# Best Practices

## Veeam Backup & Replication for VMware

## Version 8.0

**September, 2015**

VEEAM
IT JUST WORKS!™

Microsoft Partner
Gold Application Development
Gold Datacenter

vmware PARTNER
TECHNOLOGY ALLIANCE

# Authors of this book

**Preben Berg**

**Andreas Neufert**

**Tom Sightler**

**Pascal Di Marco**

**Important!**

Please read the End User Software License Agreement before using the accompanying software program(s). Using any part of the software indicates that you accept the terms of the End User Software License Agreement.

# Contents

# CONTACTING VEEAM SOFTWARE

At Veeam Software we value the feedback from our customers. It is important not only to help you quickly with technical issues, but it is our mission to listen to your input, and build products that incorporate your suggestions.

## Customer Support

Should you have a technical concern, suggestion or question, please visit our Customer Center Portal at cp.veeam.com to open a case, search our knowledge base, reference documentation, manage your license or obtain the latest product release.

## Online Support

If you have any questions about Veeam solutions, you can use the following resources:

- Full documentation set at

  www.veeam.com/documentation-guides-datasheets.html

- Community forum at forums.veeam.com

## Company Contacts

For the most up-to-date information about company contacts and office locations, please visit www.veeam.com/contacts.html.

This page intentionally left blank

# ABOUT THIS GUIDE

This document explains Veeam Backup & Replication for VMware infrastructure, components and processes, and describes deployment and configuration options and the impact of their choices. It applies to version 8.0 and all subsequent versions until it is replaced with a new release.

## Intended Audience

This document is primarily aimed at solution architects, consultants, administrators and other IT professionals involved in deployment planning and implementation. At least intermediate level of VMware virtual infrastructure knowledge and advanced understanding (Veeam VMCE level) of Veeam Backup & Replication and Veeam ONE are required.

This document was written for Medium and Enterprise customer virtualization environments. All used examples and sizing recommendations reflect that. However, the described solutions and architecture details can be used in smaller environments and branch offices as well. In most cases, less resources are needed in SMB environments. For those environments, you can use the User Guide, Evaluator's Guide and Release Notes to find the background information.

The complete set of documentation for Veeam Backup & Replication can be found on the Veeam Help Center page at http://www.veeam.com/documentation-guides-datasheets.html.

## Document Revision History

| Revision # | Date | Description of Changes |
|---|---|---|
| Revision 1 | 26/08/2015 | Initial version of the document for Veeam Backup & Replication 8.0. |

# INTRODUCTION

## Veeam Availability Suite 8.0

Veeam introduces a great number of new features and enhancements with every release, patch and update. In particular, the latest version of Veeam Availability Suite™ combines industry leading backup, restore and replication capabilities of Veeam Backup & Replication™ with the advanced monitoring, reporting and capacity planning functionality of Veeam ONE™. You can find more information on the latest updates and enhancements, following the links below. Please note that this guide is focused on the VMware platform; if you are interested in Hyper-V environment, refer to the product documents in Veeam Help Center at http://www.veeam.com/documentation-guides-datasheets.html.

**Veeam Backup & Replication**

- What`s New for v8:

  http://www.veeam.com/veeam_backup_8_whats_new_en_wn.pdf

- Patch 1 enhancements: http://www.veeam.com/kb1982
- Update 2 enhancements (Update is a new name of Veeam patches): http://www.veeam.com/kb2024

**Veeam ONE**

- What´s new v8:

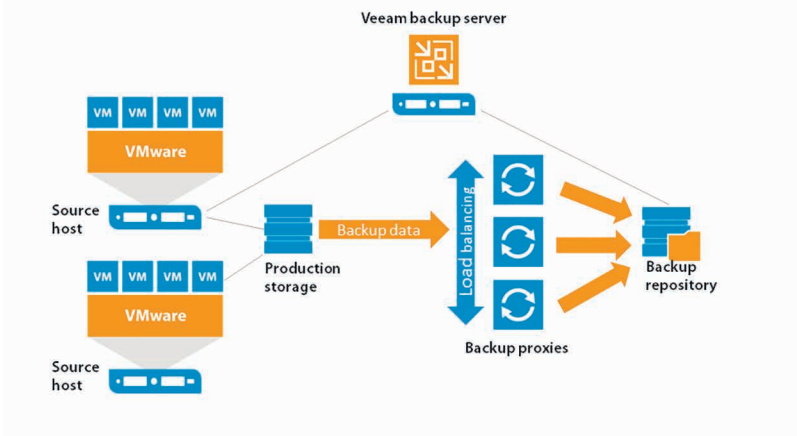
  http://www.veeam.com/veeam_one_8_whats_new_en_wn.pdf

- Patch 1 enhancements: http://www.veeam.com/kb1986
- Update 2 enhancements: http://www.veeam.com/kb2025

If you want to upgrade the existing environment, check Release Notes for the corresponding product version – they include detailed upgrade instructions.

