Configuring EMC Data Domain Boost with Veeam Availability Suite v8

Rick Vanover
VMware vExpert, MCP, MCITP, MCSA and Veeam Product Strategy Specialist
Introduction
Support for EMC Data Domain Boost is included for backups leveraging Veeam® Availability Suite™ v8. This tech note will walk through the configuration points associated with adding an EMC Data Domain target for backup jobs. This document is written from an early internal preview and is subject to change upon the final release of Veeam Availability Suite v8.

EMC Data Domain Boost
For organizations dealing with data sprawl, disk-based deduplication will be the answer for managing overall storage when it comes to data protection and archiving. EMC Data Domain deduplicating storage systems are a leader in this space and provide many options to address ways of protecting data.

EMC Data Domain Boost exists in two principle ways: For Enterprise Applications and for Backup Applications. Organizations with enterprise applications like SAP, IBM DB2, Oracle, SQL Server and others can leverage EMC Data Domain Boost to write data to the appliance for protection. Organizations can also take advantage of backup application support with EMC Data Domain Boost, including the recently announced support in Veeam Availability Suite v8 for backups of both Microsoft Hyper-V and VMware vSphere virtual machines (VMs).

Backup applications that complement the EMC Data Domain Boost ecosystem strengthen the investment in a deduplicating storage system by increasing performance for data protection. In the case of Veeam, EMC Data Domain Boost is supported over Local Area Networks (LANs) and Storage Area Networks (SANs), giving options for many different storage arrangements for virtualized environments. The diagram below (Figure 1) shows how EMC Data Domain Systems provide protection for backup and archiving.

Figure 1: EMC Data Domain provides backup and archive storage for many use cases.
Configuring EMC Data Domain Boost with Veeam Availability Suite

Within Veeam Availability Suite v8, the EMC Data Domain Boost is supported with a backup repository created in the new wizard with awareness of the new repository types. A backup repository is a storage location used to keep backup files, copies of VMs, file copy tasks and optionally replicated VM metadata.

**Note:** The screenshots of Veeam Availability Suite v8 in this document are from an early internal preview and are subject to change.

There are a few steps to configure backup repository with EMC Data Domain Boost. The following checklist includes the required steps which will be explained subsequently:

- Configure a DD Boost Storage Unit
- Configure a Veeam backup repository
- Configure Veeam backup job(s)
- (optional) Configure Veeam Backup Copy jobs

This document will not cover initial configuration of the EMC Data Domain deduplicating storage system or Veeam Availability Suite. This includes enabling the DD Boost feature in the appliance and installing Veeam Enterprise or Enterprise Plus licensing.

Configure a DD Boost Storage Unit

A DD Boost Storage Unit is the local object that will become a target for the Veeam backup jobs. Within the Data Domain server administrative console, navigate to: Data Management | DD Boost | Storage Units and click *Add*, as shown in the figure below (Figure 2).

![Figure 2: A DD Boost Storage Unit is the first step in creating a Veeam repository.](image)
In the wizard, a user is selected for the DD Boost Storage Unit, which will be used in the Veeam repository creation wizard. In this example, the DD Boost Storage Unit is called v8-ddboost-sample and the user is sysadmin.

Configure a Veeam backup repository

Once the DD Boost Storage Unit is created on the appliance, the next step is to create a repository in Veeam Availability Suite that will use this as a backup target. To create an EMC Data Domain Boost-enabled backup repository, navigate to the Backup Infrastructure section of the user interface, then select Backup Repositories and right-click to select Add Backup Repository. This step is shown below (Figure 3).

![Image of the backup repository wizard](image)

*Figure 3: The backup repository wizard will start to leverage EMC Data Domain Boost.*

The next step is to select the repository type, Deduplicating storage appliance as shown below (Figure 4).
When the actual Data Domain system is selected in this step, a few important options are presented:

- **System name**
- **Fibre Channel connectivity**
- **Gateway server**

The system name is the DNS name or TCP/IP address of the Data Domain server that will house the DD Boost Storage Unit connected in the previous steps. The communication that leverages EMC Data Domain Boost can go over the LAN or SAN as indicated earlier. If Fibre Channel connectivity is to be used, select the option to leverage the SAN. This requires that the SAN be correctly zoned for the Veeam gateway server to access the storage targets.

