VERITAS | PURESTORAGE®

Veritas NetBackup with Pure Storage FlashArray//C Solution

Technical White Paper | April 2022

Contents

Introduction

Ransomware, malware, and computer viruses have been a hot topic in current news. There are many reports on how businesses have been impacted by malware or ransomware attacks. With growing concerns over the safety of their data and business continuity, companies are looking for data protection solutions that offer rapid recovery at scale. Veritas NetBackup™ with Pure Storage® FlashArray//C provides data protection and management of digital assets with an "all-flash" solution for mass recovery from a cyberattack or disaster. Using FlashArray//C also improves the performance of NetBackup Instant Access by delivering faster third-party malware scanning of backup images.

Solution Value

NetBackup with FlashArray//C addresses the numerous challenges companies face today where cyberthreats and attacks are more prevalent and the need to protect digital assets is critical.

- **Mass recovery at scale**—Using all-flash technology like FlashArray//C lowers the recovery time objective (RTO) and enhances business continuity for organizations that require rapid restore of their data, especially in cases where mass recovery is required due to malware or disaster.
- **Improved performance**—The NetBackup Instant Access feature with FlashArray//C allows companies to restore data and make it readily available for use by third-party malware scanning or analysis tools.
- **Balanced cost and performance**—Combining NetBackup with Pure Storage's most cost-effective storage solution, FlashArray//C, balances cost and performance. Organizations can also realize additional cost savings with their deduplication and compression features.

Solution Overview

NetBackup can use various storage types such as the FlashArray//C as target storage for backup images. Figure 1 provides an illustrated view of the solution. With this solution, FlashArray//C is the target storage for the NetBackup backup images and metadata.

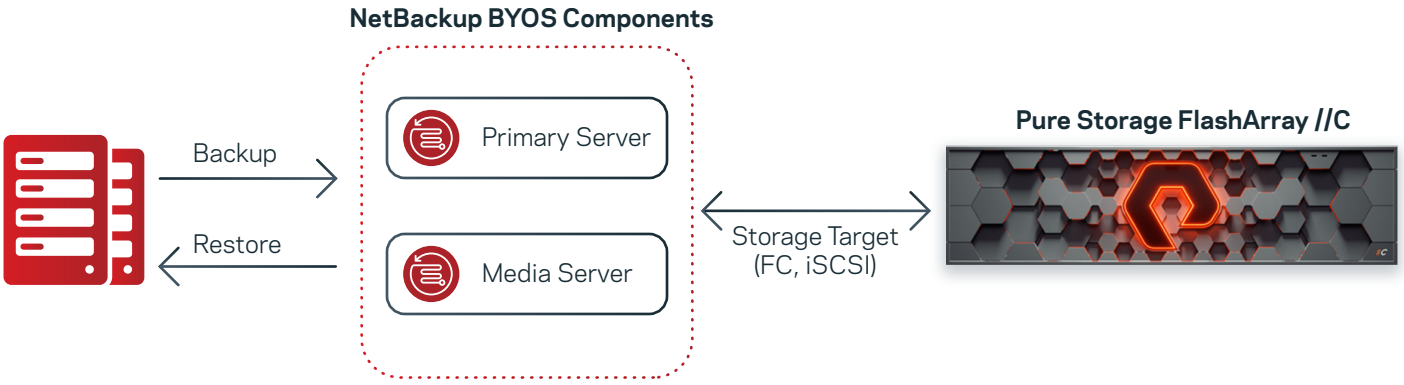


Figure 1. An overview of NetBackup with FlashArray//C Solution integration.

NOTE: If you are using NetBackup Media Server Deduplication Pool (MSDP) technology, the MSDP size for NetBackup Build Your Own Server (BYOS) is limited to 250 TB per media server.

NetBackup

NetBackup provides protection for a wide variety of data and platforms, including operating systems (OSs), virtual systems, databases and applications, files, and all kinds of structured and unstructured content. It has many add-on features to speed up and automate backups and snapshot management. NetBackup can backup data to tape, storage area network (SAN), network-attached storage (NAS), and public or private clouds. Schedules, retention periods, and the ability to tier to different types of storage are defined in policies or storage lifecycle policies (SLPs).

A typical NetBackup environment consists of three components:

1. Primary Server—Manages and controls the backup and recovery activities and hosts the catalog that contains policies and schedules, files, metadata about the backup jobs, and media, device, and image metadata.
2. Media Server—Writes client data as backup images to storage such as local disks, tape, NAS, SAN, and the cloud, and restores the data to the client as instructed by the primary server.
3. Clients—NetBackup client components are installed on hosts with the data to be backed up and are responsible for sending and receiving data to and from the media server for backup and recovery.

In this solution, Veritas recommends using FlashArray//C as an MSDP storage target. The unique blocks or deduplicated blocks are stored on FlashArray//C, and the deduplication is conducted and managed by the media server. The NetBackup MSDP deduplication architecture is composed of the following main components:

- Deduplication Plug-in—Separates the data into segments or chunks. Uses a hash algorithm to calculate fingerprints to identify each unique segment. Compares incoming data fingerprints with the fingerprints of existing data.
- Deduplication Engine (spoold)—Manages and stores the fingerprint database and metadata, stores unique segments, or uses a reference or pointer to the data already stored, and conducts integrity checks.
- Deduplication Manager (spad)—Maintains the configuration, controls, and dispatches the internal processes, security, and events handling.

NetBackup MSDP uses SHA-2 (SHA256) for the hash algorithm. The chunk segment size unit used to compute fingerprints is by default a fixed length of 128 KB or configurable to a variable-length size based on the chunk boundary. NetBackup MSDP also compresses deduplicated data for further storage efficiency and offers an option to encrypt deduplicated data. For more information on the architecture of NetBackup MSDP deduplication technology, refer to the NetBackup Deduplication Guide.

FlashArray//C

FlashArray//C is Pure Storage's most cost-effective solution for workloads that require higher capacity and performance that all quad-level cell (QLC) flash array storage provides. FlashArray//C assists in simplifying and modernizing data protection solutions with the performance and availability benefits of flash. It is designed to be modular and upgradable. FlashArray//C has capacity options between 797 TB to 7.3 PB. There are two models available:

- //C40 – up to 1.9 PB/1.4 PiB effective capacity
- //C60 – up to 7.3 PB/6.6 PiB effective capacity

Key features offered by FlashArray//C:

- Protocol supported include NVMe-oF (RoCE), iSCSI, Fibre Channel (FC), Server Message Block (SMB) and Network File System (NFS)
- 99.9999% availability
- 100% non-disruptive architecture
- QLC Direct Flash modules
- Consistent 2-4ms latency

Refer to the FlashArray//C data sheet for more detail.