# COMPONENTS OVERVIEW

This section provides a brief overview of the Veeam Backup & Replication solution architecture, as readers are expected to have a decent understanding on what makes up a Veeam Backup & Replication deployment.




As shown in the diagram, Veeam Backup & Replication comprises the following key components:

- Veeam backup server – this server is responsible for job management and scheduling, indexing tasks, and general orchestration of the backup and replication environment.
- Backup proxy - these servers read data from the VM snapshots, deduplicate and compress that data, and send it on its way. In the case of replication, they also receive the replica data and write it to the new replica, acting as the data movers to transfer data from the source to the target environment.
- Backup repository – stores backup images for future restores, and important meta-data used during backup and replication. A repository may be a Windows or Linux server or a NAS device which supports CIFS access.

Also, the following optional components can be deployed:

- WAN Accelerator – this component is intended to help you send less data over the network. To reduce the amount of data going over the WAN, Veeam Backup & Replication uses the global data deduplication mechanism.
- Veeam Backup Enterprise Manager – this component is intended to simplify the daily management and administration

of Veeam backup servers via a single web-based UI. It is typically deployed when there are multiple Veeam backup consoles/sites to manage.

- Veeam Backup Search - this component is installed on a dedicated Microsoft Search Server in case you need to search through VM guest OS files search in large-scale virtual deployments, using sophisticated search criteria. Once you find a necessary file, you can use Veeam's file-level restore to recover the file from the VM backup.

# RESOURCE PLANNING

## DNS and Name Resolution

Name resolution is critical for Veeam Backup & Replication deployment and configuration. Make sure that all VMware components can be reached by both forward and reverse name resolution, otherwise some Veeam components will not be able to work correctly.

If the name resolution system is not available, you can add VMware servers to the local hosts file on all Veeam component servers. Add short and fully qualified domain names in order to increase reliability.

Make sure that fully qualified domain names (FQDN) and domain name servers (DNS) on ESXi hosts are configured correctly. You may alter host file on Veeam proxy serves for ESXi host name resolution to bind Veeam datastreams for NBD mode to the required VMKernel interfaces.

## Veeam Backup Server

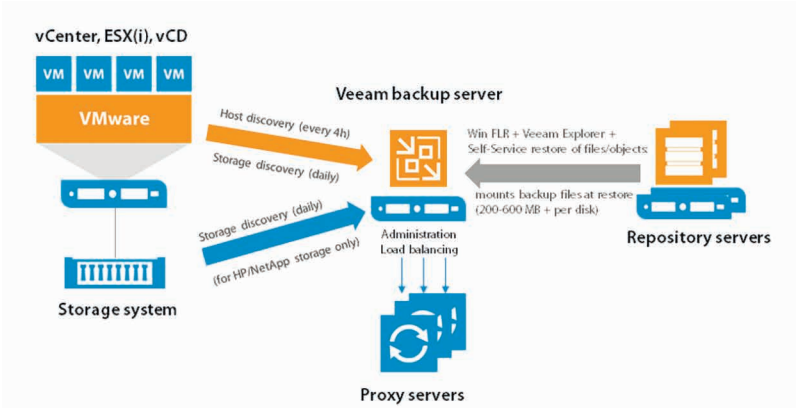
### Planning for Data Streams to Be Processed

The placement of Veeam backup server is likely to depend on the several data streams it will process, in particular:

1. Host discovery and storage discovery
2. Data flow at restore, including:

- Windows file-level restore process
- Application item restore (via Veeam Explorer for Microsoft Active Directory, Veeam Explorer for Microsoft Exchange, Veeam Explorer for Microsoft SharePoint, Veeam Explorer for Microsoft SQL)
- Self-Service Restore of files and application objects (via Veeam Backup Enterprise Manager)

3. Disaster recovery optimization for replica failover and/or for possible WAN link downtime




# Host and Storage Discovery

The Veeam backup server periodically scans all managed hosts and storage systems in order to detect new datastores and newly added VMware cluster members (ESXi hosts). Host scan runs every 4 hours, storage scan runs daily. As a result of the scans, Veeam updates VMware metadata stored in the Veeam Backup & Replication configuration database.

**Note**

If there is a single Veeam backup server deployed on the same site with vCenter Servers and VMware hosts, there are no special considerations on data flow. However, if there are multiple Veeam backup servers or if the deployment involves WAN links, you should plan appropriately for network traffic so that it does not affect the deployment.