A critical decision in this wizard is the selection of a gateway server. The gateway server in all situations should be placed from a network perspective as close as possible (from a latency perspective) to the actual Data Domain server. This is because the gateway server will send the traffic to the Data Domain server and communicate with the DD Boost Storage Unit. Backup (or Backup Copy job) data flow may originate in other sites or networks, but the gateway server role for the Data Domain server should be very close to the appliance.

Additionally the credentials set in the configuration phase of the DD Boost Storage Unit are specified here in the backup repository wizard (Figure 5).
Configuring EMC Data Domain Boost with Veeam Availability Suite v8

Figure 5: The Data Domain system name and credentials are applied in this step.

If you need the Data Domain server name on a Fibre Channel network, it is in the Data Management | DD Boost | Fibre Channel section of the administrative interface. This section is shown below (Figure 6).

Figure 6: The Fibre Channel name is defined in the Data Domain server administrative interface.

Once the Data Domain server, credentials and gateway server are selected, the DD Boost Storage Unit is selected while creating the backup repository. Note how the path to the repository will call the EMC Data Domain booth through the libraries that are present via the integration between Veeam Availability Suite and the Data Domain server in the figure below (Figure 7).
After completing the remaining steps of the backup repository (which are not specific to DD Boost), the new repository will be ready for use in a backup job.

**Configure Veeam backup jobs**

The final step will be to configure backup jobs to land on the newly created backup repository that supports EMC Data Domain Boost. The critical decision on backup jobs will be whether to do an active full backup or leverage synthetic full backups. For most environments, it is recommended to do synthetic full backups when leveraging EMC Data Domain Boost. This will save stress on primary storage for the vSphere and Hyper-V VMs and the Boost-enabled synthesizing is very fast.

The synthetic full backup is an optional configuration in Veeam backup jobs, which is synthesized from a full backup already on the Data Domain server. Veeam will access the previous full backup file along with a chain of subsequent increments on the backup repository, consolidate the VM data from these files, and write the consolidated data to a new full backup file. All of this I/O is kept off of primary storage and entirely on the Data Domain server. Configuring the synthetic full backup is shown below (Figure 8).
Configuring EMC Data Domain Boost with Veeam Availability Suite v8

Figure 8: Synthetic full backups are created via the Advanced Settings button of a backup job.

With EMC Data Domain Boost on synthetic full backups, there are benefits to be seen in the time it takes to create these full backups. In the preview edition, the following performance benefits were seen on a test VM. The backup job on the left (with DD Boost) took 11 minutes and 30 seconds for a synthetic full backup. The backup job on the right (without DD Boost) took 1 hour 27 minutes and 4 seconds to synthesize the same VM to a full backup. The highlighted sections of the user interface are shown below (Figure 9).

Figure 9: The synthetic full backups are faster when Data Domain Boost is enabled for a repository.
Configuring EMC Data Domain Boost with Veeam Availability Suite v8

Note: All virtualized environments are different. These numbers will likely differ in many ways that what you may encounter in your environment.

Configure Veeam Backup Copy jobs

If a Backup Copy job is going to be used and GFS retention (Monthly, Weekly, Quarterly and/or Annual restore points) implemented as well, the first consideration is to ensure that the gateway server is closest to the Data Domain server, since the Backup Copy job frequently involves an offsite transfer. When the Data Domain server is designated in the repository setup, ensure that consideration is given to the gateway server if it is being used off site.

GFS transformations can leverage EMC Data Domain Boost for quicker processing during synthetic full backup file creation. In this example, a transformation in the Backup Copy job with EMC Data Domain Boost takes 8 minutes and 23 seconds (as shown on the left). Without EMC Data Domain Boost, the same VM in a Backup Copy job synthesized in 1 hour 25 minutes and 48 seconds (as shown on the right) (Figure 10).

