

## Cohesity Microsoft Azure Data Box Integration

---



## Table of Contents

Introduction.....	2
Audience.....	2
Requirements.....	2
Assumptions.....	2
Order Microsoft Azure Data Box.....	3
Requesting.....	3
Order Details.....	4
Shipping Address.....	5
Notification Details.....	5
Summary.....	6
Data Box Deployment.....	6
Shipping.....	7
Data Box Installation Setup.....	8
Cohesity DataPlatform.....	8
General.....	8
Configure External Targets.....	9
Configure Protection Policy.....	9
Create Protection Job.....	10
Return Data Box to Microsoft Azure.....	11
Prepare to Ship.....	11
Physically Prepare & Ship.....	12
Discover.....	12
Cohesity - Resume Protection - After Import of data into Azure.....	13
Redirect External Targets.....	13
Resume Archive.....	13
Restore a VM.....	14
Summary.....	15
About Cohesity.....	15
About Azure Data Box.....	15

## Introduction

Data is growing at an exponential rate. It's more important than ever to reduce data silos, in addition to taking advantage of efficiencies in order to reduce your costs. Additionally, since your data holds the keys to the success of your business, it's critical that it's protected and available to be recovered or leveraged wherever you operate. Further cost reduction strategies include consolidation of secondary and tertiary data centers and involve using the public cloud as part of your data protection strategy.

The main challenge with protecting your data AND leveraging the cloud at the same time is how to get your data to the cloud in the first place. Is your data network robust enough to facilitate your initial data transfer requirements? An integral part of reducing your data transfer requirements is to reduce the amount of data you need to send to the cloud in the first place. Cohesity leverages its robust global deduplication and compression capabilities to dramatically reduce the amount of storage required to protect your data. The next thing to consider is after Cohesity has applied its storage efficiencies, is much time it will take to transfer your protected data to the cloud.

Here's a data transfer example:

You have 80TB of data you want to transfer to the cloud and you have an existing 200Mbps connection. You know you currently utilize around 50% of that connection for normal operations, leaving approximately 100Mbps of bandwidth to transfer your data. Assuming no connectivity hiccups, it will still take over 80 days to transmit your data over the network connection. For many businesses, that's simply too long.

Luckily, [Microsoft Azure](#) has a solution for you. It's called the Azure Data Box, which is an appliance that can be shipped to your location, loaded with data, and sent back to Azure to be uploaded into your account. Microsoft uses Next Day shipping to transport the Azure Data Box to and from your data center. With that in mind, and assuming the on-premises network is a 10Gbps network, for the same 80TB example above it should take approximately a week for your data to be available in your Azure subscription. That's a huge time savings!

In this document we'll cover the integration between the Cohesity DataPlatform and Microsoft Azure Data Box. We start by protecting your data with Cohesity, and then seeding the initial archival of your data onto Azure Data Box. This ensures the protection of your data and provides a simple and secure way of getting your data into Microsoft Azure.

## Audience

Readers of this document will be IT professionals interested in revolutionizing the way they currently protect their on-premises virtual machines and data in addition to taking full advantage of Microsoft Azure. Additionally, they have large amounts of data they want to move into Azure, but face various network bandwidth constraints, and need to leverage an appliance-based data transfer approach.

Requirements

## Requirements

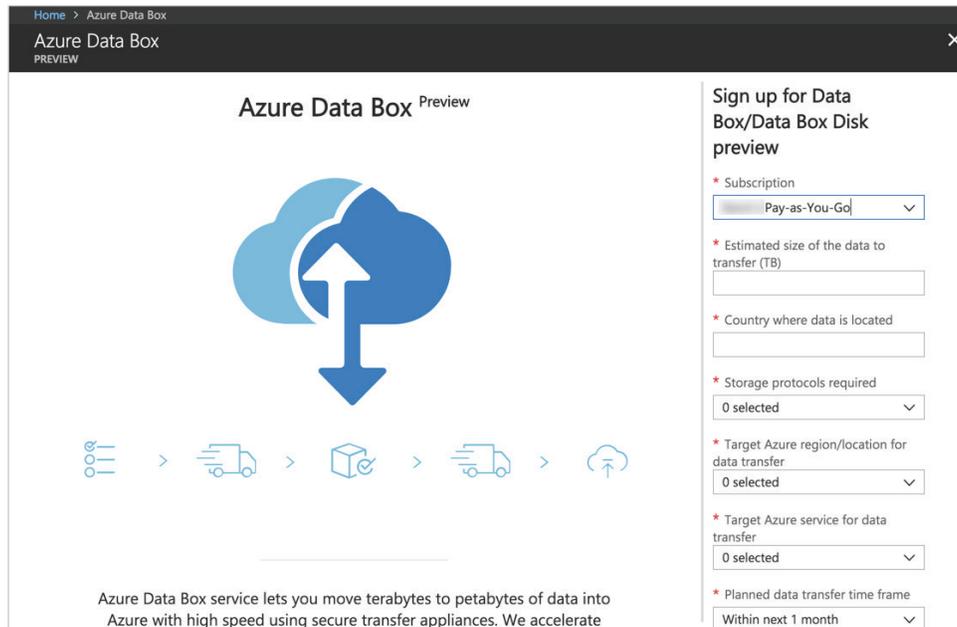
Item	Value	Provider
Microsoft Azure Subscription	Non-Free Tier	Microsoft
Microsoft Azure Data Box	4.x	Microsoft
Azure Blob Container Types	Hot, Cool	Microsoft
Azure Data Box Connectivity	NFS	n/a
Cohesity DataPlatform	5.0.1+	Cohesity

## Assumptions

The primary assumption is that you already have a Microsoft Azure cloud subscription and are familiar with how to navigate through the Microsoft Azure Portal. The second assumption is that you already have Cohesity DataPlatform deployed and running either on as a Virtual Edition, a Cloud Edition (i.e. running on Azure resources), or as a physical appliance.