Solution Best Practices

This section highlights some of the best practices for deploying NetBackup with FlashArray//C. For additional best practices, refer to the best practices of NetBackup and FlashArray//C respectively:

- NetBackup Backup Planning and Performance Tuning Guide
  - https://www.veritas.com/content/support/en_US/doc/21414900-146141073-0/v146020053-146141073
- FlashArray Host Connectivity Guide for Specific Platforms (for example, Linux, AIX, Windows, Solaris, VMware)
  - https://support.purestorage.com

Provisioning Storage for Deduplicated Data

Organizations can improve performance by using multiple FlashArray//C volumes as opposed to a single volume. Using multiple volumes allows for additional parallelism, paths, and I/O queues to FlashArray. As a best practice, the minimum file system size created should be at least 10 TB when using multiple file systems in a single MSDP. Also, the number of file systems per MSDP should be limited to no more than 12. As previously mentioned, the maximum size of an MSDP supported for NetBackup BYOS is 250 TB. When deploying MSDP in this fashion, Veritas recommends that the MSDP metadata reside in a separate volume and file system.

If multiple file systems are aggregated to create an MSDP pool, Veritas recommends that each volume is of the same size and that the file system grows at the same rate. For example, if you have created eight volumes and the MSDP is to grow by 80 TB, then you should expand each file system by 10 TB. One of the criteria used to determine which file system to send data to is based on the available capacity in that file system to provide an even distribution of the data across file systems. There is another option to grow the MSDP by adding another file system; however, we don't advise doing so due to the criteria mentioned. If you want to use this approach, we advise creating a new file system of the same size as the existing ones. Please refer to the Veritas NetBackup Deduplication Guide for guidance when resizing the MSDP storage partition. **NOTE**: Each platform (such as Linux or VMware) and file system has different methods to resize the underlying storage, and some may be disruptive. Refer to each platform's instructions on how best to resize a volume or file system after you have resized FlashArray//C.

MaxCacheSize

The MaxCacheSize setting determines the number of fingerprint indices that can be cached in memory per pool and is a factor in the deduplication ratio. Incoming data fingerprints are compared with the existing fingerprints that are present in the physical memory cache. If the fingerprint is not present, then data is not a candidate for deduplication. As a best practice, allocate 0.75 to 1 GB of RAM for each TiB of storage allocated to the pool.

Deduplication and Compression

Both NetBackup and FlashArray//C have deduplication and compression enabled by default. Having both enabled does not affect overall storage efficiency. In some scenarios where NetBackup is not able to achieve deduplication, such as when fingerprints for comparison are not in the physical memory cache of the media server, then you can use FlashArray//C to achieve further data reduction . Also, if you are using FlashArray//C as the target storage for multiple media servers, you may see data reduction in this scenario as well. As a best practice, keep the default settings for compression on NetBackup for deduplicated data.

Encryption

FlashArray//C encryption is enabled by default for data at rest and keys are distributed. If you require encryption "in flight" to the storage target, then you should implement NetBackup encryption as well. NetBackup offers policy-based encryption and encryption for deduplicated data.

Disaster Recovery

As a best practice, organizations should have a disaster recovery (DR) plan in place in case of catastrophic events. NetBackup Auto Image Replication (AIR) is supported when using FlashArray//C as an MSDP target for deduplicated data. AIR allows the NetBackup catalogs and data from one domain to be replicated asynchronously to another NetBackup domain in one or several geographical sites for DR. There are other ways to protect catalogs associated with NetBackup from accidental deletion or corruption such as conducting regular backups of the NetBackup and MSDP catalog.

Linux Tunable

If you are using the Linux OS as the platform to run NetBackup, there are some recommended settings for optimum performance of FlashArray//C. Performance is better when setting the scheduler mode to noop instead of using the default deadline. For more information on Pure Storage's recommended settings for Linux, please refer to Pure Storage documentation: https://support.purestorage.com/Solutions/Linux/Linux_Reference/Linux_Recommended_Settings.

If you are using the Veritas Volume Manager (VxVM) to create a disk group of the FlashArray//C volumes, enabling VxVm queuing improves the I/O request size and overall restore performance.

Planning and Sizing

A organization's specific requirements for performance and capacity will dictate the number of media servers and number of FlashArray//C to deploy. Some parameters that may affect these requirements include the size of the source data, the daily change rate of the data, annual storage growth, desired retention, deduplication rate, and whether you are conducting weekly, daily, or monthly incremental or full backups. It is best to work with the Veritas and Pure Storage account teams that have tools to appropriately size the environment based on your desired requirements. You can also refer to the "Sizing for capacity with MSDP" section in the NetBackup Planning and Performance Guide.

In the context of storage consumption on the FlashArray//C with NetBackup, there is overhead associated with the backup images, such as the directory structure and the metadata stored with the backup images. With NetBackup deduplication, there is additional metadata that resides by default on the same storage target as the deduplicated data. As a best practice, we recommend specifying a different volume to hold this metadata if you are using multiple volumes in one MSDP. When configuring the MSDP, there is an option to specify an alternate location for this metadata.

Using multiple media servers to load balance the fingerprint calculation does improve overall throughput; however, you should add a load-balancing media server only if you see high CPU utilization on the media server acting as a storage server.

It is important to set the concurrent jobs and maximum I/O streams appropriately for NetBackup disk pools to prevent overload or to underutilize compute and storage resources. For best practices relating to these settings, please refer to the "Best practices: Disk pool configuration – setting concurrent jobs and maximum I/O streams" section in the NetBackup Backup Planning and Performance Guide.

