Using StorSimple Data Transformation

The StorSimple Data Manager is the resource within which Data Transformation can be instantiated. The Data Transformation service lets you move data from your StorSimple on-premises device to blobs in Azure storage. Hence, in the flow that we will see, you will need to specify details about your StorSimple device and the data of interest that you want to move to the storage account.

Create a StorSimple Data Manager service

1. To create a StorSimple Data Manager service, go to https://aka.ms/HybridDataManager

2. Click the **+** icon and search for StorSimple Data Manager. Click your StorSimple Data Manager service and then click **Create**.

3. If your subscription is enabled for creating this service, you will see the following blade.

StorSimple Data Mana...
My Resource - PREVIEW

* Resource name

Enter a name

* Subscription

Microsoft Azure Internal Consumption ⌄

* Select a resource group ⓘ

◉ Create new   ◯ Use existing

* Location

West US ⌄

☐ Pin to dashboard

Create   Automation options

4. Enter the inputs and click **Create**. Note that the specified location should be the one that houses your storage accounts and your StorSimple Manager service. It takes about a minute to create the service.

Creating a data transformation job definition

Within a StorSimple Data Manager service, you need to create a data transformation job definition. A job definition specifies details of the data that you are interested in moving into a storage account in the native format.

1. Navigate to the service that you just created. Click **+ Job Definition**.

2. The new job definition blade opens up. Give your job definition a name and click **Source**. In the **Configure data source** blade, specify the details of your StorSimple device and the data of interest.

New job definition
PREVIEW

* Job definition name

vivtestjd1 ✓

1. Service
   Data transformation ✓

2. Source
   Configure source settings >

3. Target
   Configure target settings >

☐ Present an estimate of the job run duration & data and wait for my consent before executing the job. ⓘ

OK

Configure data source ...
PREVIEW

* Data repository

[⌄]

* Device name

Enter device name

* Volume

Enter volume name

Filter

Root directories

Eg: "MyRootDirectory\SubFolder"

Filter files

*

BackupChoice

Take backup now | Use latest existing backup

ⓘ A cloud snapshot for the specified volume will be taken for this job at runtime.

OK

Screensho

Past
620x

3. Since this is a new Data Manager service, no data repositories are configured. To add your StorSimple Manager as a data repository, click **Add new** in the data repository dropdown and then click **Add Data Repository**.

4. Choose **StorSimple 8000 series Manager** as the repository type and enter the properties of your **StorSimple Manager**. For the resource ID field, you need to enter the number before the ':' in the registration key of your StorSimple manager.

Configure data source ...
PREVIEW

* Data repository

Add new	⌄

```
Add data repository
Configure settings                    >
```

* Device name

```
Enter device name
```

* Volume

```
Enter volume name
```

Filter

Root directories

```
Eg: "MyRootDirectory\SubFolder"
```

Filter files

```
*
```

BackupChoice

Take backup now	Use latest existing backup

> ℹ A cloud snapshot for the specified volume will be taken for this job at runtime.

OK

Add new data repository
PREVIEW

* Data repository type

StorSimple 8000 series manager	⌄

* Data store friendly name

```
Enter a friendly name
```

* Subscription id

```
Enter the subscription id
```

* Resource group name

```
Enter the resource group name
```

* Resource id

```
Enter the resource id
```

* Device manager name

```
Enter the name of the repository
```

* Service encryption key for 8000 series

```
Enter the key
```

OK

5. Click **OK** when done. This saves your data repository and this StorSimple Manager can be reused in other job definitions without entering these parameters again. It takes a few seconds after you click OK for the StorSimple Manager to show up in the dropdown.

6. In the **Configure data source** blade, enter the device name and the volume name that has your data of interest.

7. In the **Filter** subsection, enter the root directory that contains your data of interest (this field should start with a '\'). You can also add any file filters here.

8. The data transformation service works on the data that is pushed up to the Azure via snapshots. While running this job, you have an option to choose if you want the backup to be taken every time this job is run (to work on latest data) or to use the last existing backup in the cloud (if you are working on some archived data).

9. Next, the Target settings need to be configured. There are 2 types of supported targets – Azure Storage accounts and Azure Media Services accounts. Choosing storage accounts puts files into blobs in that account. Choosing media services account puts files into assets in that account. Again, we need to add a repository. In the dropdown, select **Add new** and then **Configure settings**.

10. Here, you can select the type of repository you want to add and the other parameters associated with the repository. In both cases, a storage queue is created when the job runs. This queue is populated with messages about transformed blobs as they are ready. The name of this queue is the same as the name of the job definition. If you select **Media Services** as the repo type, then you can also enter storage account credentials where the queue will be created.

