

## Install the Domino Enterprise Al Platform in AWS

To install the Domino Enterprise MLOps Platform in AWS, follow the instructions outlined in Domino admin guide. You must deploy Domino in the same Amazon EKS cluster that you previously provisioned. Additionally, NetApp Astra Trident must already be installed and configured in this EKS cluster, and you must specify a Trident-managed storage class as the shared storage class in your domino.yml install configuration file.



Refer to the Domino install configuration reference guide for details on how to specify a shared storage class in your domino.yml install configuration file.



Technical Report TR-4952 walks through the deployment of Domino in AWS with Amazon FSx for NetApp ONTAP and may be a useful reference for troubleshooting any issues that arise.

### **Enable Domino Nexus**

Next, you must enable Domino Nexus. Refer to the Domino admin guide for details.

## Deploy a Domino Data Plane in your On-premises Data Center

Next, you must deploy a Domino Data Plane in your on-premises data center. You must deploy this data plane in the on-premises Kubernetes cluster that you previously provisioned. Additionally, NetApp Astra Trident must already be installed and configured in this Kubernetes cluster. Refer to the Domino admin guide for details.

## **Expose Existing NetApp Volumes to Domino**

This section describes the tasks that need to be performed in order to expose existing NetApp ONTAP NFS volumes to the Domino MLOps platform. These same steps apply both on-premises and in AWS.

## Why Expose NetApp ONTAP Volumes to Domino?

Using NetApp volumes in conjunction with Domino provides the following benefits:

You can execute workloads against extremely large datasets by taking advantage of NetApp ONTAP's

scale-out capabilities.

- You can execute workloads across multiple compute nodes without having to copy your data to the individual nodes.
- You can take advantage of NetApp's hybrid multicloud data movement and sync capabilities in order to access your data across multiple data centers and/or clouds.
- You want to be able to quickly and easily create a cache of your data in a different data center or cloud.

## **Expose Existing NFS Volumes that were not Provisioned by Astra Trident**

If your existing NetApp ONTAP NFS volume was not provisioned by Astra Trident, follow the steps outlined in this sub-section.

#### **Create PV and PVC in Kubernetes**

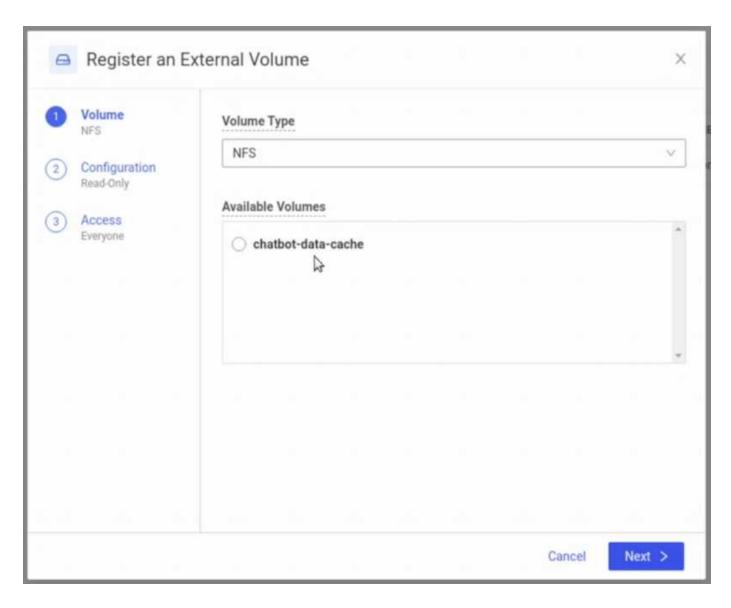


For on-premises volumes, create the PV and PVC in your on-premises Kubernetes cluster. For Amazon FSx for NetApp ONTAP volumes, create the PV and PVC in Amazon EKS.