Solution Deployment

This section describes the generic steps for configuring multiple volumes on FlashArray//C with a NetBackup BYOS. Configuration and setup will differ depending on requirements including the type of environment (hardware or virtualized), OS platform, and the physical connectivity of FlashArray//C to the server running NetBackup. Please refer to the NetBackup hardware and software compatibility matrix for a list of supported hardware and OSs. In this deployment example, FlashArray//C is directly connected to the back of ESXi servers using FC and NetBackup is installed on a virtual machine (VM) running Red Hat Enterprise Linux (RHEL) 7.8. Figure 2 shows the topology of this connectivity. The ESXi server has a host bus adapter (HBA) with two ports, and it is connected to the Pure Storage FlashArray in a redundant fashion. Multiple volumes are created on FlashArray//C, and each volume is formatted with a file system, which is then combined to create a single MSDP.

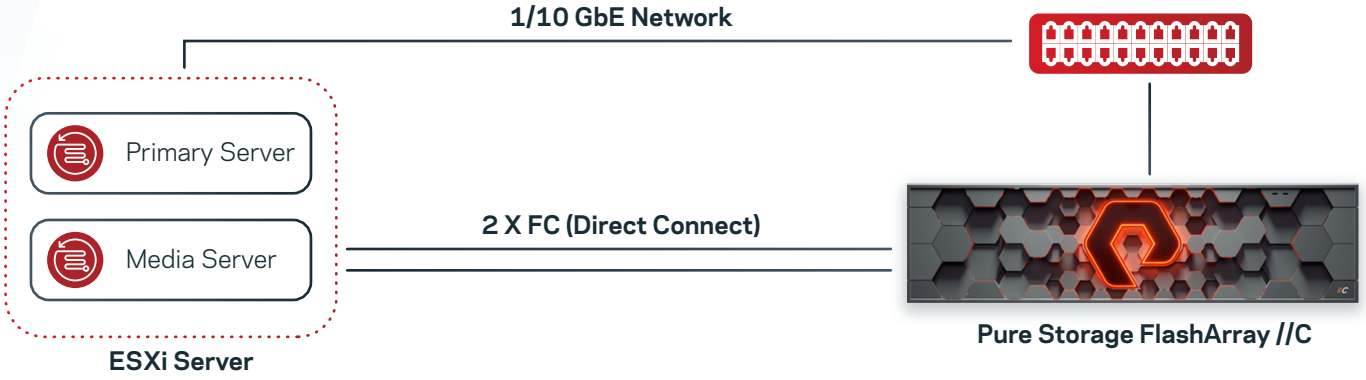


Figure 2. An overview of NetBackup with FlashArray//C topology.

To configure FlashArray//C for use with NetBackup:

1. Physically connect the FlashArray//C volumes to the ESXi servers.
2. Create volumes on FlashArray//C and connect to the ESXi hosts.
3. Rescan volumes on the server (ESXi hosts) and attach to the VMs.
4. Rescan the devices on the OS, if needed.
5. Format and mount the device.
6. Allocate the device as the target for MSDP for both data and metadata.
7. Validate the configuration.

Physical Connection

Pure Storage FlashArray//C with NetBackup supports the FC and iSCSI protocols. In this example, there were only two ports available on the ESXi server. For redundancy, one HBA port was directly connected to a port on the FlashArray//C controller 0 (CT0), and the other port was connected to a port on controller 1 (CT1). You can see this connection from the FlashArray//C web user interface (UI), as shown in Figure 3.

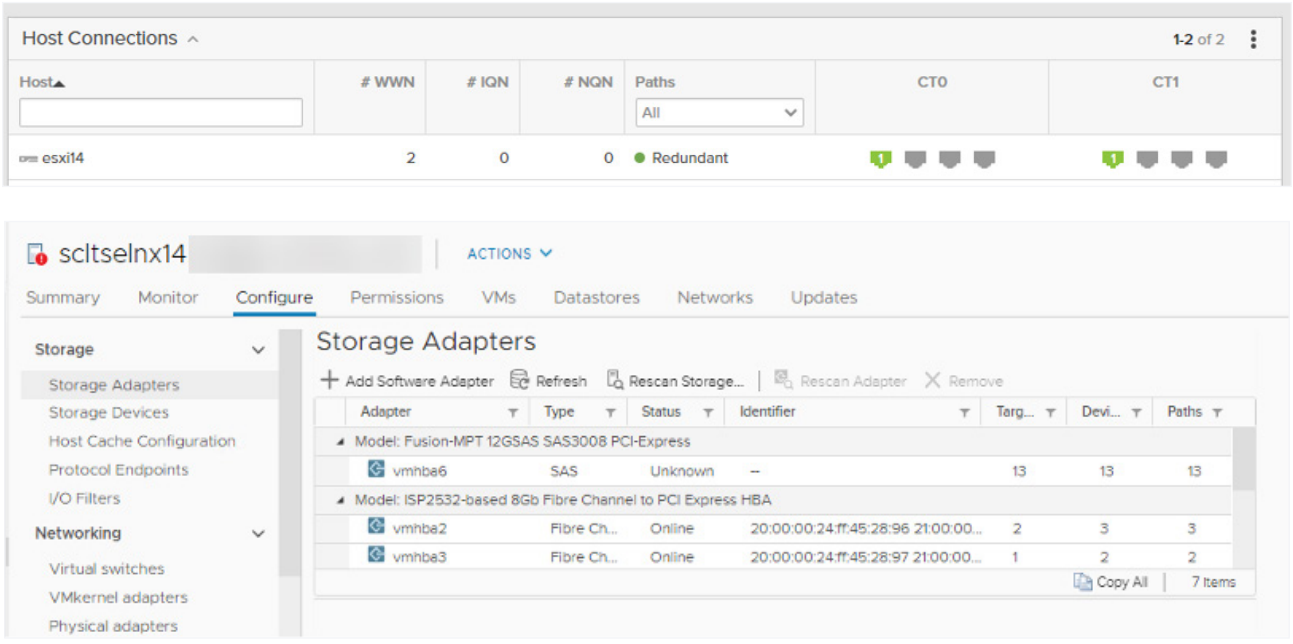


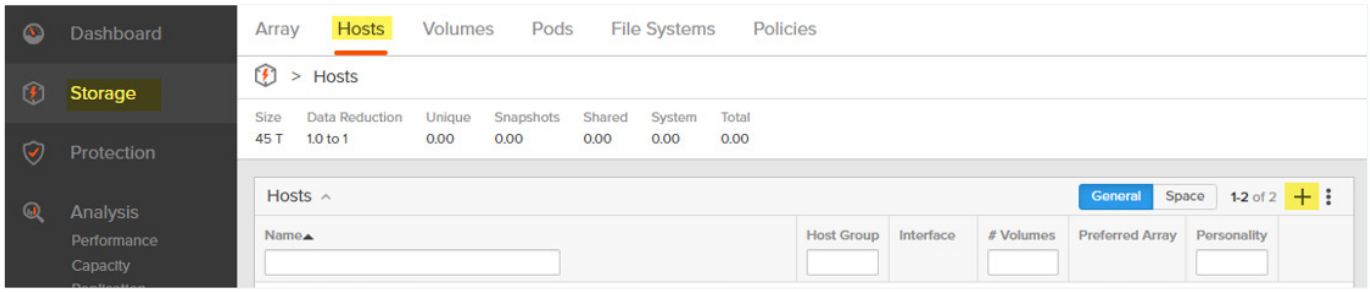
Figure 3. The FlashArray//C web UI within the ESXi vCenter.