When adding a vCenter Server to Veeam Backup & Replication management console, consider that all its managed ESXi hosts will be scanned automatically every 4 hours. So, in case of a RO/BO deployment with 100 branches, each with its own Veeam backup server, adding a central (global) vCenter Server can produce a significant traffic over the WAN links due to these periodic scans. To avoid that, the following can be suggested:

- Create a vCenter Server user account that can only access the ESXi hosts of a specific branch office. This account can be used by the local Veeam backup server to limit traffic to the local ESXi hosts when scanning.

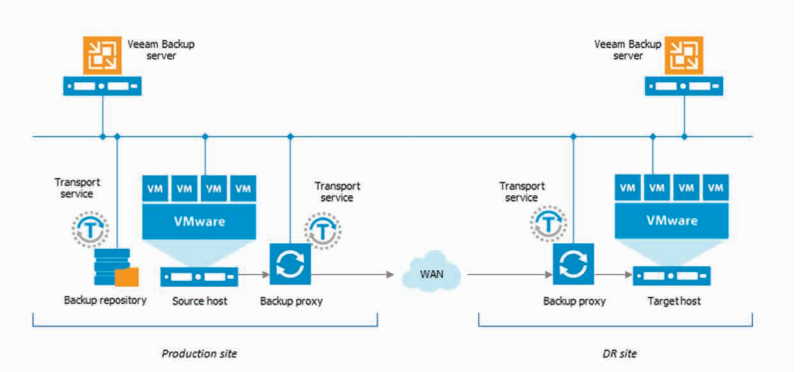
- If a site comprises standalone ESXi hosts, they can be added to Veeam Backup & Replication management console separately, as managed hosts (instead of adding the vCenter Server).

**Note**

Avoid adding individual hosts to the backup infrastructure if using shared storage in VMware vSphere.

# Disaster Recovery Optimization

Best practices recommend to deploy a Veeam backup server on the target site where replicated VMs will reside. Therefore, if you plan to perform VM replication with Veeam, consider placing a Veeam backup server at the replica target site. A backup server deployed in the DR site guarantees correct one-click failover during disaster recovery event. If it is not possible to deploy a Veeam backup server in the DR site, install Veeam Backup & Replication on a VM and replicate this VM to DR site. Then in the case of a disaster recovery event, you will be able to boot this replica VM using VMware vSphere Web Client and perform the necessary failover operations.




## Example 1: Enterprise Deployment for 50 Remote Offices, with Central Job Management

An organization comprises 50 branches, with ESXi hosts on each site, and a vCenter Server in the central location. IT require a central management point for the Veeam backup infrastructure, administration and job scheduling. They require local backups to be created at branch sites, and a backup copy job (with WAN acceleration) to consolidate these at Global HQ. An administrator can implement this example scenario in the following way:

1. Install Veeam Backup & Replication in the branch office (ROBO) and the datacenter (HQ), either on a virtual or physical server. (Do not configure WAN accelerators).

2. Add the branch office (ROBO) ESXi hosts to the corresponding (branch office) Veeam backup servers to provide fast local restore.

3. Install Veeam Backup & Replication at Global datacenter HQ and add all branch office (ROBO) Veeam backup servers as managed servers with proxy, repository and WAN accelerator.

**Note**

Remember to add appropriate resources so that all 3 roles can run on this server.

4. Configure one or more repository servers and WAN accelerators at Global datacenter (HQ) as described in the documentation.

5. Add the vCenter Server to the Veeam Backup & Replication console at Global datacenter (HQ).

6. Configure backup and backup copy jobs on Veeam backup server at Global datacenter (HQ) for all branch offices (ROBO).

This deployment scenario facilitates centralized administration via Veeam backup console at HQ.

## Constraints

However, the following considerations should be taken into account:

- Use local Veeam backup servers to rescan/import backups for file-level restore and item-level restore via Veeam Explorers.

- Enterprise Manager can be used with all of its functionality, but Enterprise Manager Self-Service file restore functions and object restore cannot be used, as it would mount the whole backup file to the central Veeam backup server through the WAN link and also transport the restore data stream twice through the WAN link. This does not affect 1-Click VM Restore/failover and job administration by Enterprise Manager.

- If a WAN link fails, no backup job will run, as the backup server will not be able to connect to the branch office Veeam and VMware components to start required tasks.

- As components are managed by multiple backup servers, always ensure that the same patch/update/version level is used for the entire Veeam backup infrastructure.

## Example 2: Enterprise Deployment for 50 Remote Offices, with Independent Job Scheduling and Self-Services

An organization features 50 branches, with ESXi hosts on each site, and a vCenter Server in the central location. They require VM backups and Self-Service Restore to function without data flowing through a WAN link for restore, as the WAN bandwidth is very low. They require

local backups to be created at branch sites, and a backup copy job (with WAN acceleration to save bandwidth) to consolidate these at Global HQ.

