COHESITY

Version 1.0

August 2018

Cohesity Microsoft Azure Data Box Integration

Table of Contents

Tables

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 that after Cohesity has applied its storage efficiencies, there 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 Platform 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

Table 1: Requirements

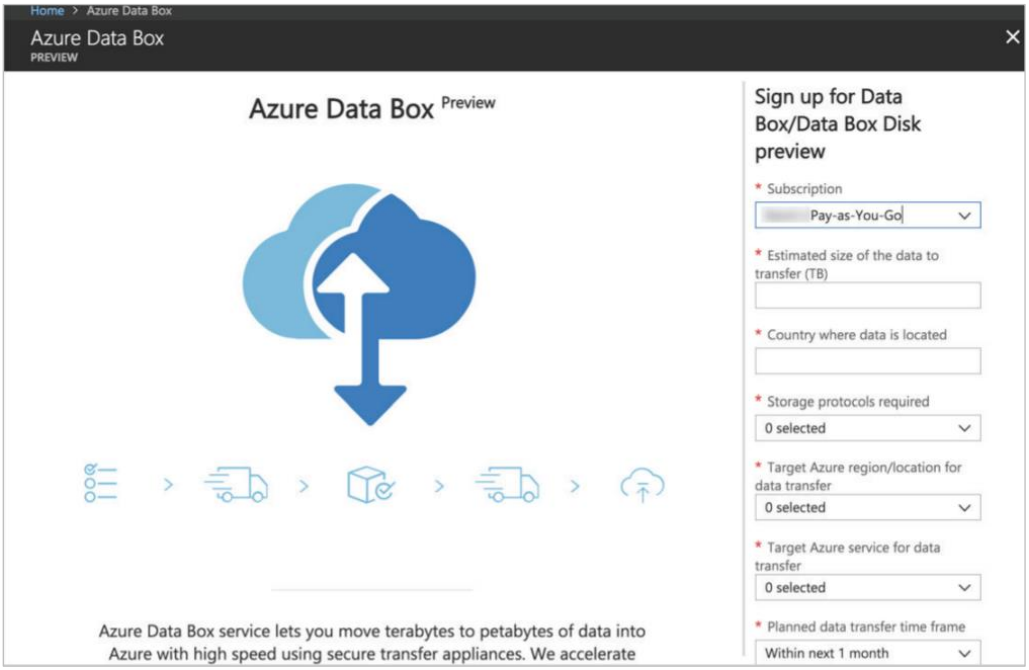
ITEM	VALUE	PROVIDER
Microsoft Azure Subscription	Non-Free Tier	Microsoft
Microsoft Azure Data Box	4.x	Microsoft
Azure Blob Container Types	Hot, Cool	Microsoft
Azure Data Box Connectivity	NFS	n/a
Cohesity Platform	5.0.1+	Cohesity

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 Platform deployed and running either as a Virtual Edition, a Cloud Edition (i.e. running on Azure resources), or as a physical appliance.

Order Microsoft Azure Data Box

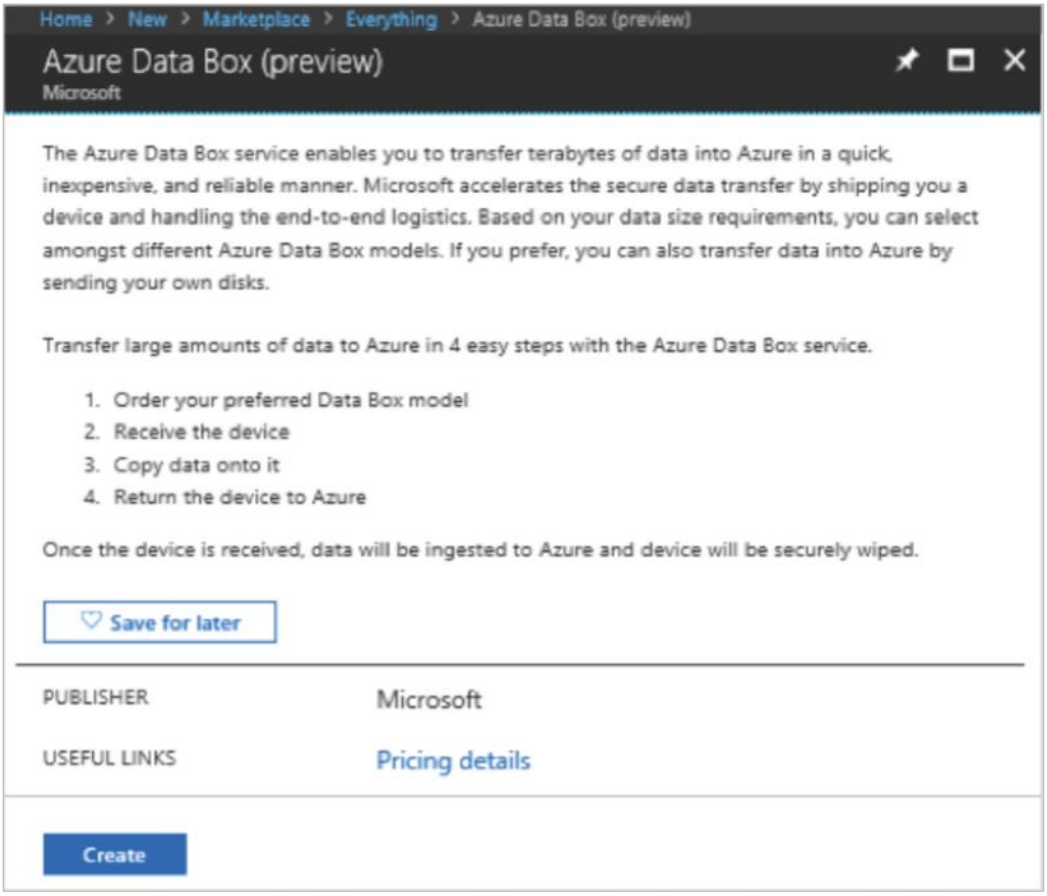
One of the first things you need to do is to request and set up an Azure Data Box. At the time of initial publication of this document, you'll first need to go to the Azure Data Box or Microsoft Azure Site, and sign up for the Preview. Clicking on the **Sign Up for the Preview >** button takes you to where you 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. See Azure documentation for the complete walk-through guide for ordering an Azure Data Box.

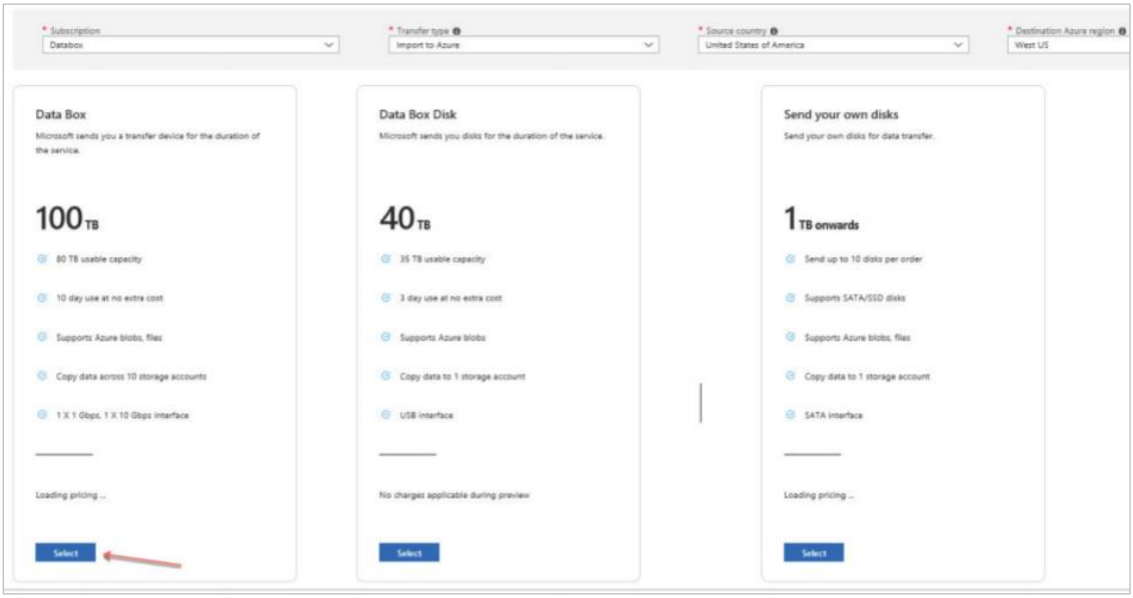
Requesting

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**, select the type of Azure Data Box you want to use in addition to the following Azure details:

- The **Subscription** you would like to use.
- The **Transfer type** (in this scenario: 'Import to Azure').
- Your **Source country** (in this scenario: 'United States of America').
- **Destination Azure region** (in this scenario: 'West US').