Creating and Connecting Hosts and Volumes

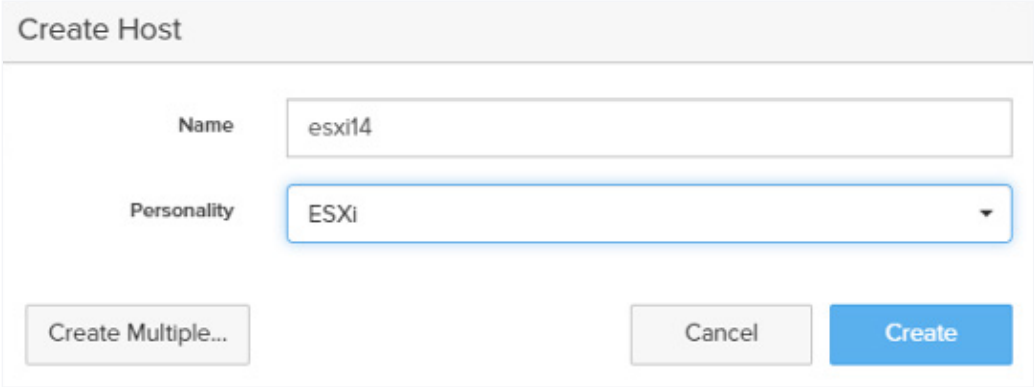
This section describes how to create the hosts and volumes and how to connect the volumes to specified hosts.

Creating Hosts

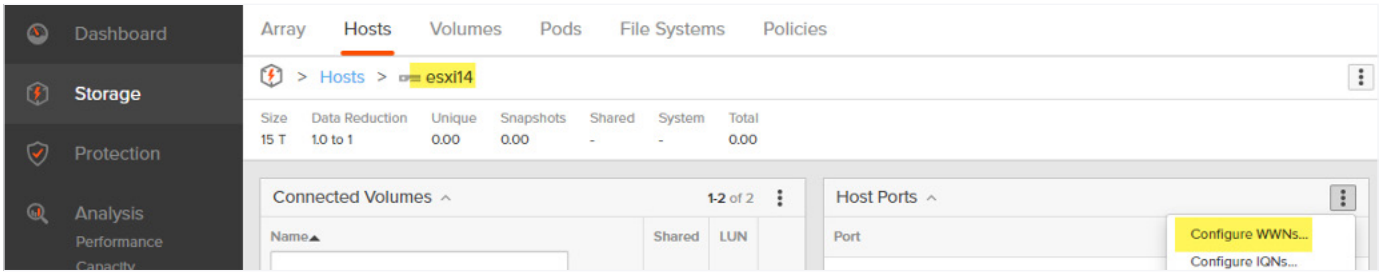
1. Log onto the FlashArray//C web UI using credentials.
2. Select **Storage** on the left pane.
3. Select the **Hosts** tab on the right pane.
4. Hit the + sign on the far right under Hosts and create the hosts.



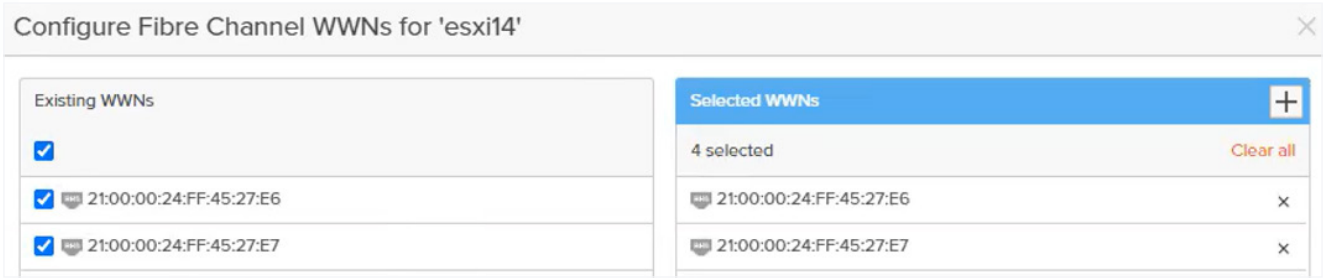
5. Enter the information relating to the hosts, such as **Name** and **Personality**.



6. After creation, click on **Hosts** created, such as esxi14, and click on **Configure WWNs** (World Wide Name) under the Host Ports. The WWN is a unique identifier associated with the FlashArray//C as an FC device.