First, you must create a persistent volume (PV) and persistent volume claim (PVC) in your Kubernetes cluster. To create the PV and PVC, use the NFS PV/PVC example from the Domino admin guide and update the values to reflect to your environment. Be sure to specify the correct values for the namespace, nfs.path, and nfs.server fields. Additionally, we recommend giving your PV and PVC unique names that represent that nature of the data that is stored on the corresponding ONTAP NFS volume. For example, if the volume contains images of manufacturing defects, you might name the PV, pv-mfg-defect-images, and the PVC, pvc-mfg-defect-images.

## Register External Data Volume in Domino

Next, you must register an external data volume in Domino. To register an external data volume, refer to the instructions in the Domino admin guide. When registering the volume, be sure to select "NFS" from the 'Volume Type' drop-down menu. After selecting "NFS", you should see your PVC in the 'Available Volumes' list.



## **Expose Existing Volumes that were Provisioned by Astra Trident**

If your existing volume was provisioned by Astra Trident, follow the steps outlined in this sub-section.

#### **Edit Existing PVC**

If your volume was provisioned by Astra Trident, then you already have a persistent volume claim (PVC) corresponding to your volume. In order to expose this volume to Domino, you must edit the PVC and add the following label to the list of labels in the metadata.labels field:

```
"dominodatalab.com/external-data-volume": "Generic"
```

### Register External Data Volume in Domino

Next, you must register an external data volume in Domino. To register an external data volume, refer to the instructions in the Domino admin guide. When registering the volume, be sure to select "Generic" from the 'Volume Type' drop-down menu. After selecting "Generic", you should see your PVC in the 'Available Volumes' list.

## Access the same Data Across Different Environments

This section describes the tasks that need to be performed in order to access the same data across different compute environments. In the Domino MLOps platform, compute environments are referred to "data planes." Follow the tasks outlined in this section if your data resides on a NetApp volume in one data plane, but you need to access it in another data plane. This type of scenario is often referred to as "bursting" or, when the destination environment is the cloud, "cloud bursting." This capability is often needed when dealing with constrained or over-subscribed compute resources. For example, if your onpremises compute cluster is over-subscribed, you may want to schedule workloads to the cloud where they can be started immediately.

There are two recommended options for accessing a NetApp volume that resides in a different data plane. These options are outlined in the sub-sections below. Choose one of these options depending on your specific requirements. The benefits and drawbacks of the two options are described in the following table.

Option	Benefits	Drawbacks
Option 1 - Cache	<ul> <li>Simpler workflow</li> <li>Ability to cache a subset of data based on needs</li> <li>Ability to write data back to source</li> <li>No remote copy to manage</li> </ul>	- Increased latency on initial data access as cache is hydrated.
Option 2 - Mirror	<ul> <li>Full copy of source volume</li> <li>No increased latency due to cache hydration (after mirror operation is complete)</li> </ul>	<ul><li>Must wait for mirror operation to complete before accessing data</li><li>Must manage a remote copy</li><li>No ability to write back to source</li></ul>

## Option 1 - Create a Cache of a Volume that Resides in a Different Data Plane

With NetApp FlexCache technology, you can create a cache of a NetApp volume that resides in a different data plane. For example, if you have a NetApp volume in your on-premises data plane, and you need to access that volume in your AWS data plane, you can create a cache of the volume in AWS. This section outlines the tasks that need to be performed in order to create a cache of a NetApp volume that resides in a different data plane.

#### **Create FlexCache Volume in Destination Environment**



If the destination environment is your on-premises data center, you will create the FlexCache volume on your on-premises ONTAP system. If the destination environment is AWS, you will create the FlexCache volume on your Amazon FSx for NetApp ONTAP instance.

First, you must create a FlexCache volume in the destination environment.

We recommend using BlueXP to create the FlexCache volume. To create a FlexCache volume with BlueXP, follow the instructions outlined in the BlueXP volume caching documentation.

If you prefer not to use BlueXP, you can use ONTAP System Manager or the ONTAP CLI to create the FlexCache volume. To create a FlexCache volume with System Manager, refer to the instructions outlined in the ONTAP documentation. To create a FlexCache volume with the ONTAP CLI, refer to the instructions

outlined in the ONTAP documentation.