An administrator can implement this example scenario, as follows:

1. Install an Enterprise Manager on the Global HQ and add all Backup & Replication Servers later.
2. Install Veeam Backup & Replication on each site, either on a virtual or physical server.
3. Configure WAN accelerators as described to the documentation.
4. Add the branch office ESXi hosts to the corresponding (local) Veeam backup servers to provide fast local restore.
5. Install Veeam Backup & Replication at Global HQ for import and restore of backups that are copied to the Global HQ using the backup copy jobs.
6. Configure backup and backup copy jobs on Veeam backup servers in the branch offices.

This deployment scenario facilitates VM backups at local sites, as well as Enterprise Manager-based Self-Service Restore. Windows VM Restore File Level Recovery and application item restore via Veeam Explorers can be done without a need to rescan/import the backups directly at the branch office Backup & Replication servers. You can use multiple Veeam Backup & Replication management consoles for independent job scheduling.

**Note**

As components are managed by multiple backup servers, always ensure that the same patch/update/version level is used for the entire Veeam backup infrastructure.

# Data Flow at Restore

## Windows File-Level Recovery

To perform file-level restores for a Windows-based VM, Veeam mounts all VM disk files from the backup files (stored on the repository server) to the Veeam backup server. This can produce a significant data flow between those components (50 - 400MB) – consider this factor if there is a WAN link here.

When the backup is mounted, the files selected for recovery are transported from the repository server through the Veeam backup server to the VM itself.

To optimize file-level and item-level restores from backup repositories in remote locations, you can deploy a backup server and import backups manually as mentioned in the "Disaster Recovery Optimization" section above.

### Veeam Explorers

Veeam Explorers are installed together with the Veeam backup server and run the Windows file-level restore process in the background. So, if you plan to use these tools, refer to the previous section for server deployment considerations.

### Self-Service File, Item and Database Restore via Enterprise Manager

The Self-Service restore process uses Veeam backup servers to mount the backup file from the repository-based primary backups to itself or to the Linux FLR appliance (for Linux file restore only). Thus, it is recommended to place a Veeam backup server on the same site as the repository where the backup files that will be used for restore are stored.

## Physical or Virtual?

Veeam Backup & Replication server can be deployed on a physical or virtual server.

- In most cases, when the Veeam backup server runs on a VM, it is replicated to a secondary location or a DR site by leveraging a separate VMware environment. In the case of a datacenter availability failure/event, this VM can be powered on via the VMware vSphere Client to ensure proper 1-click failover.
- If installed on a physical machine, the Veeam backup server runs independently from the VMware platform, which may also be an ideal solution in case of a disaster even within VMware environment.

In an enterprise environment, it is recommended to install an additional Veeam backup server to speed up the failover process when it is needed (it can be co-located with one of the Veeam proxy or repository servers). If the Veeam Backup Service has to be restored independently of the Veeam backup server, then back up the Veeam Backup & Replication configuration and keep a copy in the designated location.

# Veeam Backup Server Sizing

## CPU and Memory

Recommended Veeam backup server configuration should have 1 Core or 1 vCPU and 4 GB RAM per 10 concurrent running jobs, including any started backup or replication jobs, as well as backup copy and tape jobs running in the background.

The minimum requirement is 2 cores or 2 vCPUs.

**Note**

Best practice is to add multiple servers to a single job. For more information, refer to the job sizing recommendations later in this guide. Additional resources will be required for the Microsoft SQL Server database, backup proxy, backup repository, WAN accelerator, Enterprise Manager and other roles. For their requirements, see the corresponding sections of the Veeam Backup & Replication User Guide and this document.

## Operating System

Veeam backup server requires Microsoft Windows 2008 R2 or later. The latest version of Windows OS is recommended (currently, Microsoft Windows 2012R2 ) as it will support additional options at restore (e.g., Windows file-level restore from virtual machines with Windows Server Deduplication enabled).

For the full list of supported OS, refer to Veeam Backup & Replication Release Notes.

## Disk Space

This section explains what folders you should plan for when preparing for installation of the Veeam backup server.

The folders are detailed here as follows:

1. Installation folder - default location is *C:\Program Files\Veeam\Backup and Replication*. Plan for 40 GB (if installing on a VM, thin disks can be used).

2. Application log files – default location is *%allusersprofile%\Veeam\Backup*, which in most cases is located at *C:\ProgramData\Veeam\Backup*.

   Log file size mainly depends on the number and frequency of jobs and on the VM count. Also, consider that the logging level may also affect the log size. If you need to change the logging level or log file location, refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1825. In enterprise environments, it is recommended to change the default location to a separated disk.

3. Veeam Backup Catalog folder – default location is *C:\VBRCatalog*.

   This folder is used if VM guest indexing in backup jobs is enabled. For more information, refer to the "Indexing" section of this guide. To change the default location, refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1453.