7. Select the **WWNs** connected to the hosts.

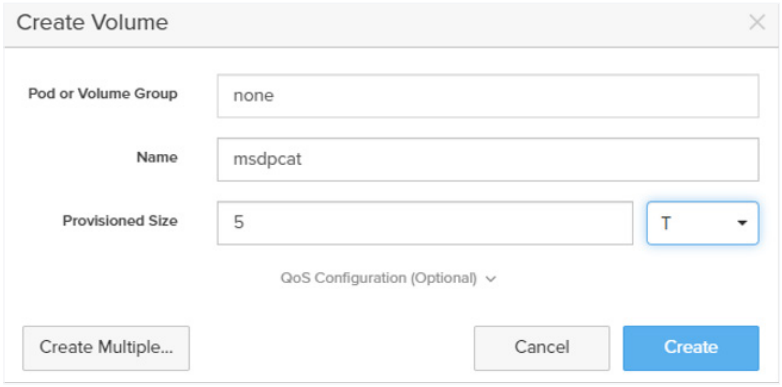


Creating Volumes

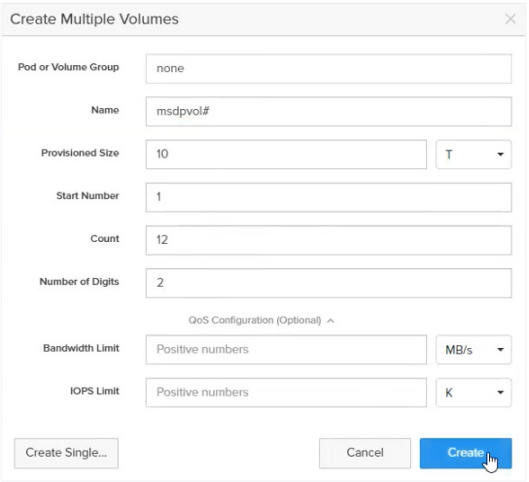
NOTE: In this example, we will create 1 volume for the MSDP metadata and 12 volumes for the MSDP data.

1. Click on the **Volumes** tab from the Storage view and click on the **+** sign on the far right of the pane.

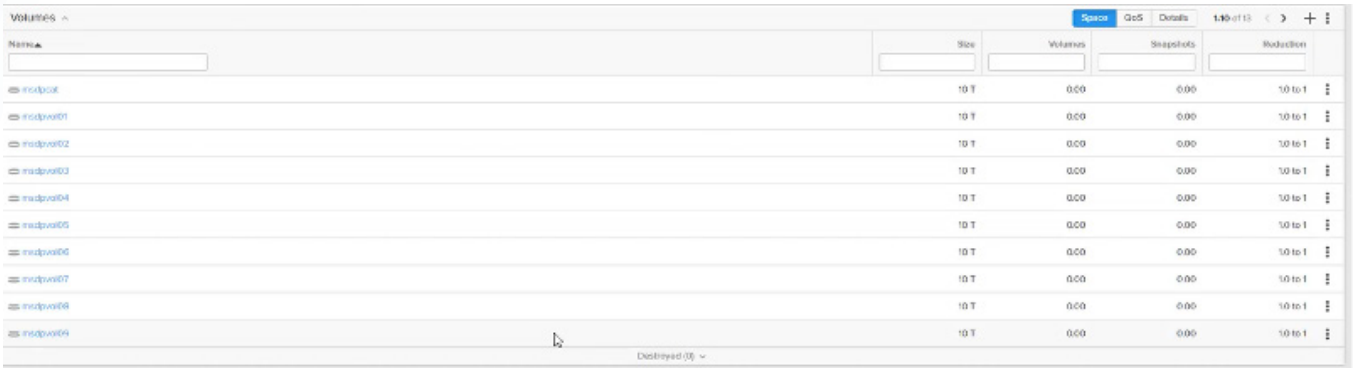
2. Create a single volume for NetBackup MSDP metadata. Specify the **Name** and **Provisioned Size**.



3. There is an option to create multiple volumes at one time. For the MSDP data, you will create 12 volumes. Click the + sign again and select **Create Multiple**. Enter the create Name, Provisioned Size, start number, count, and number of digits and click **Create**. In the example below, the Name is prefixed with msdpvol followed by a number starting with number 1 and the number of digits for the number is 2. The provisioned size is 10 TB.