If you wish to automate this process, you can use the BlueXP API, the ONTAP REST API, or the ONTAP Ansible collection.



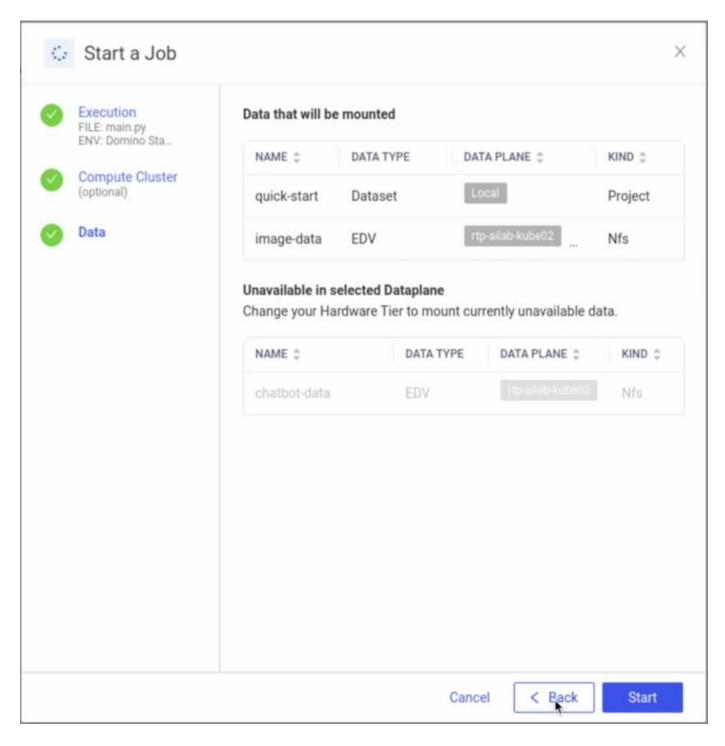
System Manager is not available in Amazon FSx for NetApp ONTAP.

### **Expose FlexCache Volume to Domino**

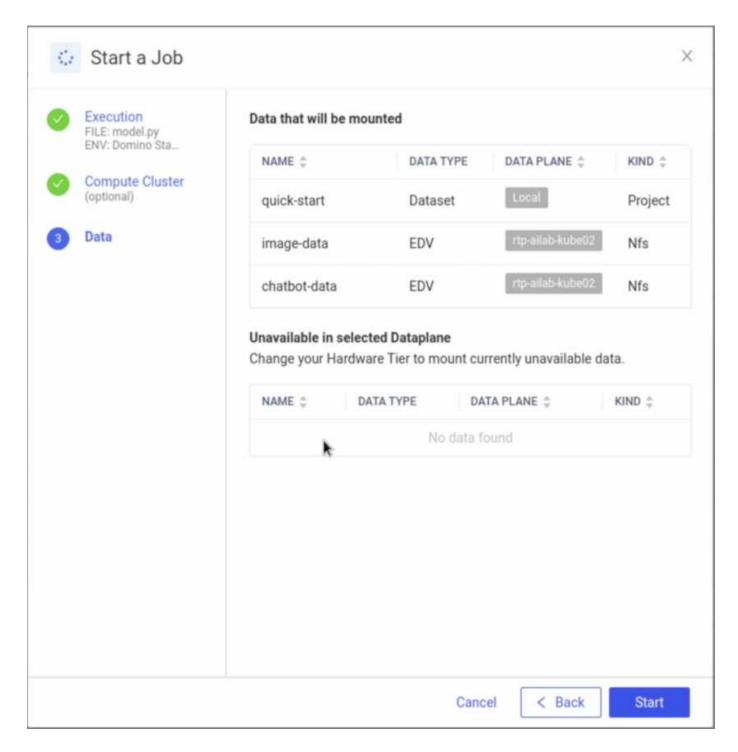
Next, you must expose the FlexCache volume to the Domino MLOps platform. To expose the FlexCache volume to Domino, follow the instructions outlined in the 'Expose Existing NFS Volumes that were not Provisioned by Astra Trident' sub-section of the 'Expose Existing NetApp Volumes to Domino' section of this solution.