4. Veeam vPower NFS folder for NFS Service – default location is *C:\ProgramData\Veeam\Backup\NfsDatastore*.

   If you start a VM with Instant VM Recovery, this folder is used by default to store all configuration files and data change of the running VM. To offload the storing of these changes to a specific datastore, use the Instant VM Recovery wizard. Best practices recommend to implement vPower NFS Services on each Windows-based backup repository. For CIFS shares, it is recommended to locate it on a proxy gateway, and for Linux repositories — on a Windows machine next to the Linux repository (remember to use fast and low-latency network connections).

   It is recommended to reserve at least 10 GB space for this folder. If you plan to start a significant number of VMs or run VMs over a long time, increase the space according to the produced/estimated block-level changes of the VMs.

The vPower NFS server is bound to backup repositories, and the folder location is defined per server. To change the folder location, go to a repository where you want to use this vPower NFS Server, and modify its properties:

5. In the **Backup Infrastructure** view, click **Backup Repository**.
6. Then right-click the necessary repository, select **Properties** and navigate to the **vPower NFS** step of the wizard.

**Important!**

Make sure the vPower NFS Service is configured correctly on the Veeam backup server itself, as it will be used (if selected) on a repository server. Otherwise, you will be unable to leverage Veeam background functionalities like distribution of Virtual Lab FLR appliances or disk mapping for the Linux FLR appliance.

For information on folders required for Enterprise Manager, backup proxy and repository servers (backup targets) and WAN accelerators, as well as for recommendations on their sizing, please refer to the corresponding sections of this guide.

There are no special sizing considerations for PowerShell SDK and tape proxy server.

## Other Software

It is strongly recommended that no highly-transactional and business-critical software (like Microsoft Active Directory Domain Controller, Exchange server or production databases on the SQL server instance) are deployed on the same machine as the Veeam backup server.

However, you can install the dependancies, like staging Microsoft SQL Server for Veeam Explorer for SQL or for Sharepoint, or Microsoft Outlook (64-bit) for mail export to PST files via Veeam Explorer for Exchange, or Acrobat Reader for reading Veeam documentation.

## Installing Veeam Backup & Replication Updates

New Veeam releases and updates can be installed on the Veeam Backup Enterprise Manager and Veeam backup servers by the setup wizard or by an installation script (so called "silent installation"). For detailed instructions, check the latest release notes.

**Note**

In most cases, Veeam Backup Enterprise Manager should be updated before updating Veeam backup servers.

After installing the update, open the Veeam Backup & Replication management console. The **Update** screen will be displayed, then you can update distributed components on other Veeam managed servers (like proxy and repository servers, Veeam vPower NFS servers, WAN accelerators and tape proxy servers).

**Note**

As Veeam deploys no agents on the virtual machines, you will not need to update any software (agents) on the VMs.

# Veeam Backup & Replication Database

Veeam Backup & Replication stores all information about backup infrastructure, jobs settings, job history, sessions, files and other configuration data in its SQL Server database (default database name is *VeeamBackup*).

When planning for Veeam Backup & Replication deployment, examine a number of options offered for the configuration database setup. To make an appropriate choice, follow the recommendations provided in this section.

## SQL Server: Express Edition or Standard/ Enterprise Edition?

Microsoft SQL Server 2012 Express Edition is included in the Veeam Backup & Replication setup, which is a convenient option; on the other hand, it has several limitations:

- Uses 1 GB of RAM for the SQL Express Server see

  https://msdn.microsoft.com/en-us/library/cc645993(v=SQL.110).aspx#CrossBoxScale

- SQL Express uses only the first CPU. If possible, configure 4 Cores for the CPU. (see https://msdn.microsoft.com/en-us/library/cc645993(v=SQL.110).aspx#CrossBoxScale)

- Database size cannot exceed 10GB. That is enough to handle the environments with 400 – 500 VMs, as the configuration database will likely to be below 10 GB. However, if you plan to back up more VMs, it is recommended that you use Standard or Enterprise Edition of Microsoft SQL Server, otherwise, the Veeam Backup & Replication console and job processing will work slower than needed (due to memory and CPU limitations).

Another consideration relates to disaster recovery plans. By default, Veeam Explorer for Microsoft SharePoint and Veeam Explorer for Microsoft SQL will use a local SQL Server instance as a staging system for database restore (for details, see http://helpcenter.veeam.com/backup/80/explorers/) – and SQL databases in production environment are likely to exceed 10 GB, so Express Edition will not be able to support them.

Thus, it is necessary to use Standard or Enterprise Edition if you plan to do any of the following:

- Process more than 4000 VMs
  (also recommended for 400 – 4000 VMs)

- Use Veeam Explorer for Microsoft SQL Server or Veeam Explorer for Microsoft SharePoint with databases bigger than 10 GB. If using any special features of Microsoft SQL Server or Microsoft Sharepoint that are only available with Standard or Enterprise SQL version, you have to use the corresponding version with Veeam.