## Order Microsoft Azure Data Box

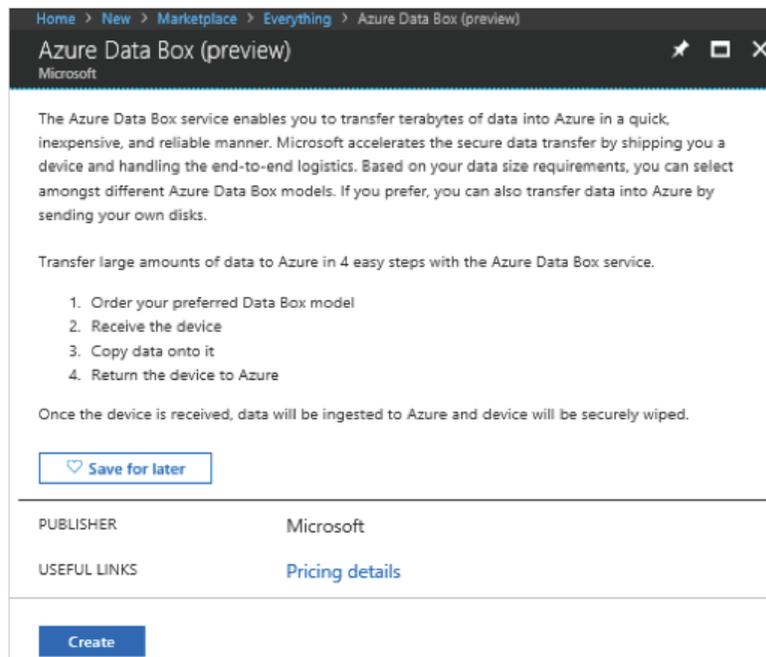
One of the first things we need to do is to request and set up an Azure Data Box. At the time of initial publication of this document, we'll first need to go to the Azure Data Box | Microsoft Azure Site, and sign up for the Preview. Clicking on the "Sign Up for the Preview >" button takes us to where we can provide the details specific to our Azure Subscription:



Please note that these instructions are a high-level summary of the necessary steps required to order an Azure Data Box. Please refer to this Azure documentation [HERE](#) for the complete walk-through guide for ordering an Azure Data Box.

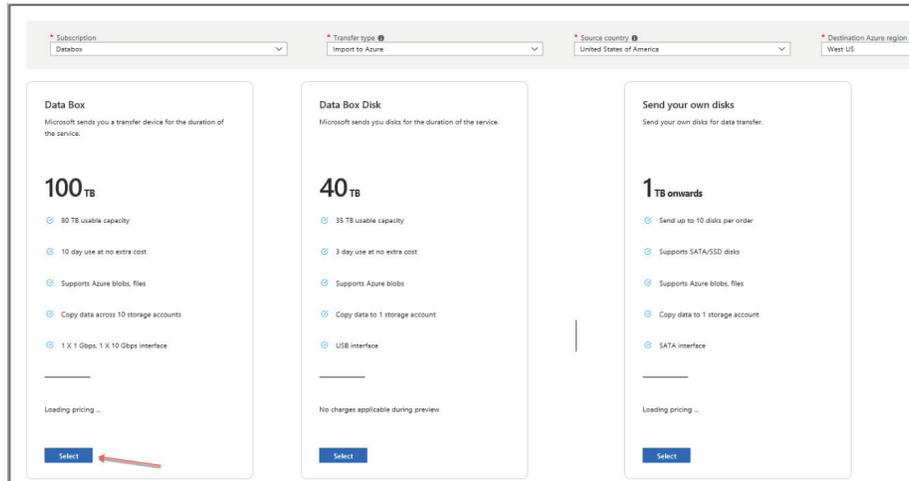
### Requesting

Once registered for the Azure Data Box preview, you are provided with an overview of how to use the Azure Data Box before creating a Data Box order.



After clicking on 'Create' we need to select the type of Azure Data Box we want to use in addition to the following Azure details:

- The Subscription we would like to use
- The Transfer Type (in this scenario: 'Import to Azure')
- Our Source Country (in this scenario: 'United States of America')
- Destination Azure Region (in this scenario: 'West US')