11. After adding the data repository (which takes a few seconds), you will find the repo in the dropdown in the **Target account name**. Choose the target that you need.

12. Click **OK** to create the job definition. Your job definition is now set up. You can use this job definition multiple times via the UI.

New job definition
PREVIEW

* Job definition name

vivtestjd1 ✓

1 Service
Data transformation ✓

2 Source
abc ✓

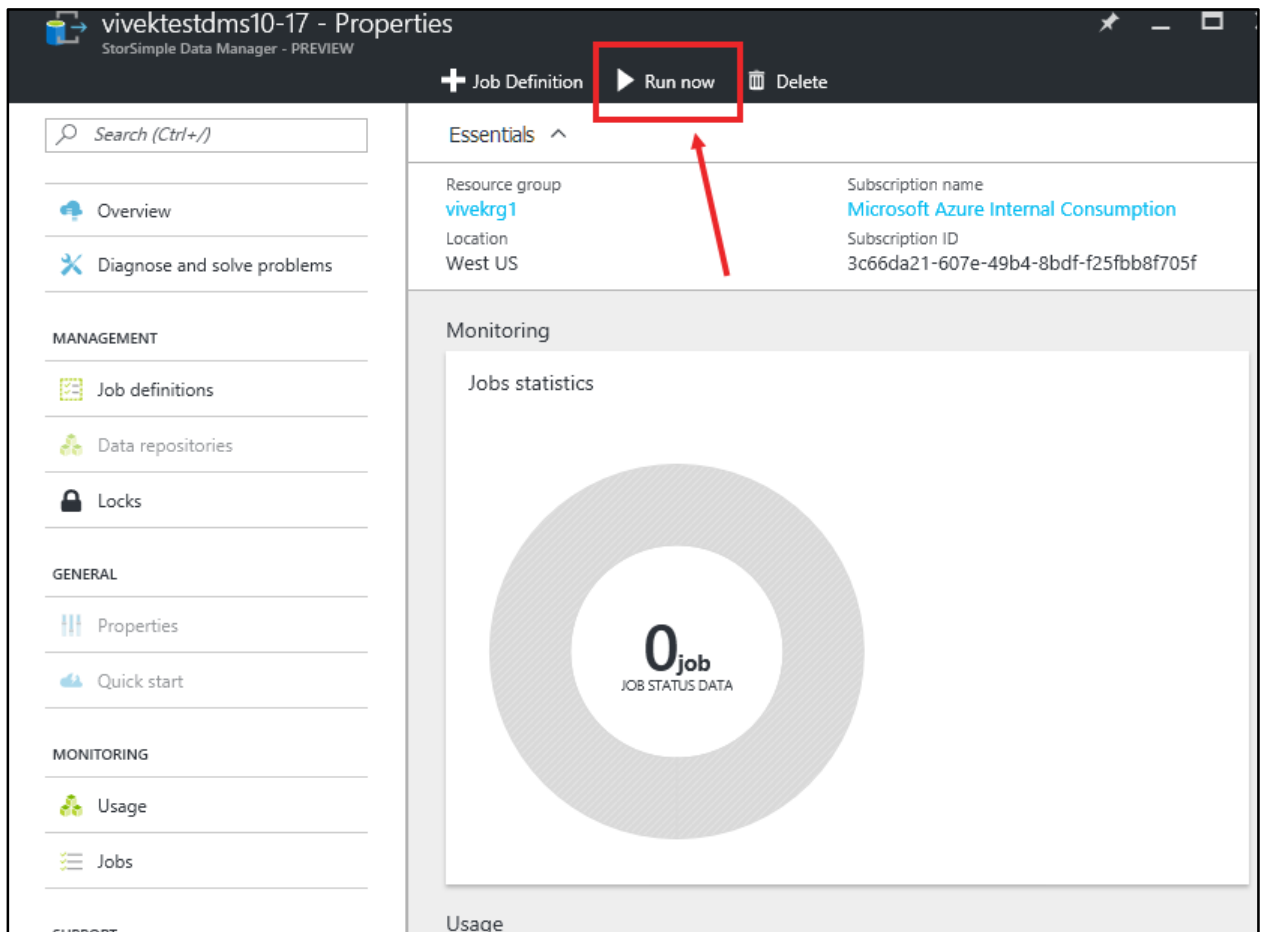**3** Target
Target account name ✓

☐ Present an estimate of the job run duration &
data and wait for my consent before
executing the job. ⓘ

OK

Running the job definition

Whenever you need to move data from StorSimple to the storage account that you have specified in the job definition, you will need to invoke it. There is some flexibility in changing the parameters every time you invoke the job. The steps are as follows:

1. Select your StorSimple Data Manager service and go the Monitoring. Click **Run Now**.

2. Choose the job definition that you want to run. Click **Run settings** to modify any settings that you might want to change for this job run.