In other deployment cases, accepting the limitations, it is possible to use Microsoft SQL Server Express Edition.

**Tip**

To improve performance of Microsoft SQL Express version, it is recommended that you add multiple cores to the first CPU. Veeam Backup & Replication leverages Microsoft SQL Server 2014 enhancements, so it is recommended, where possible, to update the database server to Microsoft SQL Server 2014 or Microsoft SQL Server Express 2014.

## SQL Server: Local or Remote?

If you plan to use Veeam Explorer for Microsoft SQL and/or Veeam Explorer for Microsoft SharePoint, you may need a local Microsoft SQL Server instance as a staging system. However, the Veeam Backup & Replication configuration database can be placed on another Microsoft SQL Server instance. Such deployment scenarios have the following benefits:

- It can help to increase the products performance if the Veeam backup server is expected to handle a heavy load (for example, if the backup server is used as a backup proxy and/or repository).
- SQL Clustering and AlwaysOn Availability Group on external SQL Servers can be used for high availability.
- If the virtual environment includes certain ESXi servers or VMware clusters licensed for Microsoft SQL Server, you can place the database(s) on those VMs to reduce overall TCO.

## Reaching Optimal Performance: Fast Disks or more RAM/CPU?

Veeam Backup & Replication may consume high amounts of CPU and RAM while processing backup or replication jobs. Therefore, to achieve better performance and load balancing, it is necessary to provide sufficient RAM and CPU resources to Veeam components. If possible, follow these guidelines:

| Number of concurrent running jobs (including enabled backup copy jobs) | Min CPU cores | RAM |
|---|---|---|
| Up to 25 | 2 | 4GB |
| Up to 50 | 4 | 6GB |
| Up to 100 | 8 | 8GB |

If you plan to run more than 100 jobs at the same time, add more ressources.

It is also recommended to place the configuration database on on a fast disk storage subsystem, such as 10, 000 rpm or faster SAS/FC disks. While disk I/O is not critical, lower latency may improve overall UI experience.

# SQL Server Configuration Settings

Veeam Backup & Replication does not require any specific Microsoft SQL Server configuration settings. However, if you plan to use Veeam Explorer for Microsoft SharePoint or Veeam Explorer for SQL Server, consider that since they utilize the local SQL Server as a staging system, the corresponding requirements should be met:

- For Veeam Explorer for Microsoft SharePoint:

  http://helpcenter.veeam.com/backup/80/explorers/index.html?vesp_staging_microsoft_sql_server.html

- For Veeam Explorer for Microsoft SQL Server:

  http://helpcenter.veeam.com/backup/80/explorers/index.html?vesql_systemreqs.html

Enable and configure all features used by the production database. Consider using the highest license level and latest version/patch level installed in the whole production environment.

If you plan to restore encrypted databases with Veeam Explorer for Microsoft SQL Server or Veeam Explorer for SharePoint, you will need a valid encryption certificate on the local Microosft SQL Server used as a staging system. For details, see this Veeam Knowledge Base article: http://www.veeam.com/kb2006.

Some installations of Microsoft SQL Server 2012 (and later) do not set the rights for log file processing by default, so make sure that the **sysadmin** server role is enabled for the Microsoft SQL Server. For details, see this Veeam Knowledge Base article: http://www.veeam.com/kb1978 Other prerequisites for SQL Server utilized by Veeam backup server can be found at http://helpcenter.veeam.com/backup/80/vsphere/index.html?system_requirements.html.

**Tip**

Follow Microsoft general recommendations for optimal SQL performance. For example, place the SQL tempdb on the fastest disks for best performance: http://blogs.msdn.com/b/cindygross/archive/2009/11/20/compilation-of-sql-server-tempdb-io-best-practices.aspx

## Modifying Database Connection Settings

To modify database connection settings or connect to another Veeam configuration database, use the **DBConfig** utility, as described in the product documentation at http://helpcenter.veeam.com/backup/80/vsphere/index.html?dbconfig_utility.html.

- If using SQL authentication, consider that all Veeam UI and Veeam PowerShell changes are communicated using this authentication.
- If using Windows Authentication for the Microsoft SQL Server connection, ensure the actual user of the Veeam UI or Veeam PowerShell has proper rights to the database, as described in the Veeam Knowledge Base article:

  http://www.veeam.com/kb1853.