Now, you will be able to mount the FlexCache volume when launching jobs and workspaces in the destination data plane as shown in the following screenshots.

**Before Creating FlexCache Volume** 



After Exposing FlexCache Volume to Domino



## Option 2 - Replicate a Volume that Resides in a Different Data Plane

With NetApp SnapMirror data replication technology, you can create a copy of a NetApp volume that resides in a different data plane. For example, if you have a NetApp volume in your on-premises data plane, and you need to access that volume in your AWS data plane, you can create a copy of the volume in AWS. This section outlines the tasks that need to be performed in order to create a copy of a NetApp volume that resides in a different data plane.

#### **Create SnapMirror Relationship**

First, you must create a SnapMirror relationship between your source volume and a new destination volume in the destination environment. Note that the destination volume will be created as part of the process of creating

the SnapMirror relationship.

We recommend using BlueXP to create the SnapMirror relationship. To create a SnapMirror relationship with BlueXP, follow the instructions outlined in the BlueXP replication documentation.

If you prefer not to use BlueXP, you can use ONTAP System Manager or the ONTAP CLI to create the SnapMirror relationship. To create a SnapMirror relationship with System Manager, refer to the instructions outlined in the ONTAP documentation. To create a SnapMirror relationship with the ONTAP CLI, refer to the instructions outlined in the ONTAP documentation.

If you wish to automate this process, you can use the BlueXP API, the ONTAP REST API, or the ONTAP Ansible collection.



System Manager is not available in Amazon FSx for NetApp ONTAP.

#### **Break SnapMirror Relationship**

Next, you must break the SnapMirror relationship in order to activate the destination volume for data access. Wait until the initial replication is complete before performing this step.



You can determine whether or not the replication is complete by checking the mirror state in BlueXP, ONTAP System Manager, or the ONTAP CLI. When the replication is complete, the mirror state will be "snapmirrored".

We recommend using BlueXP to break the SnapMirror relationship. To break a SnapMirror relationship with BlueXP, follow the instructions outlined in the BlueXP replication documentation.

If you prefer not to use BlueXP, you can use ONTAP System Manager or the ONTAP CLI to break the SnapMirror relationship. To break a SnapMirror relationship with System Manager, refer to the instructions outlined in the ONTAP documentation. To break a SnapMirror relationship with the ONTAP CLI, refer to the instructions outlined in the ONTAP documentation.

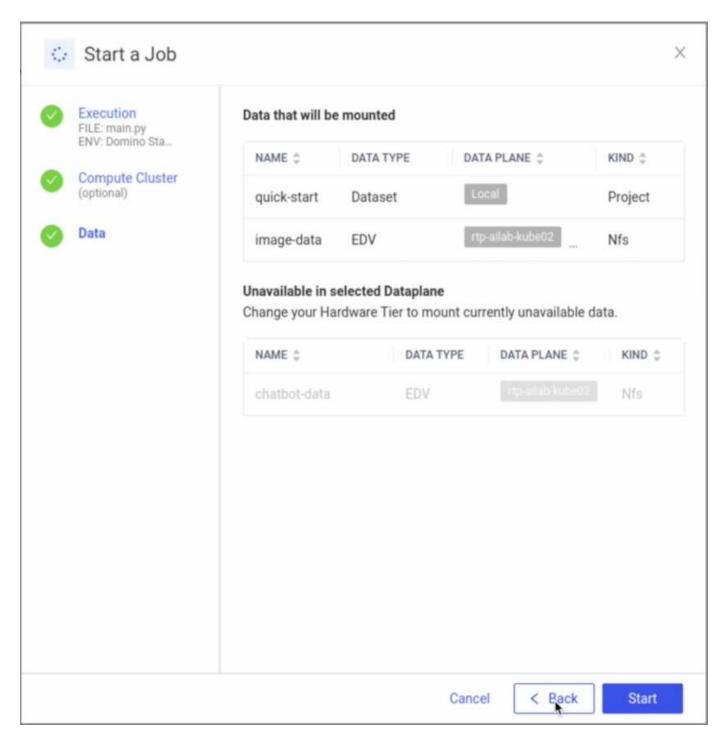
If you wish to automate this process, you can use the BlueXP API, the ONTAP REST API, or the ONTAP Ansible collection.