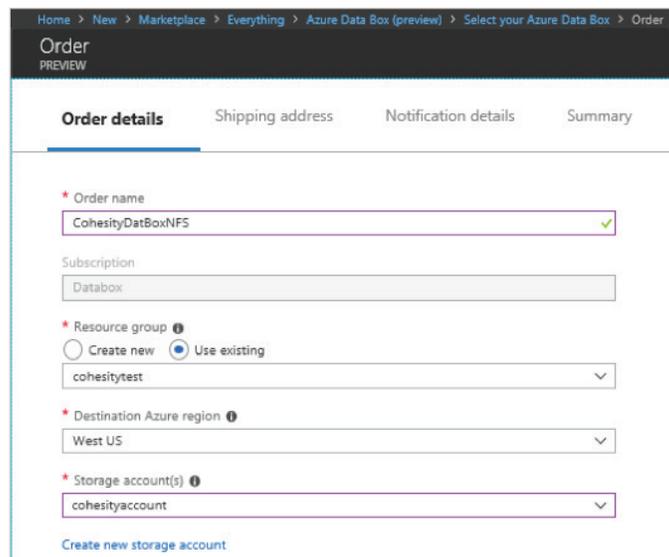


For purposes of our testing we've selected the 100TB Data Box (formatted: 80TB usable)

## Order Details

We now need to provide information about our order:

- A Name for our Order
- The Azure Resource group we want to use (or if we want to create a new Resource Group)
- The Destination Azure Region
- The Azure Storage Account we want to use (or if we want to create a new Storage Account)



## Shipping Address

Next, we need to specify the shipping information for our Data Box.

The screenshot shows the 'Shipping address' step of an order preview. The breadcrumb trail is: Home > New > Marketplace > Everything > Azure Data Box (preview) > Select your Azure Data Box > Order. The page title is 'Order PREVIEW'. The navigation tabs are 'Order details', 'Shipping address' (selected), 'Notification details', and 'Summary'. The form contains the following fields:

- \* Contact person: John ✓
- Company name: Cohesity ✓
- \* Address: 300 Park Ave, Suite 800 ✓  
Apartment, suite, unit, building, floor
- Department, c/o: (empty)
- \* Zip / Postal code: 95110 ✓
- \* City: San Jose ✓
- \* State / Province / Region: CALIFORNIA ✓
- \* Country: United States of America
- Work phone: +1 859264374 ✓ Extension: (empty)

A green checkmark icon and a message state: 'Successfully validated the address. Depending on the availability, a device will be shipped.' At the bottom are 'Next' and 'Back' buttons.

## Notification Details

Now we provide the email address(es) that we would like notified about our order status

The screenshot shows the 'Notification details' step of an order preview. The breadcrumb trail is: Home > New > Marketplace > Everything > Azure Data Box (preview) > Select your Azure Data Box > Order. The page title is 'Order PREVIEW'. The navigation tabs are 'Order details', 'Shipping address', 'Notification details' (selected), and 'Summary'. The form contains the following text and field:

We will keep you updated regarding your order progress. Specify the email address(es) to receive updates regarding your order status.

\* Enter up to 10 email addresses separated by semicolons

john@contoso.com ✓

At the bottom are 'Next' and 'Back' buttons.

## Summary

Lastly, we're provided a summary for our request before we click on 'Order'

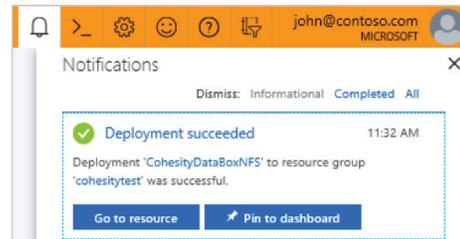
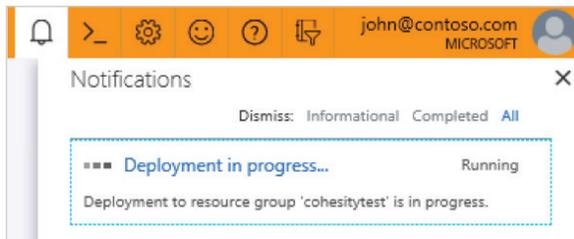
The screenshot shows the 'Order' summary page in the Azure Marketplace. The page is titled 'Order PREVIEW' and has tabs for 'Order details', 'Shipping address', 'Notification details', and 'Summary'. The 'Summary' tab is active. The page is divided into several sections:

- SELECTED MODEL:** Model: Data Box; Capacity (Usable/Total): 80 TB / 100 TB; Service fee: USD; Shipping fee: USD; Estimated total fee\*: USD.
- Pricing details:** \* Extra charges are incurred if the device is retained beyond 10 days.
- ORDER DETAILS:** Order name: CohesityDataBoxNFS; Subscription: Databox; Resource group: cohesitytest; Destination Azure region: West US; Storage account(s): cohesityaccount.
- SHIPPING DETAILS:** Contact person: John; Company name: Cohesity; Address: San Jose, 95110, CA, United States of America; Work phone: +1855-XXXX-XXXX.
- NOTIFICATION DETAILS:** Email(s): john@contoso.com.
- SHARE DETAILS:** cohesityaccount: \\<ipAddress>\cohesityaccount\_BlockBlob, \\<ipAddress>\cohesityaccount\_PageBlob, \\<ipAddress>\cohesityaccount\_AzFile.

At the bottom, there is a checkbox for 'Provided information is correct. I agree to the privacy terms.' and two buttons: 'Order' and 'Back'.