## Migrating Veeam Database to Another SQL Server

To migrate Veeam configuration database to another SQL Server, follow the recommendations provided in these Veeam Knowledge Base articles:

- http://www.veeam.com/kb1448
- http://www.veeam.com/kb1250

# Protecting Veeam Backup & Replication Configuration

## Protecting Job Settings

As recommended by best practice for disaster recovery, you can place Veeam Backup & Replication installation on a virtual machine and protect it with backups or replicas. Out-of-the box, Veeam automatically

creates configuration backups on the default backup repository. These configuration backups contain all the information about Veeam Backup & Replication jobs (passwords are not stored by default) and can be used to restore all jobs and their metadata (you will be asked for all required passwords during the restore process). Please refer to the Veeam Backup & Replication User Guide for further details:

http://helpcenter.veeam.com/backup/80/free/vbr_config.html

**Tip**

If encryption is enabled for configuration backup, passwords are also stored in the configuration backup files.

# Planning for Disaster Recovery of Veeam Backup Server

If possible, it is recommended to replicate Veeam backup server VM to the secondary site (verify network and IP mapping settings before you begin; refer to http://helpcenter.veeam.com/backup/80/vsphere/index.html?replica_job.html for details). If the server is replicated successfully, then in case of a disaster recovery situation, you may start its replica in the secondary location without having to reinstall Veeam Backup & Replication. This will help to reduce overall Recovery Time Objective (RTO).

If performing a configuration backup, it is also recommended to place one copy of the backup on the secondary site. You can configure another repository for that purpose; use Veeam's File Copy Job to place a copy of the configuration backup at the DR site.

**Note**

Backup copy jobs do not process configuration backups. Remember that configuration backups are not processed with backup to tape jobs; if you want to store configuration backups on tape, use file to tape jobs instead.

Again, since all data required for a restore is directly placed within the backup file (which VMs are in the backup file, as well as deduplication and encryption information), even in the event that configuration database is lost or damaged, you can set up a new Veeam backup server and import the backup files there, or even use the stand-alone "Extract" utility (both a command line and a graphical version are provided). Then you will be able to restore VMs, files and application data without restoring the configuration database.

# Antivirus on Veeam Servers

Antivirus software monitors all 'write' operations on the operating systems, and this also extends to Veeam backup files. Data that is processed by a backup proxy and repository can overload the antivirus system so that it blocks the backup files. This can slow down the backup process or even lead to backup file corruption. To avoid this, it is recommended to add the following items to the list of antivirus exclusions on all Veeam servers (including Veeam backup server, proxy server, repository server, WAN accelerator server, tape server, and others).

## Folders

- *C:\Program Files\Veeam*
- *C:\Program Files(x86)\Veeam*
- *C:\Program Files\Common Files\Veeam*
- *C:\Program Files(x86)\Common Files\Veeam*
- *VBRCatalog* ([**HKLM\SOFTWARE\Veeam\Veeam Backup Catalog**] **CatalogPath** value)
- *NFS* (Configured in each repository, stored in **[HKLM\SOFTWARE\Wow6432Node\Veeam\Veeam NFS] RootFolder** value)
- *C:\VeeamFLR**
- All Veeam repository folders
- All Veeam WAN accelerator folders

## Folder on VM Guest OS (if VSS is used)

*C:\Windows\VeeamVssSupport*

## Files

- VeeamAgent.exe
- VeeamAgent64.exe
- *.vmdk

Some additional exclusions may be also needed.

**Tip**

If the antivirus has a logging or history system, you can review its logs to detect whether it has taken any actions that affected Veeam Backup & Replication operations.

Also, consider that another service or process may be using ports configured for the Veeam vPower NFS Service. To avoid possible issues, it is recommended to stop the Veeam vPower NFS Service if you do not plan to use it; make sure that none of the NFS ports are used by other software (including antivirus systems). For more information, refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1055.

# Veeam Backup Enterprise Manager

## Whether to Deploy?

Veeam Backup Enterprise Manager is intended for centralized reporting and management of multiple Veeam backup servers; it also provides delegated restore and self-service restore capabilities, as well as the ability for users to request Virtual Labs from Veeam backup administrators. Starting with v8, Enterprise Manager is also a part of the data encryption and decryption processes implemented in the Veeam solution. Thus, best practices recommend deploying Enterprise Manager in the following scenarios :

- It is recommended to deploy Enterprise Manager if you are using encryption for backup or backup copy jobs. If you have enabled password loss protection for the connected Veeam backup servers, backup files will be encrypted with an additional private key which is unique for each instance of Enterprise Manager. This will allow Enterprise Manager administrators to unlock backup files using a challenge/response mechanism, effectively acting as a Private Key Infrastructure (PKI).
- If an organization has a Remote Office/Branch Office (ROBO) deployment, leverage Enterprise Manager to provide site administrators with granular restore access via web UI (rather than providing access to Veeam Backup & Replication console).
- In enterprise deployments, delegation capabilities can be used to elevate the 1st line support to perform in-place restores without administrative access.

- For deployments spanning multiple locations with stand-alone instances of Veeam Backup & Replication, Enterprise Manager will be helpful in managing licenses across these instances to ensure compliance.

