

VMware Data Recovery Dedupe Store Setup Guideline

Introduction

VMware Data Recovery (VDR) is a disk-based backup and recovery solution for protecting your virtual machines. It incorporates capabilities such as block based data deduplication and performing only incremental backups after the first full backup to maximize storage efficiency. As a result of deduplication and incremental backups, VDR significantly reduces the space requirements in the destination disk where the backup data is stored - we will refer to the destination disk as dedupe stores . (For more information on the VDR architecture, please refer to the administrator's guide at www.vmware.com/pdf/vdr_10_admin.pdf.)

This article provides guidelines surrounding the setup of dedupe stores to assist you in setting up your VDR environment.

Intended Audience

This article is intended to provide guidelines to IT professionals who plan to set up a VDR environment to protect their virtual machines in their vSphere infrastructure.

Author

Desmond Chan, VMware Technical Marketing Manager focusing on Business Continuity and Disaster Recovery.

Guidelines

General Sizing Guidelines

You can attach at most 2 dedupe stores to a VDR Appliance. Each dedupe store can be at most 1 TB. In each dedupe store, VDR reserves 5 to 8 GB of space for internal usage. Thus, we highly recommend that you allocate at least 50 GB for your dedupe stores to leave enough space for backup data.

Each VDR Appliance can protect up to 100 VMs. Depending on the number of VMs you are protecting, their space requirements and your retention policy, the space requirements in the dedupe stores can be different. As a general rule-of-thumb, we recommend that you

leave as much space in the dedupe stores as the sum of the amount of space used on all the virtual hard disks of your protected VMs. For instance, you should have a dedupe store of at least 100 GB if you are protecting 10 VMs with one 20 GB virtual disk each, when the virtual disks are on average 50% full.

Start Sufficient then Grow Incrementally as Needed

The statistics that we gathered from our lab indicate that the space requirements will plateau (i.e. remain more or less the same) after some time as long as there are no more additions of new protected VMs. Eventually, the amount of space freed by exercising the retention policy will balance the amount of new data added to the dedupe store during backups. If you are concerned about over-provisioning your dedupe store you can start with barely sufficient space in your dedupe stores and grow as needed, leveraging the VDR capability to **extend** the dedupe stores. In this release, VDR requires that you unmount the dedupe store, extend it in the storage configuration tab of vSphere Client, extend it in the VDR plug-in interface, then mount it again.

We recommend that you grow by increments (e.g. 10% of the original size) each time since that increment may be all you need to accommodate all the backups thereafter.

Note that VMware does not support shrinking of dedupe stores.

Thin Provisioning vs. Thick Provisioning

A VDR dedupe store can reside on either a thin-provisioned or a thick-provisioned disk. One difference between the two provisioning methods of VMDKs is that there may be performance impact resulting from growing a thin-provisioned disk. Since the space requirement of a dedupe store plateaus after a certain threshold, we recommend that you use thick-provisioned disk sized at the threshold to avoid the potential performance impact from growing a thin-provisioned disk. If you run out of space on your dedupe store you can grow it by extending the disk it resides on using the VI Client (as described in last section).

Dedupe Store Format

We highly recommend that you use virtual disks (VMDKs) or RDMs for dedupe stores since the performance behavior is well-understood and consistent. CIFS, on the other hand, gives varying performance characteristics across CIFS providers. In many cases, our lab statistics show that virtual disks and RDMs yield better performance than network-based dedupe stores.

Dedupe stores can be stored in RDM with either virtual or physical compatibility. If you plan to save the dedupe store to tape by taking snapshots, you will want to use RDM with virtual compatibility. Snapshots cannot be taken with RDM with physical compatibility.

Maximize the Number of VMs per Dedupe Store for Space Efficiency

You lose space efficiency by striping VMs across dedupe stores. You can save more space if you put all 100 VMs in one dedupe store of 1 TB than if you put 50 VMs in one 500 GB dedupe store and another 50 in another 500 GB dedupe store.

Deduplication as a standard feature in VDR

Deduplication is a standard feature in VDR. Your protected data are deduplicated automatically in the dedupe stores. The dedupe stores can be stored in all HCL storage and CIFS based network shares, and they are compatible with storage that is capable of deduplication.