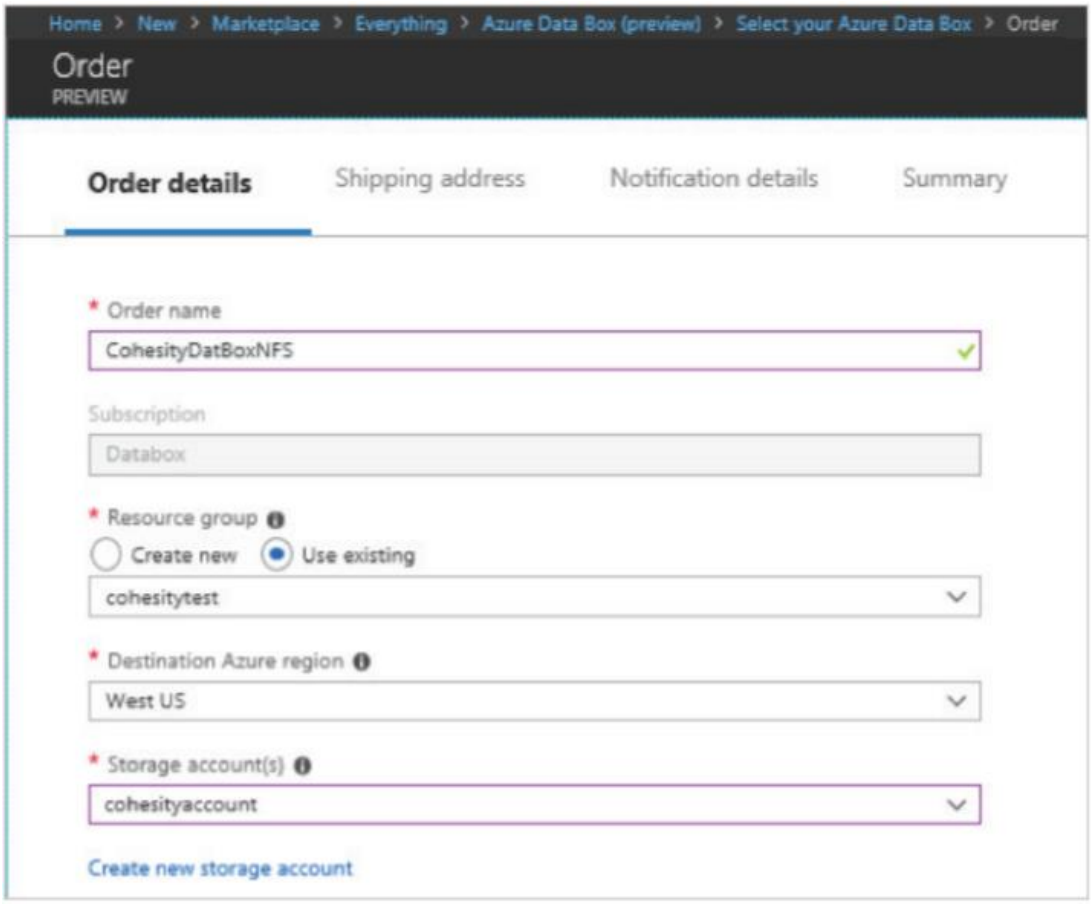


For testing purposes, 100TB Data Box (formatted: 80TB usable) is selected by default.

Order Details

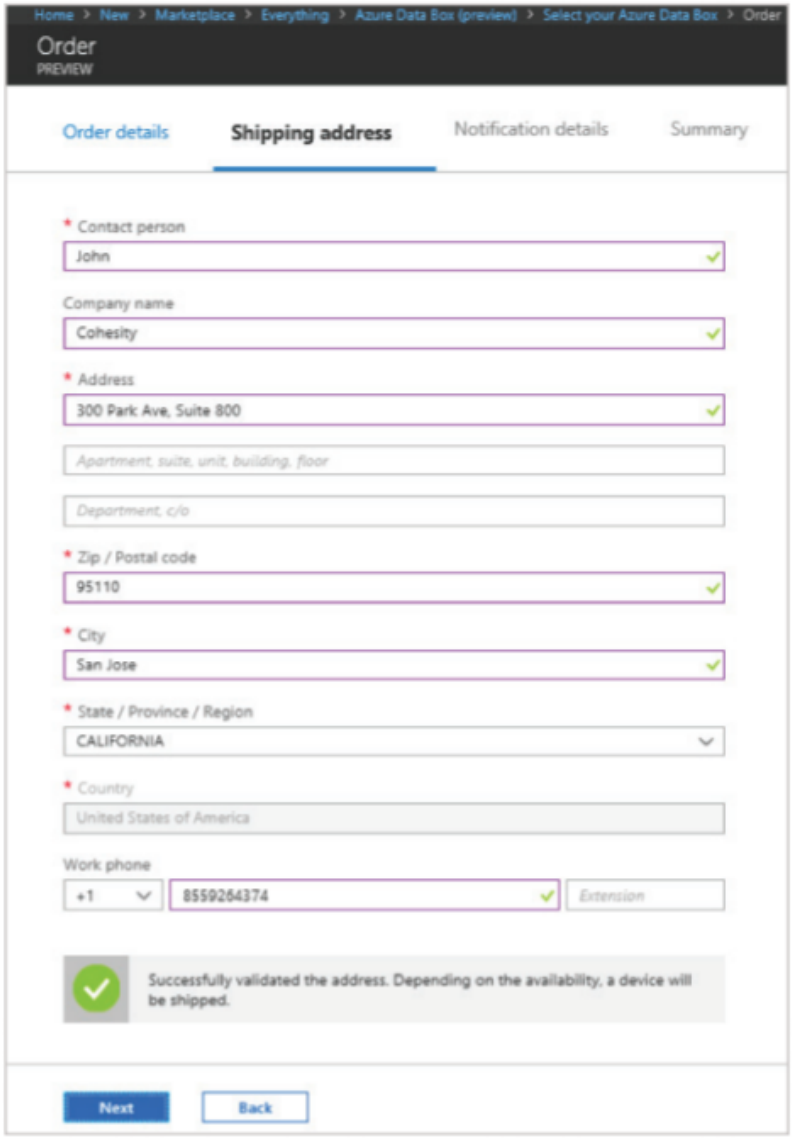
Provide information about your order:

- An **Order name**.
- The Azure **Resource group** you want to use (or if you want to create a new Resource Group).
- The **Destination Azure region**.
- The Azure **Storage account(s)** you want to use (or if you want to create a new Storage account).