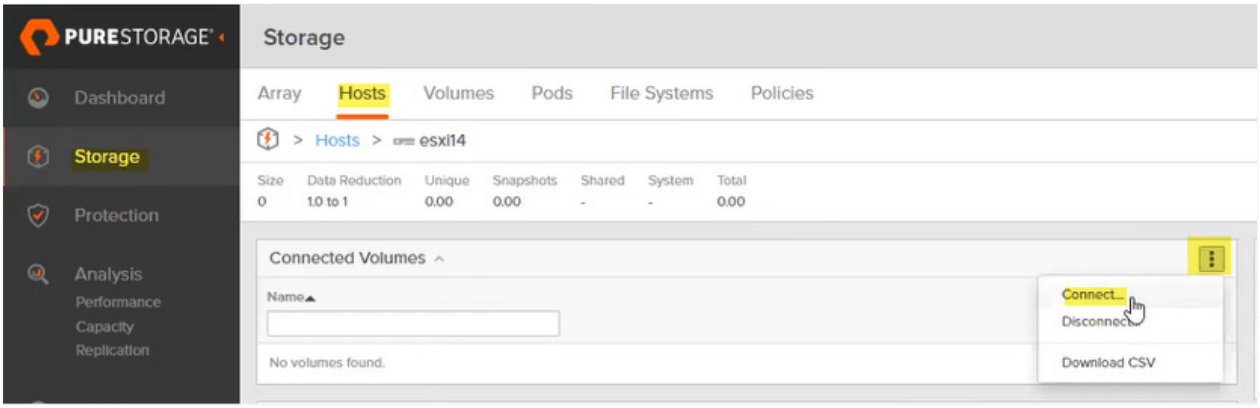


4. Validate the creation. **NOTE**: Only 1 of 10 volumes is shown below.

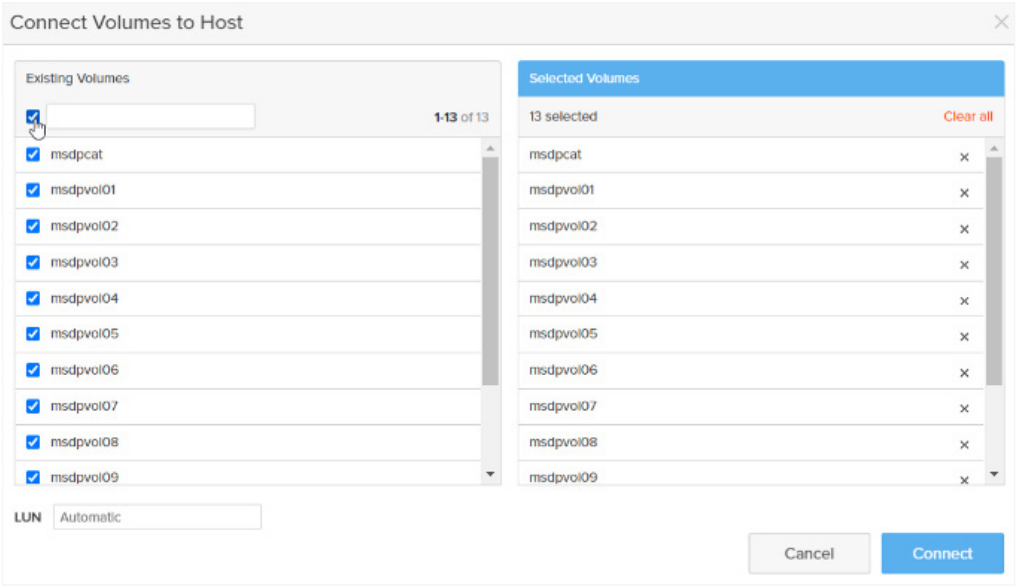


Connecting Volumes to Hosts

1. For each of the volumes created in the previous steps, you need to connect the volumes to the hosts. Click on the **Hosts** tab from the Storage view. Click on the 3 dots on the top right of Connected Volumes and select **Connect**.



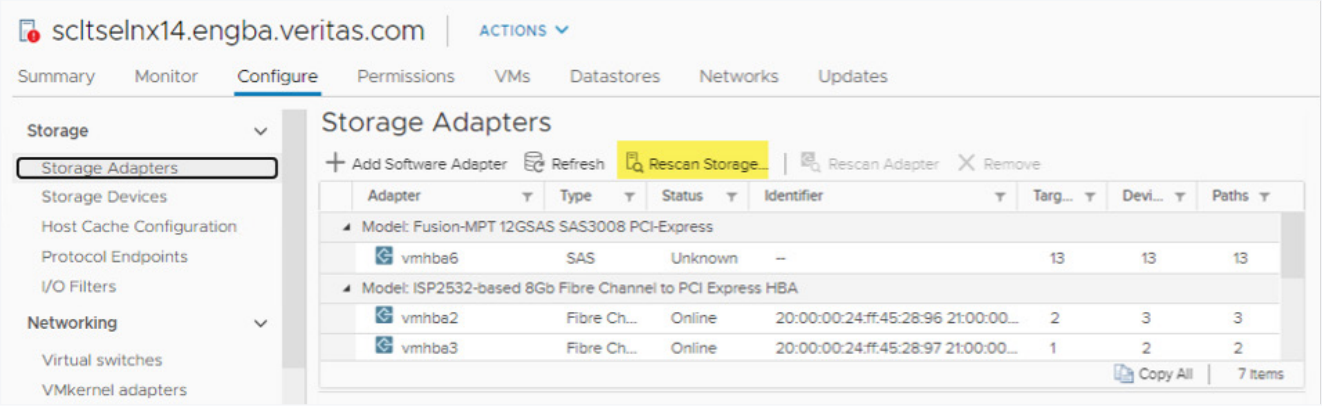
2. Select the devices to connect to the hosts. **NOTE**: You can connect all volumes at once. Click in the top check box under Existing Volumes to select all volumes and click **Connect**.



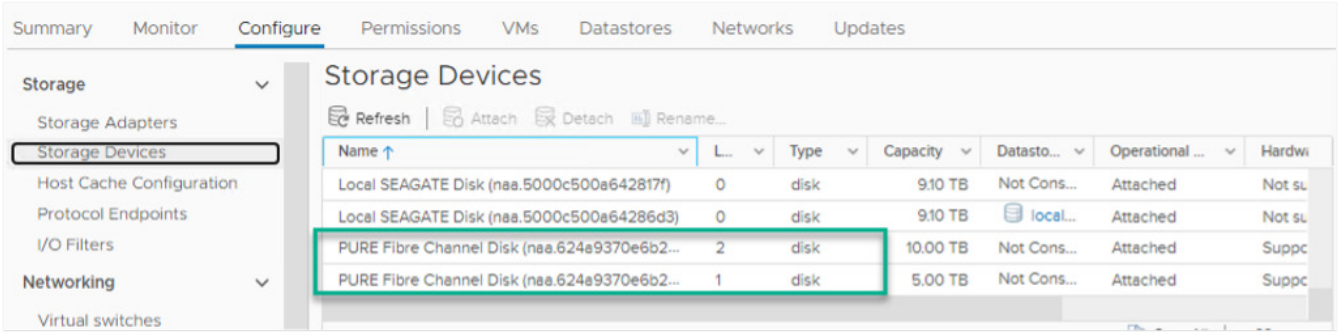
Rescanning Devices on Hosts and Connecting Devices to VMs

In this example, the FlashArray is connected directly to an ESXi server. The ESXI server will do the rescan of the volumes and then the guest OS running on the VM will rescan, if needed.

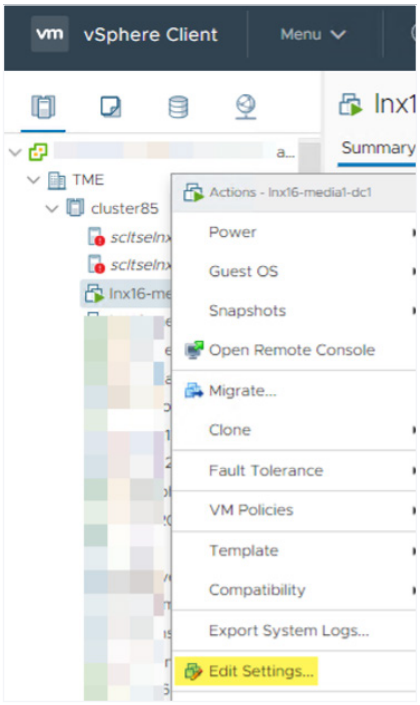
1. From vCenter, click on the **ESXi server**, select the **Configure** tab. Under **Storage** in the left-hand menu, select **Storage Adapters**, and then click on **Rescan Storage**.