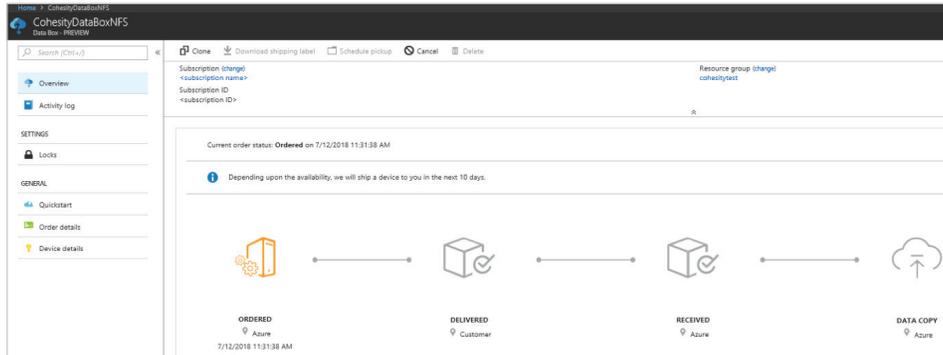
## Data Box Deployment

Once we've ordered our Data Box, we'll see that our order is being processed and being deployed to the Azure Resource Group we selected:

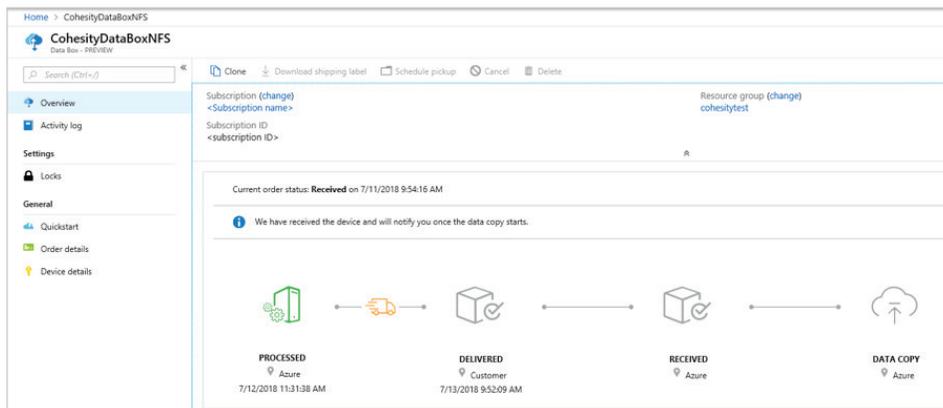


## Shipping

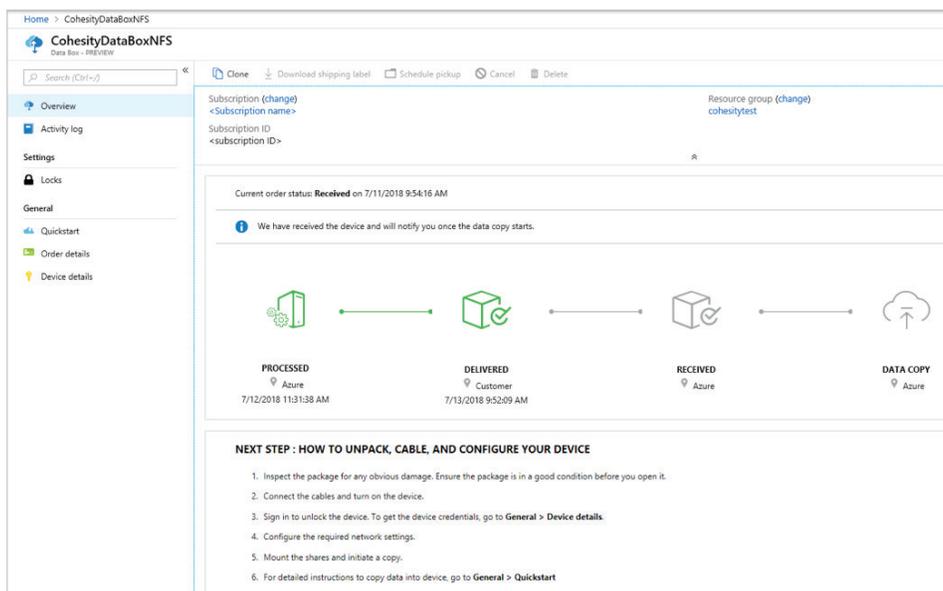
Now that our order has been processed, we can see the status of our order. Please note that it may take a few days to ship an Azure Data Box to your data center, so please plan accordingly.



Once the Data Box has been ordered, you will see that status shows the device in transit to your data center.



When the Data Box has been received at your location, the status will be updated to 'Delivered'.



## Data Box Installation / Setup

In addition to being marked as 'Delivered' in the Azure portal you are also provided information on how to unpack and configure your device.