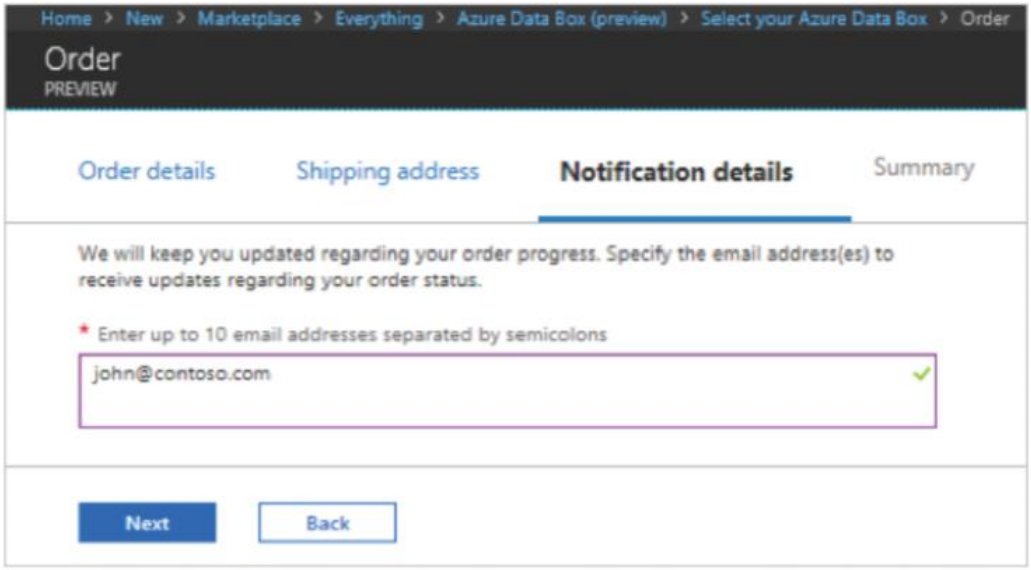
Shipping Address

Next, you need to specify the shipping information for your Data Box.



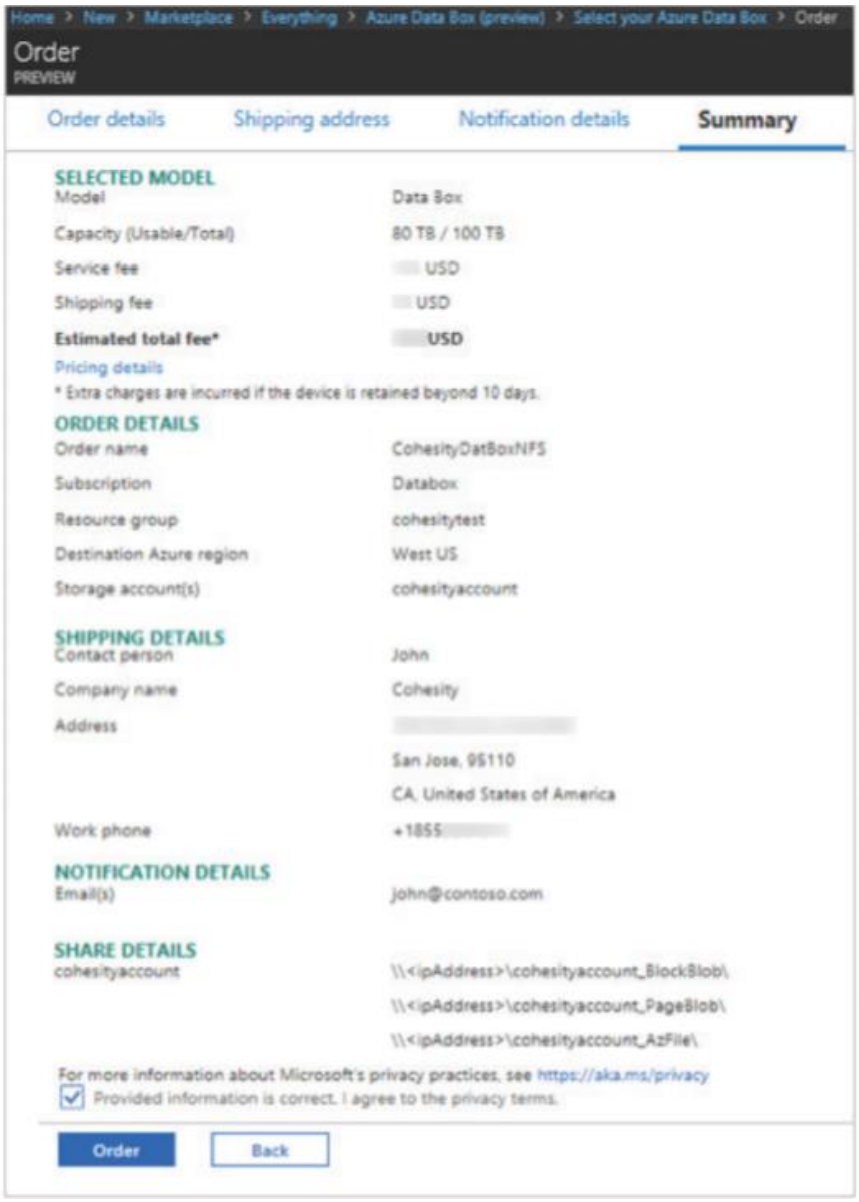
Notification Details

Now you provide the email address(es) to get notified about your order status.



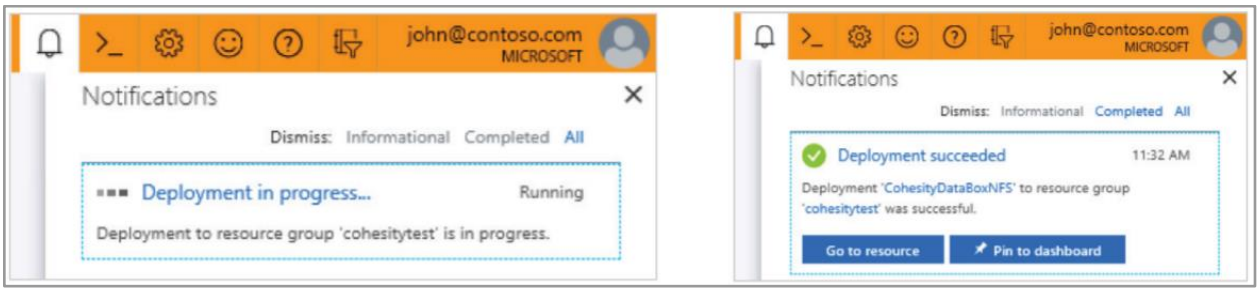
Summary

In the Order **Summary**, click **Order**.



Data Box Deployment

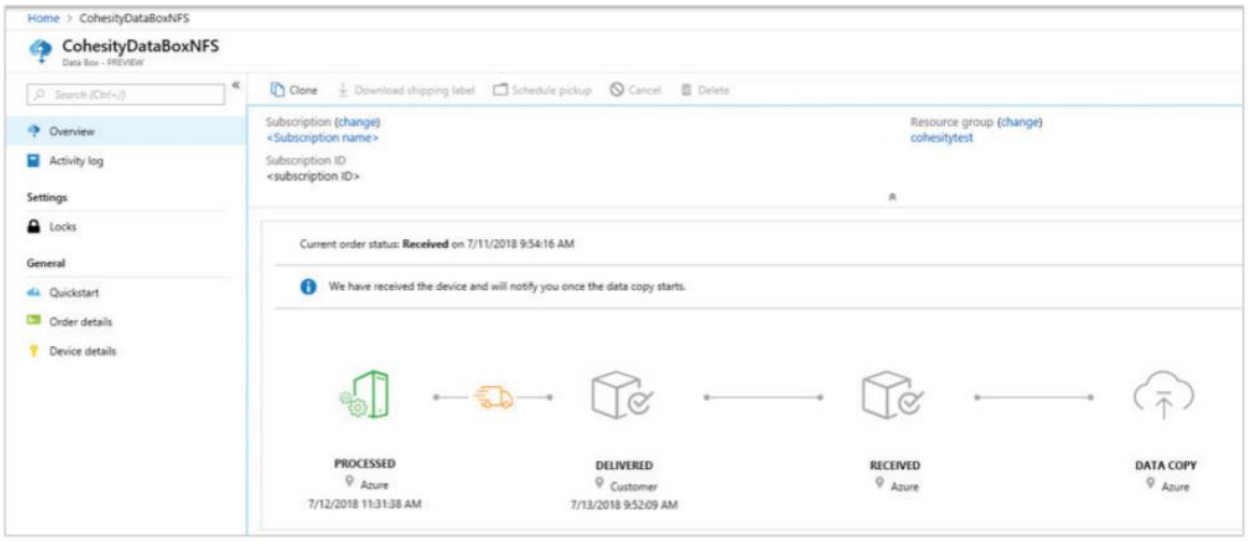
Once you've ordered your Data Box, you'll see that your order is being processed and being deployed to the Azure Resource Group you selected:



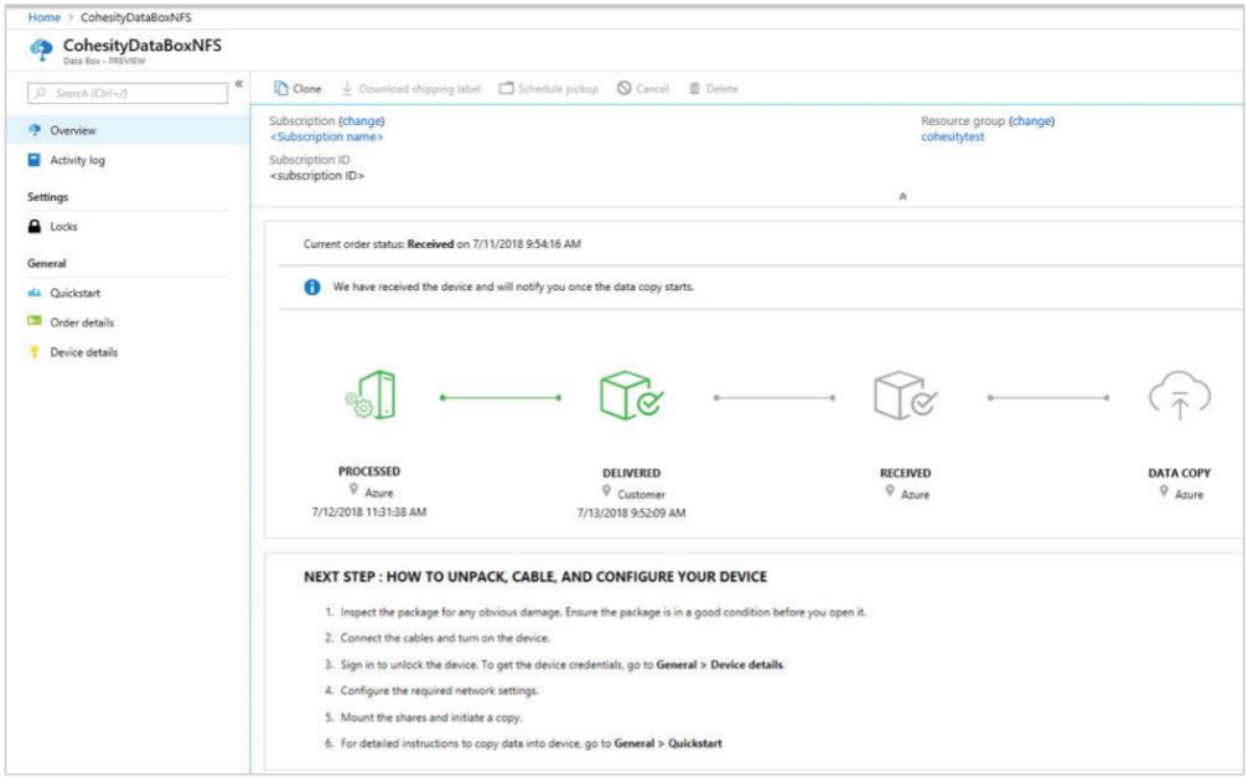
Shipping

Now that your order has been processed, you can see the status of your 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

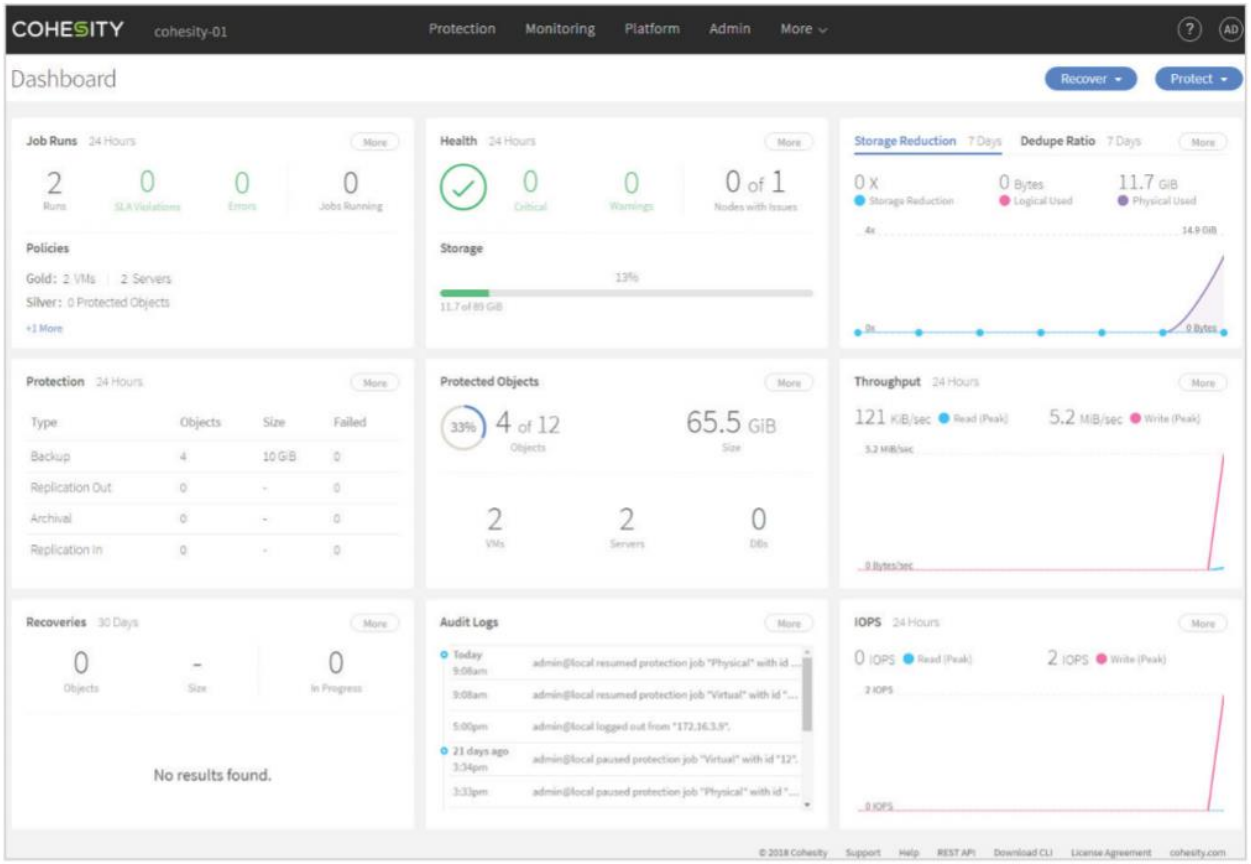
In addition to being marked as **DELIVERED** in the Azure portal you are also provided information on how to unpack and configure your device. The device is shipped in a rugged case, but please do inspect everything for any evidence of damage or tampering. After confirming the Data Box can be installed, please connect, follow the information provided in the Azure portal, and see Quickstart for detailed instructions.