Run job definition
PREVIEW

* Data service

Data transformation

* Job definition name

vivektestjd1

---

*

Run settings   >

1

Run

Run now
PREVIEW

Device name

TechReadyDevice1

Volume

techreadyvol2

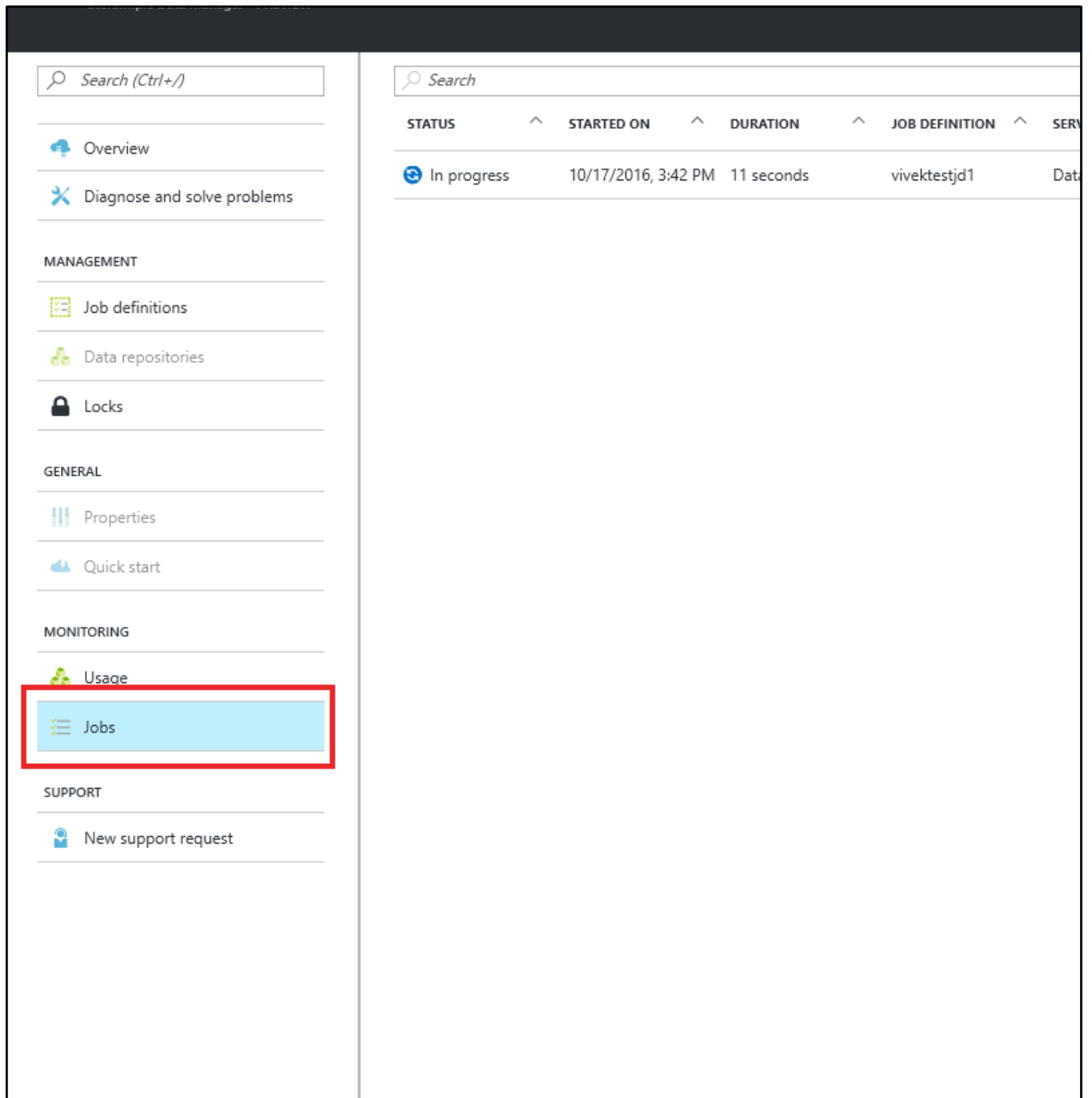Root directory

\

Filter files

*.mp3

Backup choice

Take backup now   Use latest existing backup

☐ Present an estimate of the job run duration &
   data and wait for my consent before
   executing the job. ⓘ

2

OK

3. Click OK and then click Run. This will launch your job. To monitor this job, go to Jobs page in your StorSimple Data Manager.

4. In addition to monitoring in the jobs page, you can also listen on the storage queue where a message is added every time a file is moved from StorSimple to the storage account.