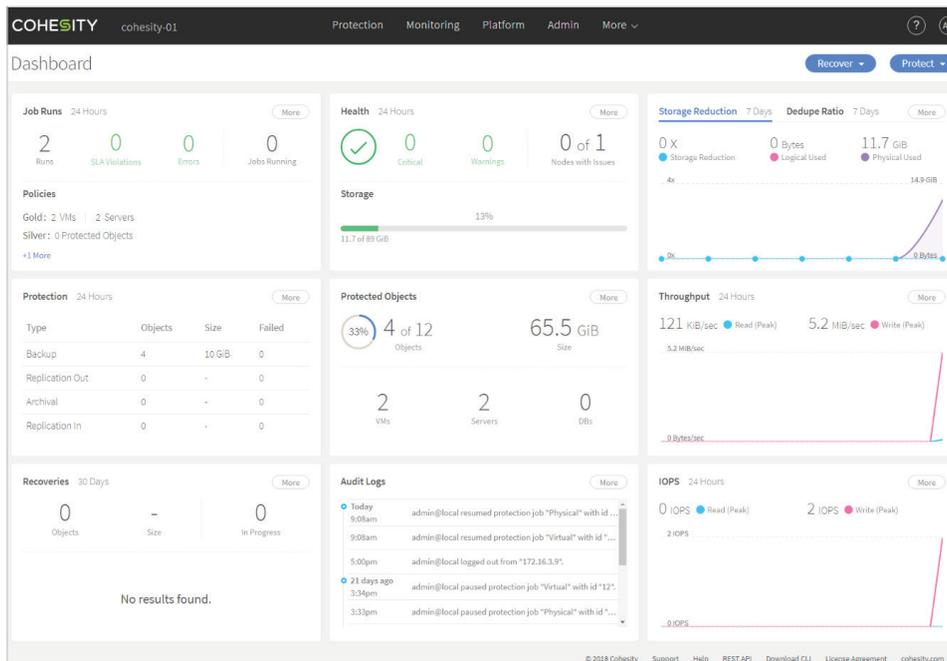
The device is shipped in a rugged case, but please do inspect everything for any evidence of damage or tampering. After confirming the Data Box can be installed, please connect follow the information provided in the Azure portal, and reference the Quickstart for detailed instructions.

## Cohesity DataPlatform

Once the Azure Data Box has been installed, configured, and you can confirm the Data Box is visible on the network (i.e 'ping'), we can configure settings on the Cohesity DataPlatform.

### General

For purposes of this integration testing we deployed a Cohesity Virtual Edition (VE). Since Cohesity is software defined, it can run on a Virtual Machine, on native public cloud resources, or as a physical appliance. Accessing the Cohesity DataPlatform cluster we will see something like this in our dashboard:



## Configure External Targets

For this step we want to configure two external targets for our Archive. One for our Azure Data Box (DataBoxBlock), and one for our Azure Storage Account and Container (AzureLTR).

DataBoxBlock <a href="#">Back to External Targets</a>	
Type	NAS
Purpose	Archival
NAS Host	10.126.76.36
Mount Path	/cohesitydatabox_BlockBlob/nfsarchive
Share Type	NFS
Encryption	Enabled
Additional security by managing key manually	Disabled
Compression	Enabled
Source Side Deduplication	Enabled
Incremental Archival	Enabled

AzureLTR <a href="#">Back to External Targets</a>	
Type	Azure Standard
Description	Azure
Purpose	Archival
Container Name	archive
Storage Account Name	azurestandardarchive
Encryption	Enabled
Additional security by managing key manually	Disabled
Compression	Enabled
Source Side Deduplication	Enabled
Incremental Archival	Enabled

## Configure Protection Policy

First and foremost, Cohesity is focused on protecting your data. The protection policies provide granular levels of control for backups and archive schedules, in addition to replication to alternate sites, or replication of on-premises VMs into public cloud VMs. You can find out more about Cohesity's data protection and secondary storage capabilities at [www.cohesity.com](http://www.cohesity.com).

While it's very easy to edit an existing protection policy to add an archive destination, we've created a new protection policy with an Archive to our Azure Data Box (DataBoxBlock) external target.

### Edit Policy: DataboxTestPolicy [Description](#)

Backup

**Schedule**

Backup every **6** hour(s) **Retain for 90** day(s)

Incremental only **▼**

[+ Add Log Backup](#) (MS SQL)

**BMR Backup** (Physical Server)

Backup every week **▼** On **Monday** **▼** Retain for **90** day(s) **✕**

Extended Retention [+ Add](#)

Retry Options [Edit](#)

Blackout Window [+ Add](#)

Replication

[+ Add Replication](#)

Archival

Archive to: **DataBoxBlock** **▼**

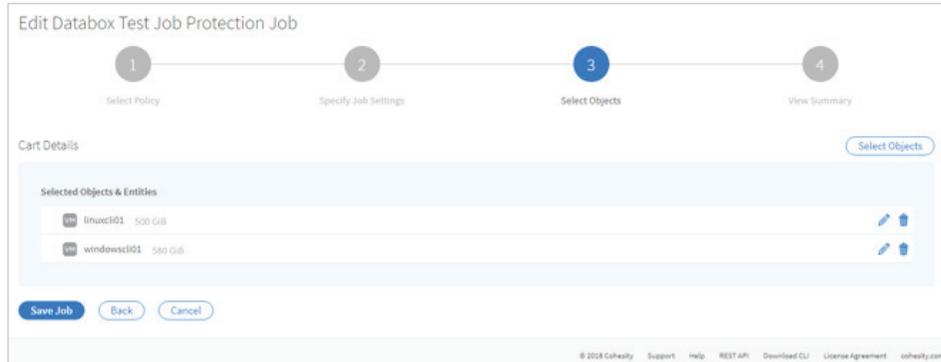
After every run **▼** Retain for **365** day(s) **✕**

[+ Add Archival](#)