Cohesity Platform

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

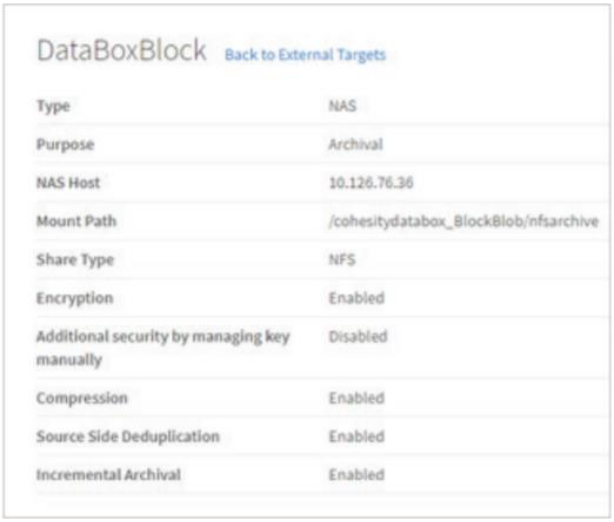
General

For integration testing purposes, Cohesity 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 Platform cluster, the following dashboard appears:



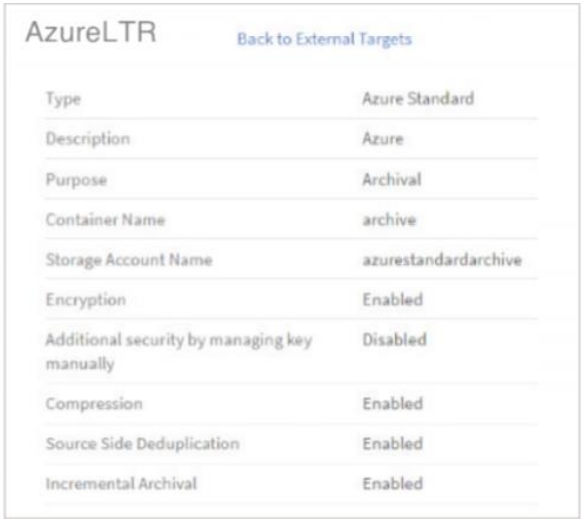
Configure External Targets

In this step, you want to configure two external targets for your Archive. One for your Azure Data Box (DataBoxBlock), and another for your Azure Storage Account and Container (AzureLTR).



DataBoxBlock Back to External Targets

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



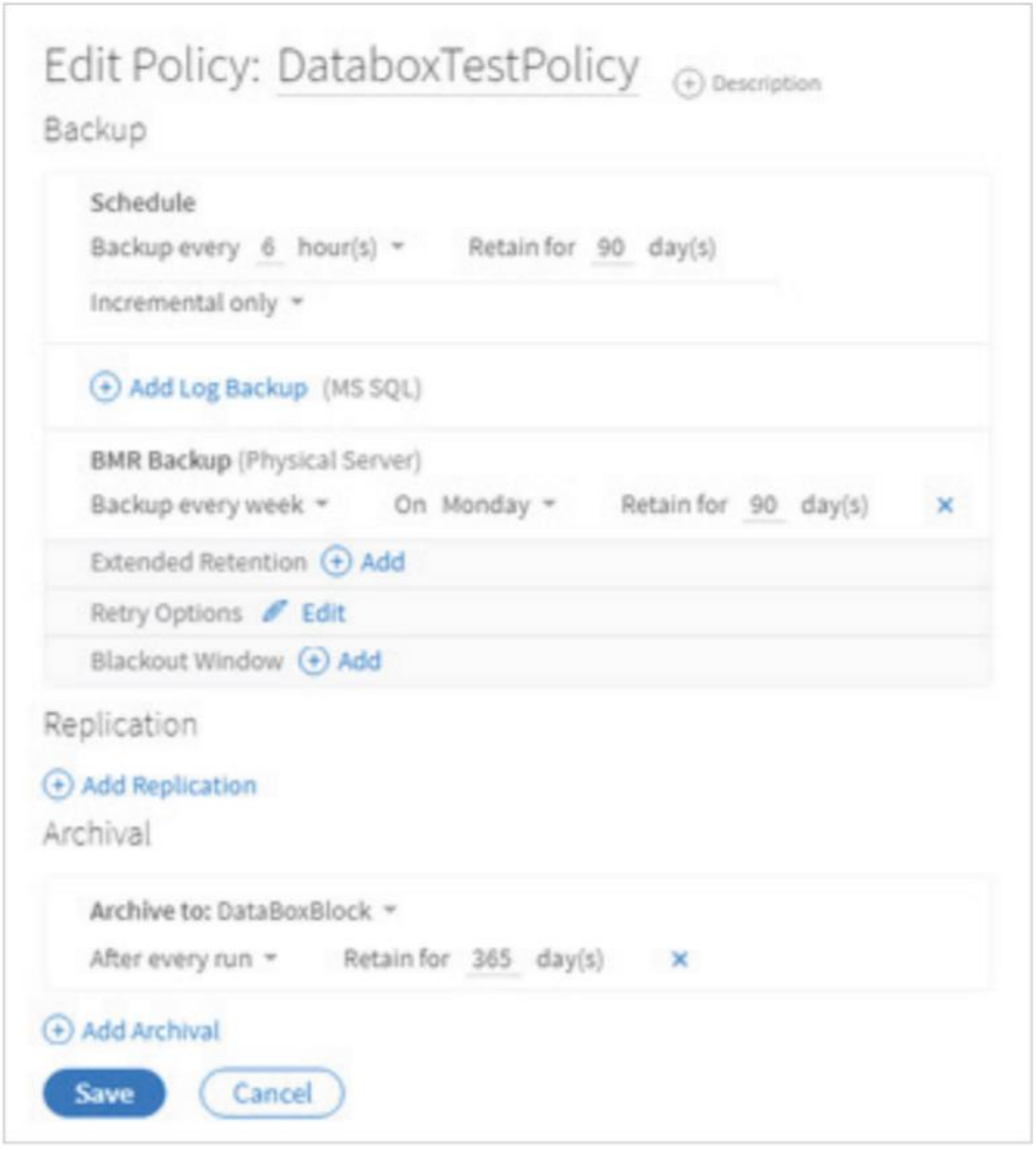
AzureLTR Back to External Targets

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

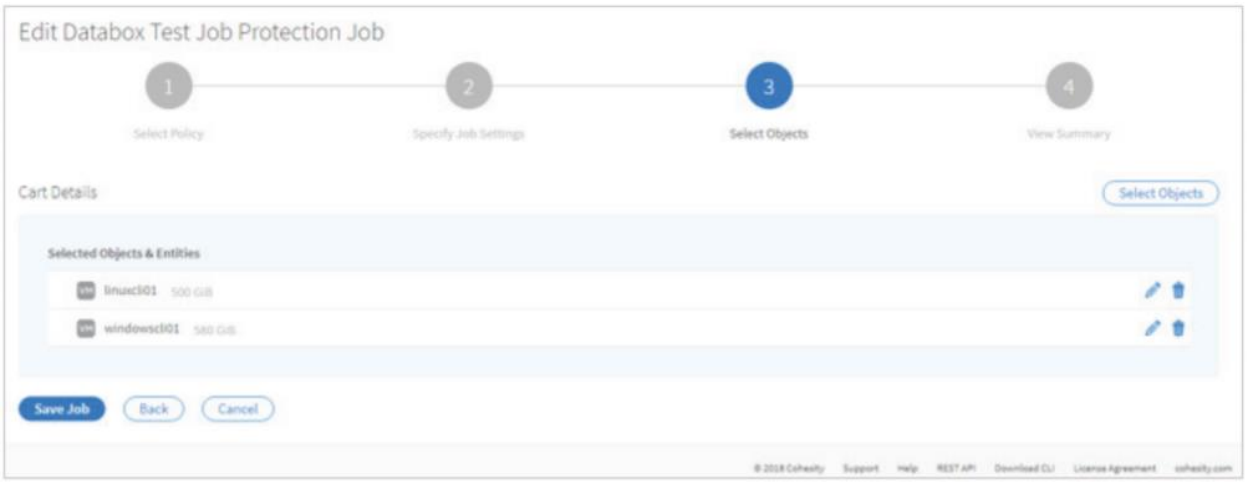
First and foremost, Cohesity is focused on protecting your data. The protection policies provide granular levels of control for backups and archive schedules, in addition to replication to alternate sites, or replication of on-premises VMs into public cloud VMs. You can find out more about Cohesity's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