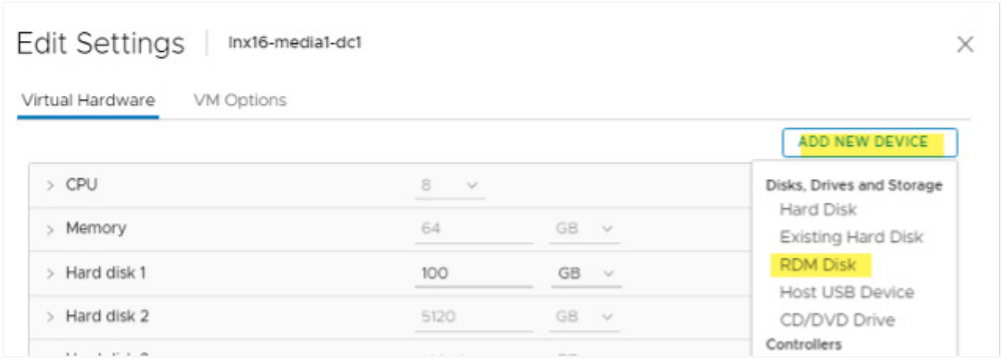
2. **Refresh** and validate the devices are recognized under **Storage Devices**.



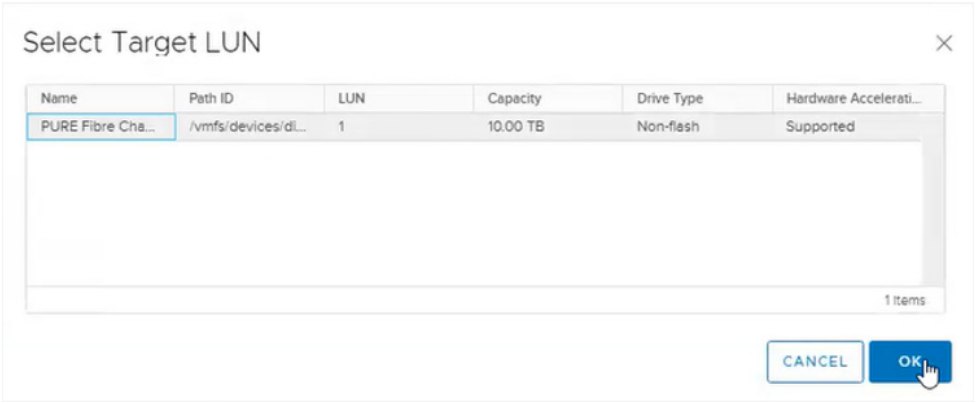
3. Add the device to the VM. Select the VM and then choose **Edit Settings** from the Actions drop-down.



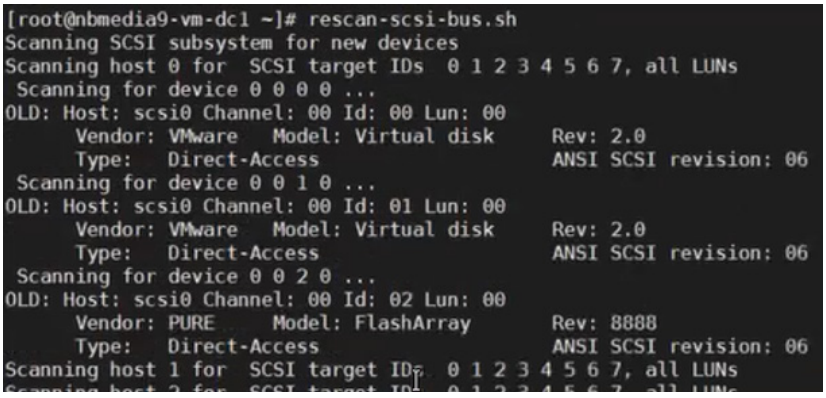
4. Add a new device of the **RDM Disk** type.



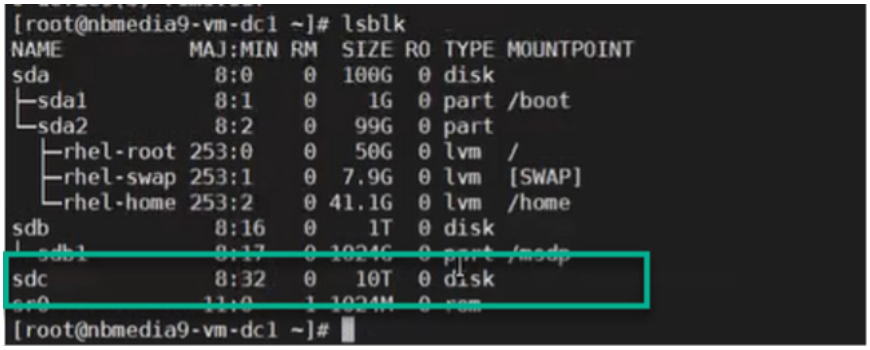
5. Select the target LUN and click **OK**.



6. Use the Secure Shell protocol (SSH) to get onto the VM guest OS (RHEL) and do a **rescan-iscsi-bus.sh (located in /usr/bin)**.
   NOTE: Some OSs may automatically recognize the attached disks and may not need this step.



7. Validate the device is discovered using the **lsblk** command.

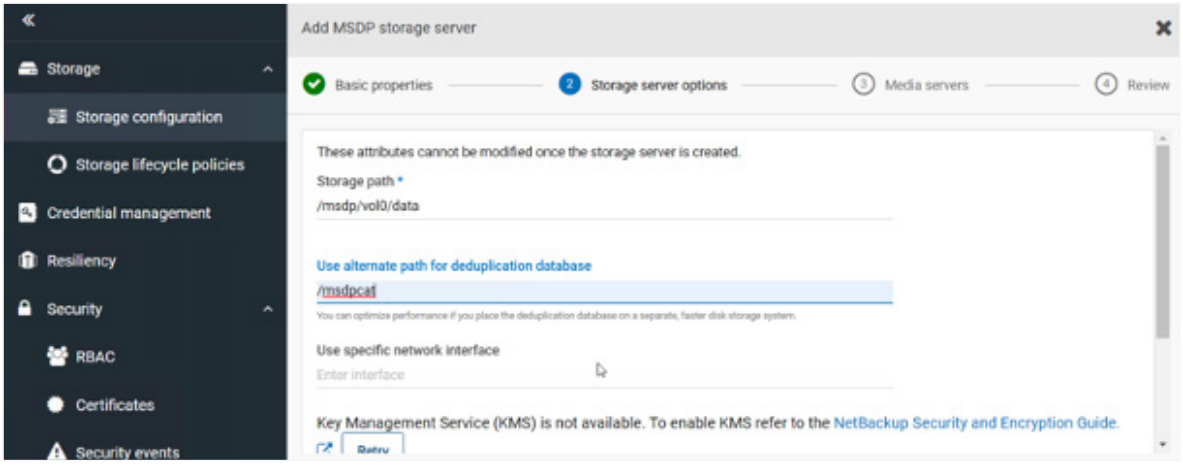


Formatting and Mounting Volumes

1. Create a file system on each volume/device. In this example, you will use an xfs file system. Therefore, run "mkfs -t xfs /dev/sdc."
2. Next, prepare the volumes and assign them to the MSDP. Refer to this Knowledge Base article for detailed instructions or refer to the product documentation.
   a. Modify the /etc/fstab file so one volume is mounted to /msdp/cat and the other 12 file systems are mounted on /msdp/vol0, /msdp/vol1, and so on until you have mounted each volume.
   b. Mount each of the file systems.
   c. Create a sub-directory called 'data' under each mounted volume: for example, /msdp/vol0/data, /msdp/vol1/data, /msdp/vol2/data, and so on.