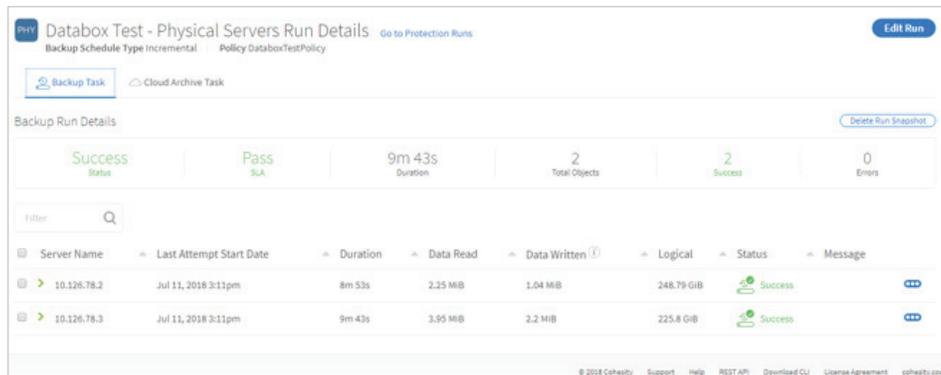
**Save** **Cancel**

## Create Protection Job

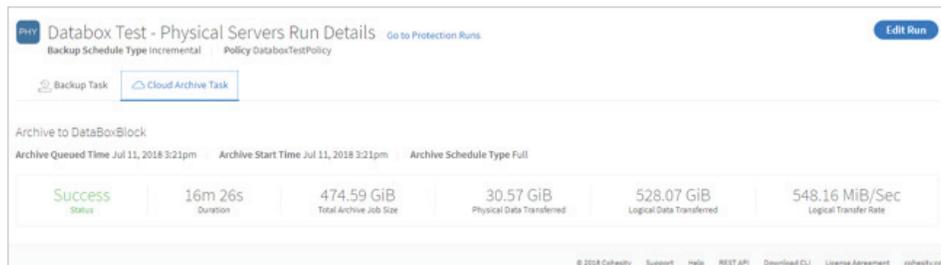
The next step in the process is to identify what to protect. In this case, we're interested in protecting some VMs from our VMware environment that we've already discovered as a data Source. In this scenario, based on our protection policy, we will be backing up a Linux and Windows VM to our Cohesity cluster and archiving them to our Azure Data Box.



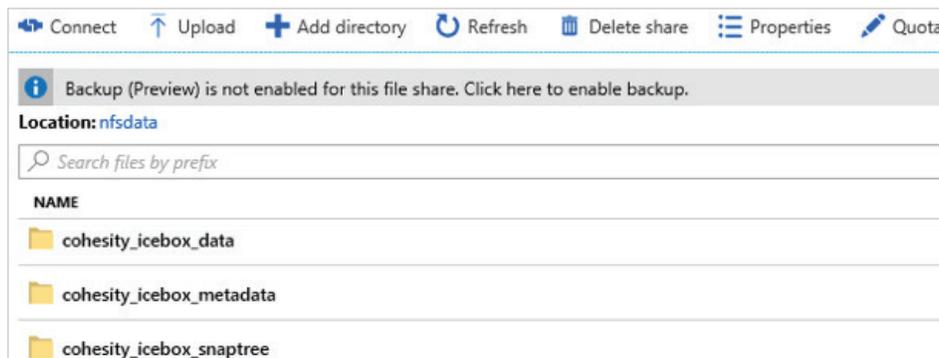
Looking at Job Details while this protection job is running will show us something like this where there is a breakdown between protection processes. In this view we can see the data that's been protected on Cohesity via the backup task.



Taking a look at the Cloud Archive Task we can see the status of our Archive to Azure Data Box:



Accessing the Data Box, we can see the directories placed onto the Data Box for the Cohesity Archive:



The data in a Cohesity Archive contains the data, metadata, or SnapFS SnapTree information. This way, it's possible to launch a new Cohesity system anywhere... point to the External Target... and recover our data and operations immediately, if needed. With our scenario, we've successfully archived our data onto the Azure Data Box.

## Return Data Box to Microsoft Azure

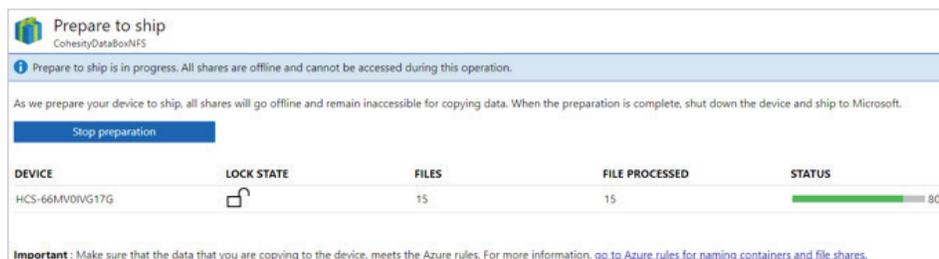
Now that we've seeded our Data Box with our Archival data, we can return the databox to Azure.