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

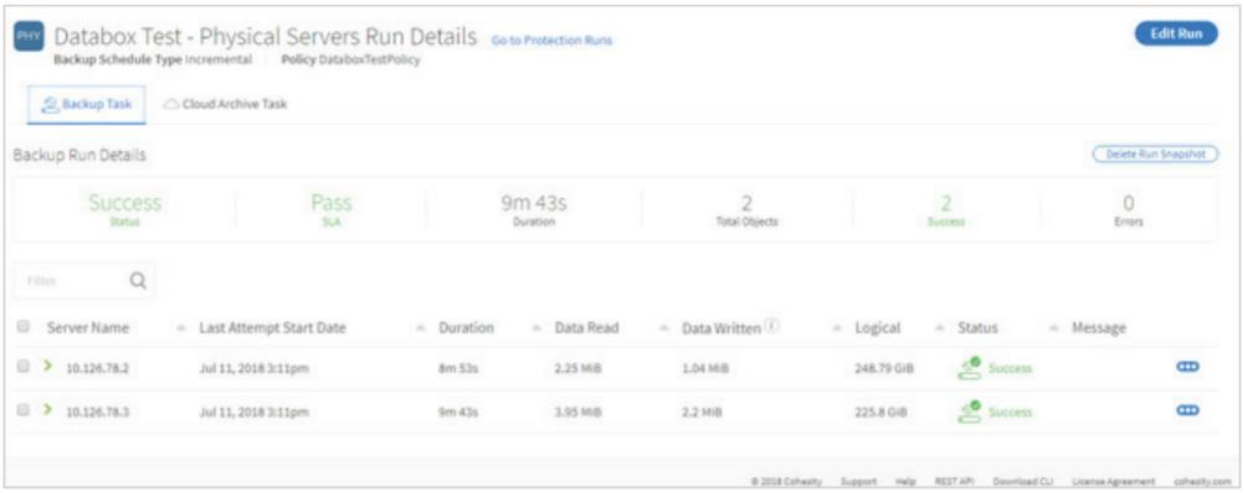


Create Protection Job

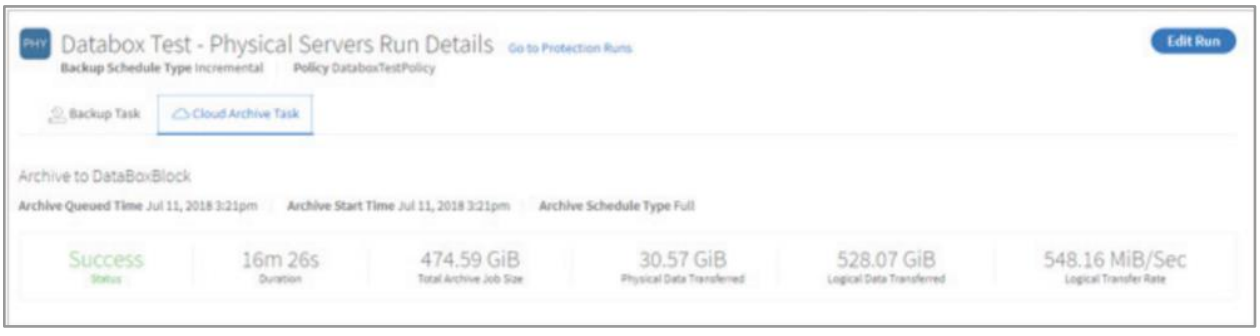
The next step in the process is to identify what to protect. In this case, Cohesity is interested in protecting some VMs from your VMware environment that Cohesity has already discovered as a data Source. In this scenario, based on our protection policy, Cohesity will be backing up a Linux and Windows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



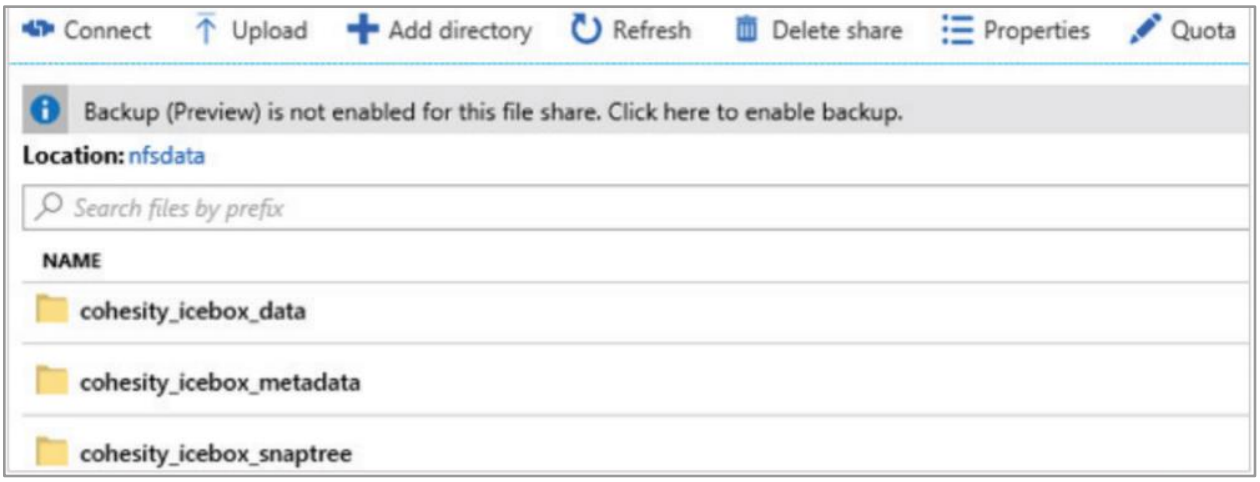
Looking at Job Details while this protection job is running will show you something like this where there is a breakdown between protection processes. In this view you can see the data that's been protected on Cohesity via the backup task.



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



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



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

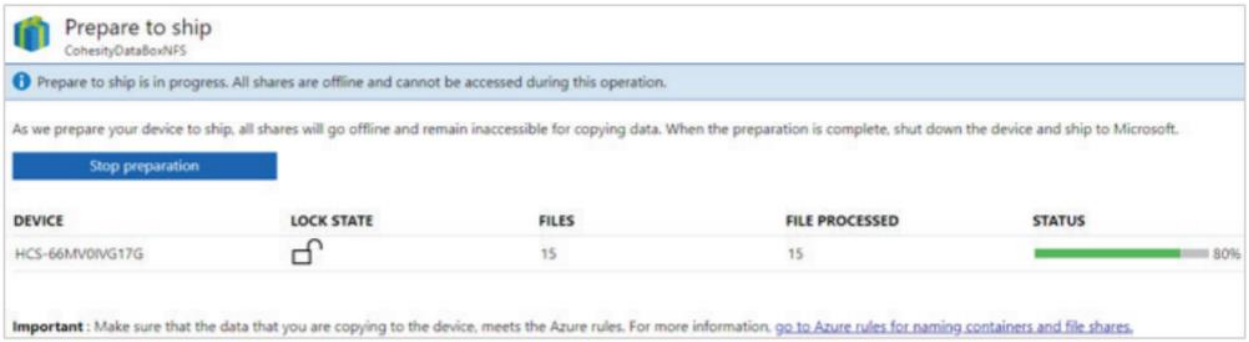
Return Data Box to Microsoft Azure

Now that you've seeded Data Box with your Archival data, you 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.
   b) 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.