Allocating a Device as a Target for MSDP for Data and Metadata

1. Configure the MSDP through the **NetBackup web UI**. Specify the **Storage path** and **alternate path for deduplication database**. So, in this example, provide the storage path of /msdp/vol0/data and the database path of /msdpcat.

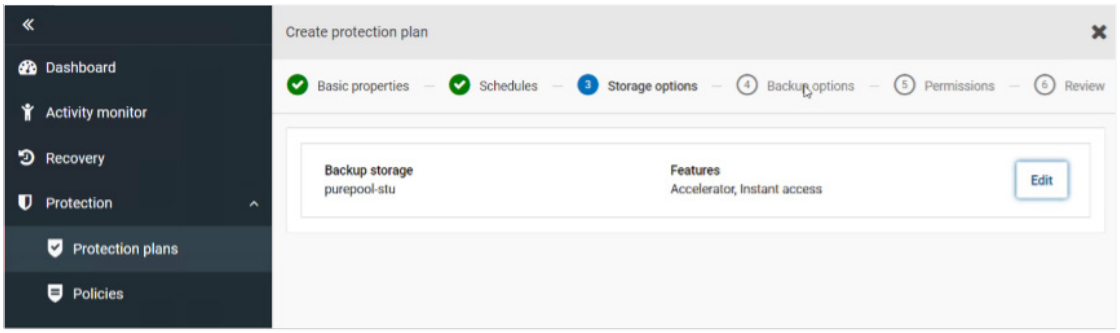


2. Add the other file systems to the deduplication pool by entering the following commands on the media server:
   /usr/openv/pdde/pdcr/bin/crcontrol --dsaddpartition /msdp/vol1/data
   /usr/openv/pdde/pdcr/bin/crcontrol --dsaddpartition /msdp/vol2/data
   .... and so on through vol12...
   /usr/openv/pdde/pdcr/bin/crcontrol --dsaddpartition /msdp/vol12/data
3. The MSDP configuration is now complete. Next, verify the deduplication pool contains the new volumes. Review the following command output to verify these volumes:
   /usr/openv/pdde/pdcr/bin/crcontrol --dsstat 2 | grep Mount
   point count: 12

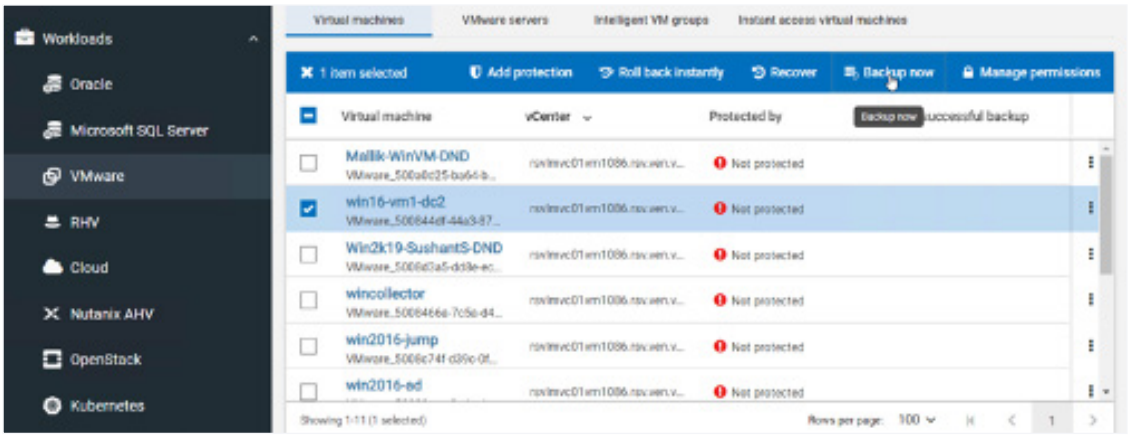
   NOTE: Another option is to use the Veritas Volume Manager to create a disk group of volumes and then create a single file system on that disk group. Also, if you want to use only one volume, then you just need to create a file system on the volume, mount the file system, and modify the /etc/fstab to mount the file system after reboots.

Validating the Configuration

1. Define a policy that would back up to the storage unit created in the previous steps.



2. Conduct a **Backup** now of a source.



References

- Pure Storage
  - https://www.purestorage.com/products/nvme/high-capacity/flasharray-c.html
  - https://support.purestorage.com/Solutions/Veritas
- Veritas Product Documentation
  - https://sort.veritas.com/documents
  - NetBackup 9.1 https://sort.veritas.com/documents/doc_details/nbu/9.1/Windows%20and%20UNIX/Documentation/

About Veritas

Veritas Technologies is a global leader in data protection and availability. Over 80,000 customers—including 87 percent of the Fortune Global 500—rely on us to abstract IT complexity and simplify data management. The Veritas Enterprise Data Services Platform automates the protection and orchestrates the recovery of data everywhere it lives, ensures 24/7 availability of business-critical applications, and provides enterprises with the insights they need to comply with evolving data regulations. With a reputation for reliability at scale and a deployment model to fit any need, Veritas Enterprise Data Services Platform supports more than 800 different data sources, over 100 different operating systems, more than 1,400 storage targets, and more than 60 different cloud platforms. Learn more at www.veritas.com. Follow us on Twitter at @veritastechllc.

VERITAS

2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
veritas.com

For global contact information visit:
veritas.com/company/contact

V1538 04/22