Depending on your preference, you can opt to do one of two things on the Cohesity side of things:

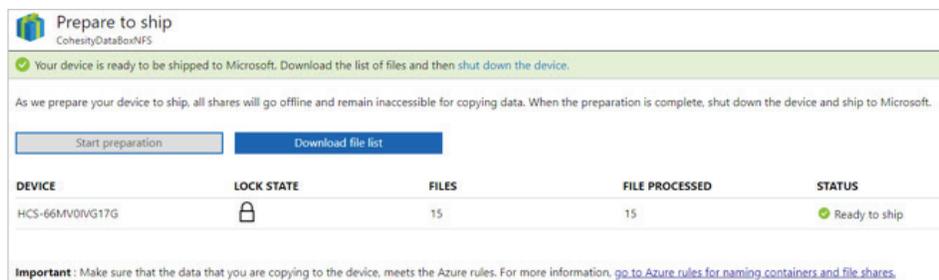
1. Leave the protection policy and protection job as-is while the Data Box is being returned to Azure
  - a. Note that this will result in a failure of the Archival portion of the protection job (via the protection policy)
2. Edit the protection policy to remove the archival component of the protection policy
  - a. As you will no longer have an archival component, you will not receive any archive errors, but you will need to edit the policy to add in an archival component once more when the Data Box data has been uploaded into your Microsoft Azure subscription.

### Prepare to Ship

When ready to ship the Azure Data Box back to Azure, you simply need to access the IP address for your Azure Data Box, and select 'Prepare to ship'.



Once the Data Box is prepped, it will change to a locked state, and you can also download the file list (recommended)

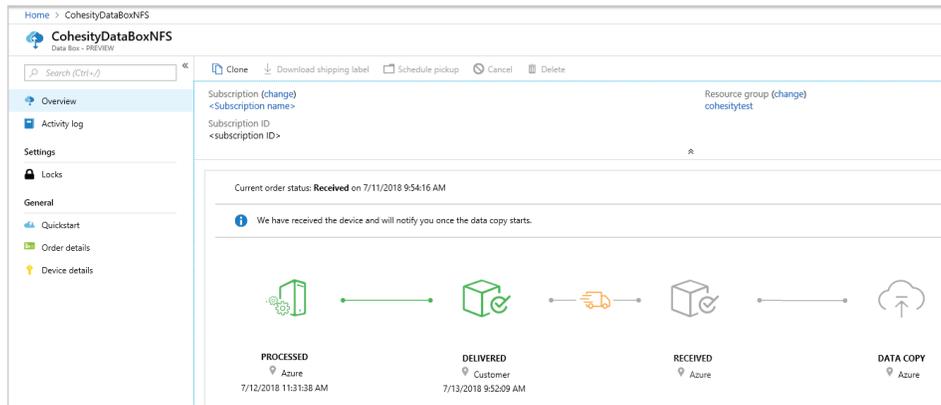


When ready, you can select to 'shut down the device' in the green banner



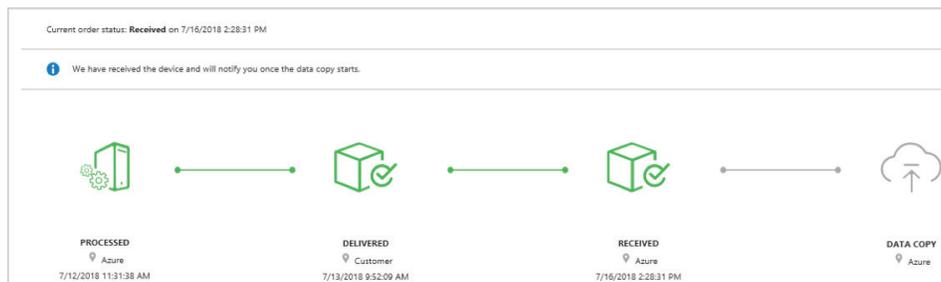
## Physically Prepare & Ship

Once you've completed the logical shutdown of the Data Box, please physically power off the system, disconnect the Data Box from your network and power, and then place into the rugged case it which it was originally shipped. Complete preparation and Data Box shipping instructions can be found [HERE](#).

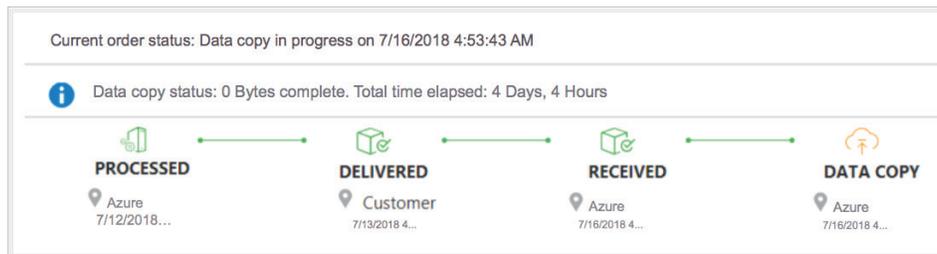


## Discover

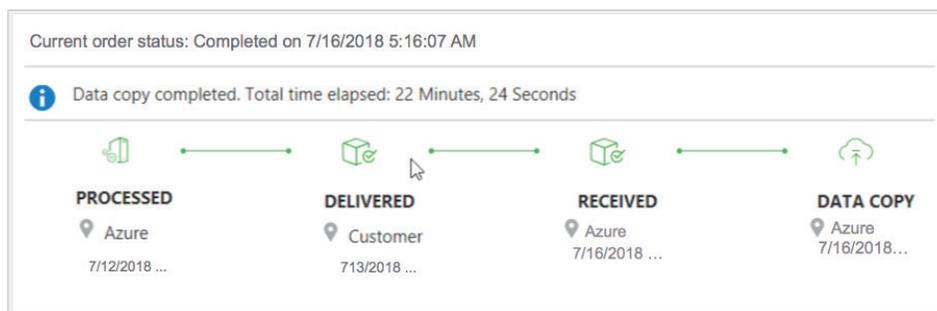
Upon shipping the Azure Data Box back to Azure your status via the Azure Portal for your Data Box will be updated accordingly. Once received, you will see a status like this:



Upon receiving the databox, our data is automatically uploaded into the Azure Storage Account we've originally associated.