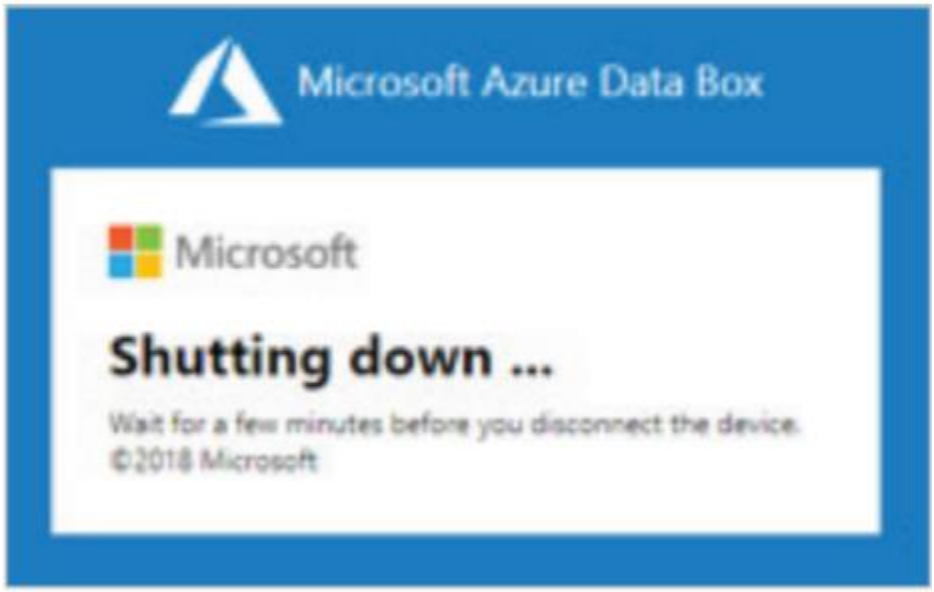
Prepare to Ship

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**.



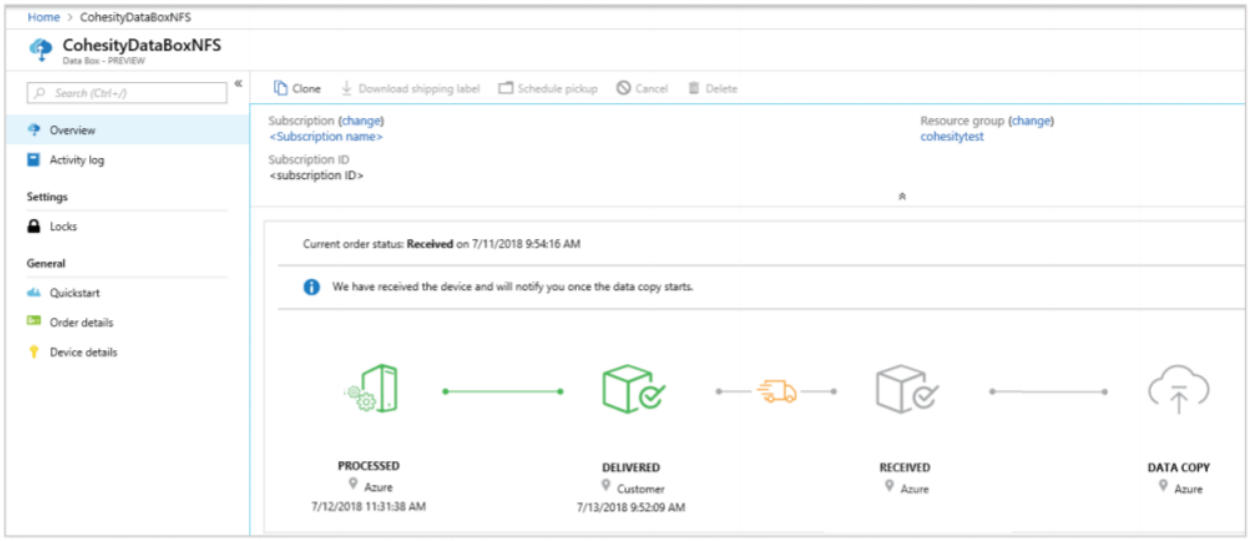
Once the Data Box is prepared, it will change to a locked state, and you can also download the file list (recommended).

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



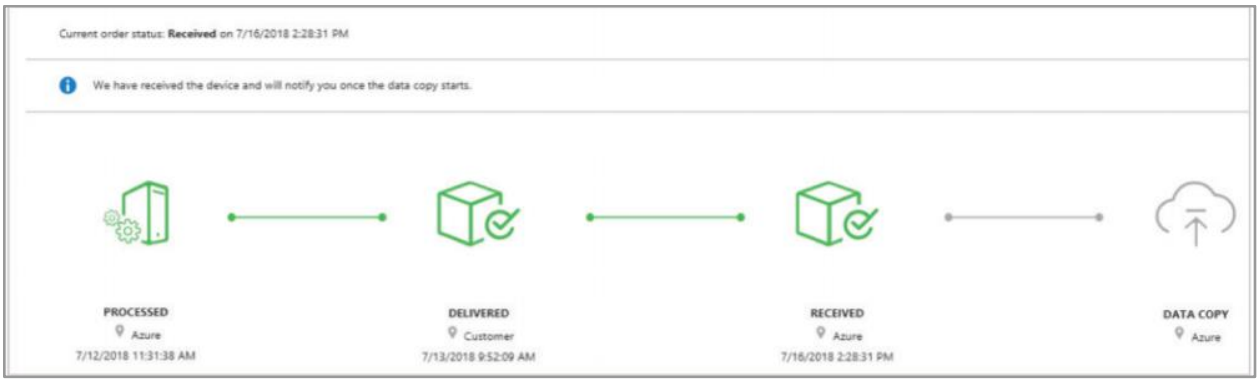
Physically Prepare and Ship

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 was originally shipped. Complete preparation and Data Box shipping instructions can be found here.

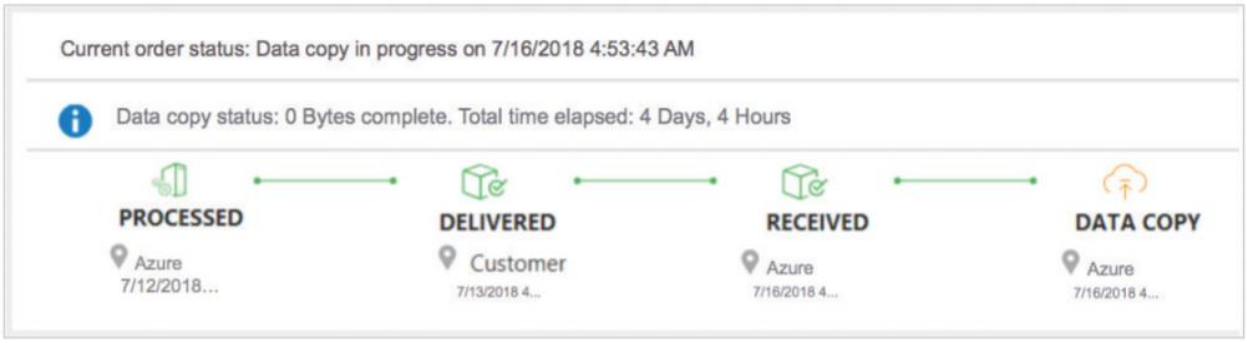


Discover

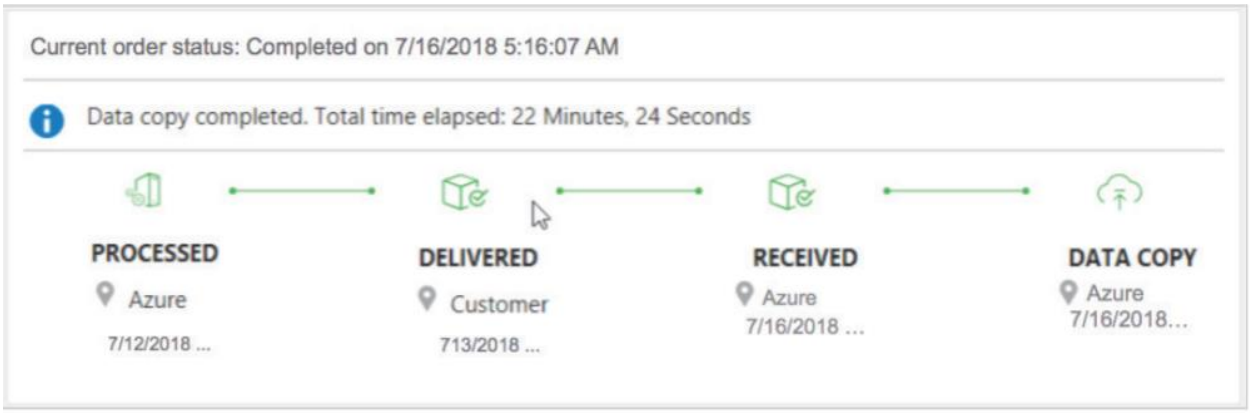
Upon shipping the Azure Data Box back to Azure, your status via the Azure Portal for your Data Box will be updated accordingly. Once received, you will see a status like this:



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



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



Cohesity—Resume Protection—After Import of Data into Azure

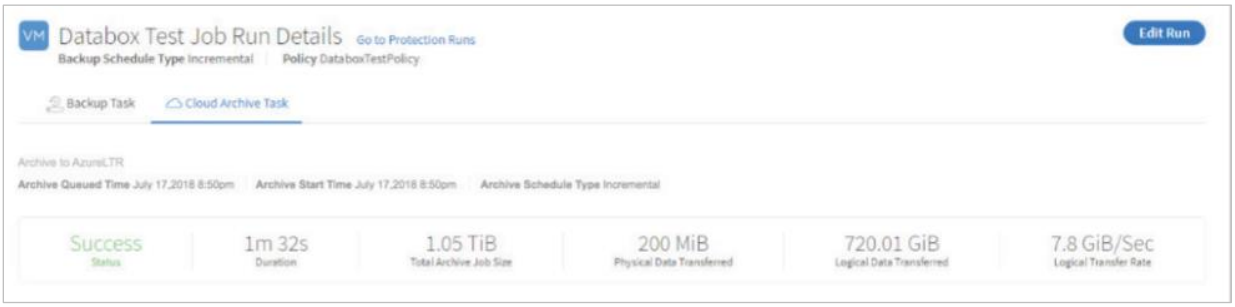
Once the data has been uploaded into the Azure Storage account we can resume the archive of our data, but now picking up where we left off with the data from the Data Box. This is where we will now use the 'AzureLTR' External Target we configured earlier.

Redirect External Targets

In order to leverage the data that is now in our Microsoft Azure account you need to redirect the original External Target (DataBoxBlock) to the new External Target (AzureLTR). To do this, you need to contact Cohesity to perform the redirect. More than likely you will already be collaborating with a person from Cohesity. If not, you can contact Cohesity Support at https://support.cohesity.com/.

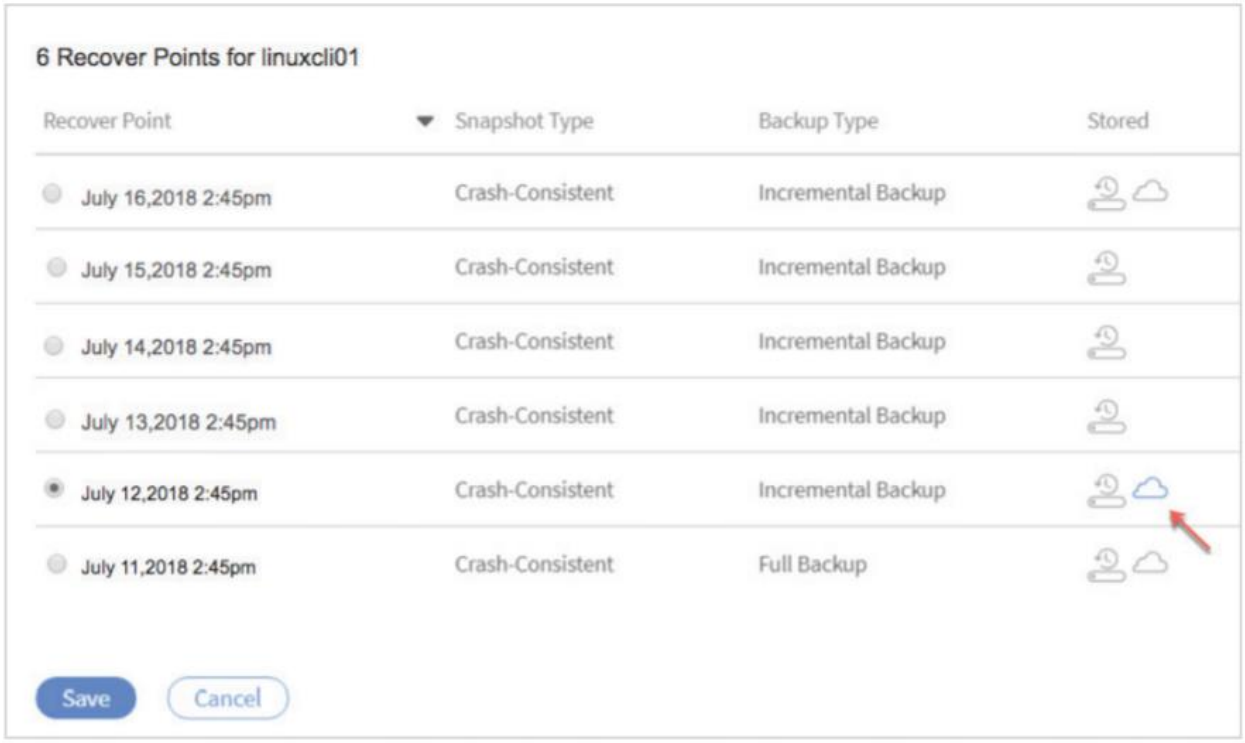
Resume Archive

After the External Target redirect, you can simply resume archiving your 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):

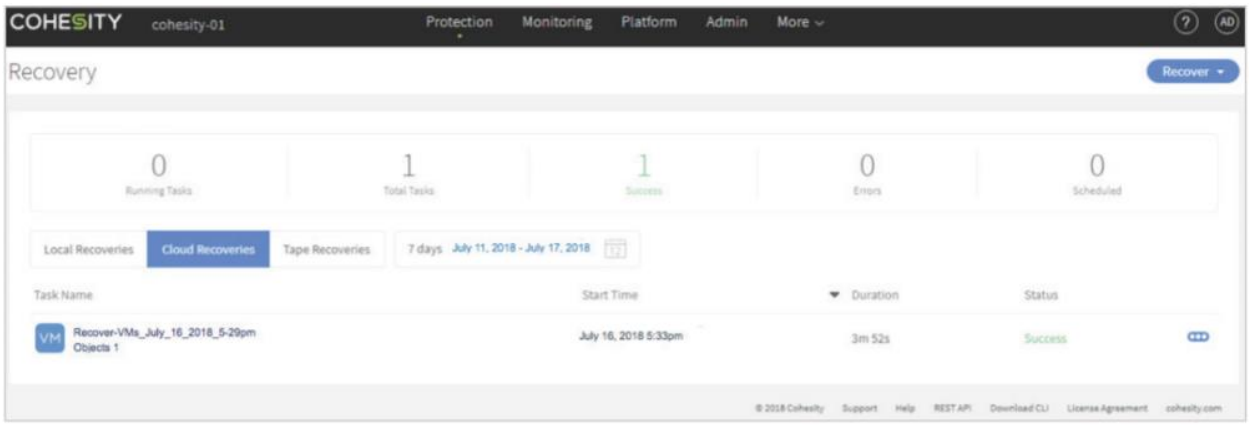


Restore a VM

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.



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



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 Platform 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, see Cohesity's Native Cloud Integration with Azure.

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.

Your Feedback

Was this document helpful? Send us your feedback!

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.0	Aug 2018	First release

ABOUT COHESITY

Cohesity radically simplifies data management. We make it easy to protect, manage, and derive value from data -- across the data center, edge, and cloud. We offer a full suite of services consolidated on one multicloud data platform: backup and recovery, disaster recovery, file and object services, dev/test, and data compliance, security, and analytics -- reducing complexity and eliminating mass data fragmentation. Cohesity can be delivered as a service, self-managed, or provided by a Cohesity-powered partner.

Visit our website and blog, follow us on Twitter and LinkedIn and like us on Facebook.