Figure 10: The GFS transformations are quicker when EMC Data Domain Boost is leveraged.

Note: All virtualized environments are different. These numbers will likely differ in many ways than what you may encounter in your environment.

Additional benefits of DD Boost with Veeam Availability Suite v8

There are supplementary benefits to using DD Boost with Veeam Availability Suite v8 that provide performance and resiliency improvements in other areas of the VM backup process. While the above scenarios will apply to many use cases, there are additional benefits that can be realized.
The first of which applies to VMs that may have multiple backup jobs and will land on the same DD Boost Storage Unit. The first time that VM is backed up with the first of the two (or more) jobs, it will function like a full backup as expected. The next time that VM is backed up with the second job, it will have to perform a full backup as it is the first time the Veeam job is called. Because the first job has the bulk of the blocks of the vSphere or Hyper-V VM on the DD Boost Storage Unit, it will only need to transfer metadata and any possible changed blocks. This can be a significant improvement on the active full backup process when there is a fast source storage resource in place.

Additionally when using EMC Data Domain Boost, there is Advanced Load Balancing and Link Failover for connectivity to the Data Domain server. When multiple network interfaces are in use, the Data Domain server will distribute connections from Veeam components over these links. This provides improved data transfer performance with Veeam jobs operating in parallel and additional resiliency by transparently switching failed links to active links if there is an interruption in connectivity.

**Configuration Summary**

Configuring EMC Data Domain Boost with Veeam Availability Suite v8 is easy. This document provides a tour of the configuration steps as well as where to find performance benefits. Additionally, Veeam and EMC customers can leverage these additional resources for more material:

- Veeam Forums [http://forums.veeam.com](http://forums.veeam.com)
- Veeam social media @Veeam
- EMC PowerLink [https://powerlink.emc.com/](https://powerlink.emc.com/)
About the Author

Rick Vanover (vExpert, MCITP, VCP) is a product strategy specialist for Veeam Software, based in Columbus, Ohio. Rick is a popular blogger, podcaster and active member of the virtualization community. Rick’s IT experience includes system administration and IT management, with virtualization being the central theme of his career recently.

Follow Rick on Twitter @RickVanover or @Veeam and on Google+

About Veeam Software

Veeam® enables the Always-On Business™ by providing solutions that deliver Availability for the Modern Data Center™, which provides recovery time and point objectives (RTPO™) of less than 15 minutes for all applications and data. Veeam recognizes the challenges in keeping a business up and running at all times and addresses them with solutions that provide high-speed recovery, data loss avoidance, verified protection, risk mitigation and complete visibility. Veeam Backup & Replication™ leverages technologies that enable the modern data center, including VMware vSphere, Microsoft Hyper-V, NetApp storage, and HP 3PAR StoreServ and StoreVirtual Storage, to help organizations meet RTPOs, save time, mitigate risks, and dramatically reduce capital and operational costs. Veeam Availability Suite™ provides all of the benefits and features of Veeam Backup & Replication along with advanced monitoring, reporting and capacity planning for the backup infrastructure. Veeam Management Pack™ extends Microsoft System Center monitoring to enterprise vSphere environments and also offers monitoring, reporting and capacity planning for the Veeam Backup & Replication infrastructure. The Veeam Cloud Provider Program (VCP) offers flexible monthly and perpetual licensing to meet the needs of hosting, managed service and cloud service providers. VCP currently has more than 4,500 service provider partners worldwide. Monthly rental is available in more than 70 countries from more than 50 Veeam aggregators.

Founded in 2006, Veeam currently has 25,000 ProPartners and more than 101,000 customers worldwide. Veeam’s global headquarters are located in Baar, Switzerland, and the company has offices throughout the world. To learn more, visit http://www.veeam.com.
To learn more, visit http://www.veeam.com/backup