Finally when our data has been fully uploaded into Azure, the status will look something like this:



## Cohesity – Resume Protection – After Import of data into Azure

Once the data has been uploaded into the Azure Storage account we can resume the archive of our data, but now picking up where we left off with the data from the Data Box. This is where we will now use the 'AzureLTR' External Target we configured earlier.

### Redirect External Targets

In order to leverage the data that is now in our Microsoft Azure account we need to redirect the original External Target (DataBoxBlock) to the new External Target (AzureLTR). To do this, we need to contact Cohesity to help us perform the redirect. More than likely you will already be collaborating with a person from Cohesity. If not, it's possible to contact Cohesity Support at <https://support.cohesity.com/>, or by calling 1-855-9COHESITY.

### Resume Archive

After the External Target redirect we can simply resume archiving our data. The Archive portion of the Protection job will identify all of the changes since the last Archival, and only send the changed blocks to the External Target (AzureLTR):

VM Databox Test Job Run Details [Go to Protection Runs](#) [Edit Run](#)

Backup Schedule Type Incremental | Policy DataboxTestPolicy

Backup Task | Cloud Archive Task

Archive to AzureLTR

Archive Queued Time July 17, 2018 8:50pm | Archive Start Time July 17, 2018 8:50pm | Archive Schedule Type Incremental

Success Status	1m 32s Duration	1.05 TiB Total Archive Job Size	200 MiB Physical Data Transferred	720.01 GiB Logical Data Transferred	7.8 GiB/Sec Logical Transfer Rate
----------------	-----------------	---------------------------------	-----------------------------------	-------------------------------------	-----------------------------------

© 2018 Cohesity | Support | Help | REST API | Download CLI | License Agreement | cohesity.com

## Restore a VM

To round out our testing, we've also restored a VM from our Azure Archive. Cohesity provides you the granular level of control to choose not only the time-stamp of the data you want to restore, but also the location you would like to restore from. In this scenario we're going to request to recover our data from our AzureLTR external target.

6 Recover Points for linuxcli01

Recover Point	Snapshot Type	Backup Type	Stored
<input type="radio"/> July 16,2018 2:45pm	Crash-Consistent	Incremental Backup	
<input type="radio"/> July 15,2018 2:45pm	Crash-Consistent	Incremental Backup	
<input type="radio"/> July 14,2018 2:45pm	Crash-Consistent	Incremental Backup	
<input type="radio"/> July 13,2018 2:45pm	Crash-Consistent	Incremental Backup	
<input checked="" type="radio"/> July 12,2018 2:45pm	Crash-Consistent	Incremental Backup	
<input type="radio"/> July 11,2018 2:45pm	Crash-Consistent	Full Backup	

Saving off our recovery point and running our recovery job will restore our linuxcli01 VM from our Archive up in Azure

The screenshot shows the Cohesity Recovery dashboard. At the top, there are navigation tabs: Protection, Monitoring, Platform, Admin, and More. Below the navigation, there are statistics: 0 Running Tasks, 1 Total Task, 1 Success, 0 Errors, and 0 Scheduled. The dashboard is divided into Local Recoveries, Cloud Recoveries, and Tape Recoveries. A date range of 7 days from July 11, 2018 to July 17, 2018 is selected. A table lists recovery tasks with columns for Task Name, Start Time, Duration, and Status. One task is listed: VM Recover-VMs\_July\_16\_2018\_5:28pm Objects 1, with a start time of July 16, 2018 5:33pm, a duration of 3m 52s, and a status of Success. The footer contains copyright information: © 2018 Cohesity, Support, Help, REST API, Download CLI, License Agreement, cohesity.com.

## Summary

Cohesity and Microsoft partner together to offer solutions and services that meet your ever-growing and ever-evolving data needs. Solutions like the integration between the Cohesity DataPlatform and Azure Data Box for protecting and moving massive amounts of data safely and securely is only one example of this partnership.

If you're interested in learning more about the Cohesity and Microsoft Azure partnership and integration, please come check out [Cohesity's Native Cloud Integration with Azure](#).

## About Cohesity

Cohesity delivers web-scale simplicity for secondary data that eliminates silos and puts data to work. Enterprises rely on our software-defined, hyperconverged platform to radically streamline backup and data protection, converge file and object services, quickly deliver test/dev instances, and provide analytic functions on a single, global data store. Customers and partners, including Global 1000 companies and federal agencies, tired of data center complexity are modernizing and scaling secondary data protection with Cohesity.

Find out more about Cohesity [HERE](#).

## About Azure Data Box

Transferring terabytes of data to the cloud is hard. Azure Data Box provides a secure, tamper-resistant method for quick and simple transfer of your data to Azure, offering 100TB capacity. You can order the Data Box through the Azure portal. Easily connect it to your existing network, then load your data onto Data Box using standard NAS protocols. Your data is automatically protected using 256-bit AES encryption. Data Box is returned to the Azure Data Center to be uploaded into Azure, then the device is securely erased.

Find out more about Azure Data Box [HERE](#).