#### **Expose Destination Volume to Domino**

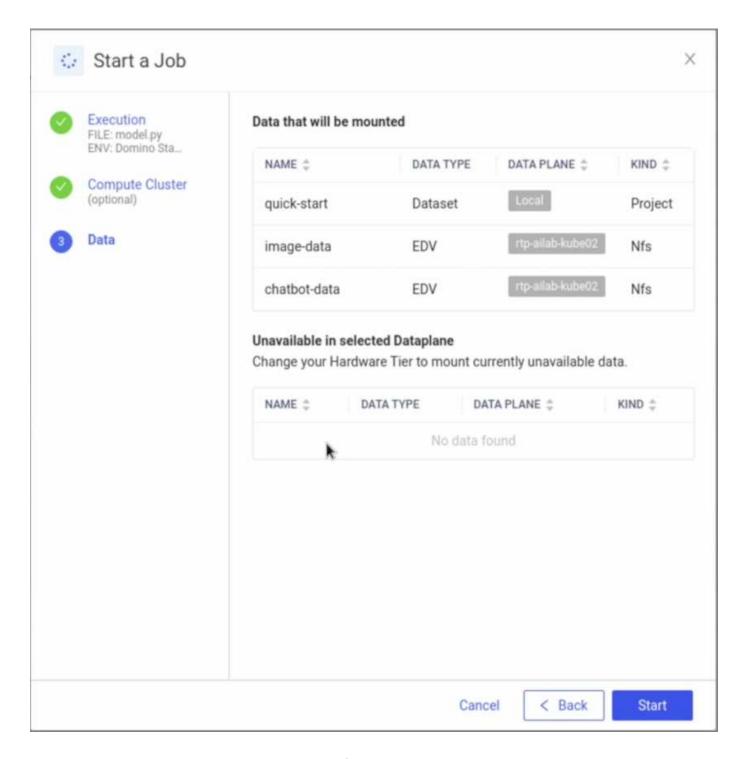
Next, you must expose the destination volume to the Domino MLOps platform. To expose the destination volume to Domino, follow the instructions outlined in the 'Expose Existing NFS Volumes that were not Provisioned by Astra Trident' sub-section of the 'Expose Existing NetApp Volumes to Domino' section of this solution.

Now, you will be able to mount the destination volume when launching jobs and workspaces in the destination data plane as shown in the following screenshots.

#### **Before Creating SnapMirror Relationship**



After Exposing Destination Volume to Domino



## Where to Find Additional Information

To learn more about the information described in this document, refer to the following documents and/or websites:

Domino Data Lab

https://domino.ai

Domino Nexus

https://domino.ai/platform/nexus

NetApp BlueXP

https://bluexp.netapp.com

• NetApp ONTAP data management software

https://www.netapp.com/data-management/ontap-data-management-software/

NetApp Al Solutions

https://www.netapp.com/artificial-intelligence/

## **Acknowledgments**

- Josh Mineroff, Director of SA for Tech Alliances, Domino Data Lab
- Nicholas Jablonski, Field CTO, Domino Data Lab
- Prabu Arjunan, Solution Architect, NetApp
- Brian Young, Global Alliance Director, Technology Alliance Partners, NetApp

#### Copyright information

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

#### **Trademark information**

NETAPP, the NETAPP logo, and the marks listed at <a href="http://www.netapp.com/TM">http://www.netapp.com/TM</a> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.