- Enterprise Manager is required when automation is essential to delivering IT services — to provide access to the Veeam RESTful API.

Remember, however, that if the environment includes a single instance of Veeam Backup & Replication, you may not need to deploy Enterprise Manager, especially if you want to avoid additional SQL Server database activity and server resource consumption (which can be especially important if using SQL Server Express Edition).

**Note**

If Enterprise Manager is not deployed, password loss protection will be unavailable.

# Using Enterprise Manager for Restore Operations

## 1-Click File-level Restore

With Enterprise Manager, you can restore VM guest OS files with a single click. To support this capability, the VM restore point must be created with guest OS indexing enabled; guest file indexing allows you to search for VM guest OS files inside VM backups.

Veeam Backup Catalog on the Enterprise Manager server will be used to store indexing data replicated from the Veeam Backup Catalog on Veeam backup server(s). For more information about the process, refer to the Enterprise Manager User Guide. To learn more about Veeam Backup Catalog sizing, refer to the "Indexing and Search" section of this document.

## 1-Click Application Item-level Restore

Currently, you can restore items from Microsoft Exchange and Microsoft SQL Server with a single click using Veeam Backup Enterprise Manager. These capabilities were developed to elevate the 1st level support engineers, enabling them to recover mail items and other Microsoft Exchange objects, without any direct visibility of the mailbox or database content. Database administrators are now able to restore Microsoft SQL Server databases without addressing the backup team.

## Microsoft Exchange Mailbox Items Restore

The process of restoring an Exchange mailbox is described in the Backup and Restore of Microsoft Exchange Items section of the Veeam Backup Enterprise Manager User Guide.

To create an application-aware image backup of Microsoft Exchange database VM, ensure you back up at least one server holding the Client Access Server (CAS) role (this can be Exchange Server with the Mailbox Database role, or a dedicated server; contact Exchange administrator if necessary). A server holding the CAS role is used to discover the mailbox location for the corresponding user. You should supply credentials for authentication with the CAS server on the **Configuration** > **Settings** page, as described here.

## Microsoft SQL Server Database Restore

To perform database-level restores of SQL Server databases using Enterprise Manager, ensure you enable application-aware image processing for the corresponding backup job. To use point-in-time recovery, enable log file backups of the Microsoft SQL Server VM. For more details, refer to the Backup and Restore of Microsoft SQL Server Databases section of the Veeam Backup Enterprise Manager User Guide.

## Self-Service File Restore

In addition to 1-Click File-Level Restore, Veeam Backup & Replication v8 allows VM administrators to restore files or folders from a VM guest OS, using a browser from within the VM guest OS, without creating specific users or assigning them specific roles at the Veeam Enterprise Manager level. For that, an administrator of the VM can access the self-service web portal using the default URL: https://<enterprise_manager_server_name>:9443/selfrestore.

**Note**

Currently, this feature is available only for the Windows-based VMs and requires Veeam Backup & Replication Enterprise *Plus* license. Also, the VM should be in the same domain with the Enterprise Manager, or in a trusted one (for SID resolution)

The process goes as follows:

1. During the backup of a VM with indexing enabled, Veeam detects users who have local administrator access rights to that machine, and stores this information in the Enterprise Manager database.

2. User enters the self-service web portal URL in the web browser and enters the account name and password to access the necessary VM guest OS.

3. After logging in, the user is presented with the most recent restore point for that VM (the one this user authenticated to) on the **Files** tab of the web portal.

For more information on using this feature, refer to the Self-Restore of VM Guest Files section of the Veeam Backup Enterprise Manager User Guide.

## RESTful API Service

RESTful API service is installed as part of Veeam Backup Enterprise Manager. To provide access to the API, consider that authentication will take place through Enterprise Manager; Enterprise Manager user role assignments (**Portal User**, **Restore Operator**, **Portal Administrator**) and their access scopes access will be inherited by the RESTful API service. For more information on role assignment, see the Configuring Security Settings section of the Veeam Backup Enterprise Manager User Guide.

# Search Server and Veeam Indexing

## Indexing and Search Overview

Veeam Backup & Replication performs backups at the image-level using APIs available from the underlying hypervisor. Thus, it has no direct visibility of the file structure after backup is finished. However, it is possible to use the Veeam File Level Restore wizard to mount VMs from within a backup file and access/restore VM guest files.

If a user wants to perform file restore from the central Enterprise Manager, it is not possible within a acceptable timeframe to mount all backup files and VMs in it to find a file that the Enterprise Manager user wants to restore.

To support advanced file-level restore scenarios, Veeam also offers the capability to index files on VMs being backed up. Indexing is available for both Microsoft Windows and Linux VMs and allows users of Veeam Backup Enterprise Manager to browse and search for the necessary files and to perform one-click file restores. The sections below will outline some specific use cases for indexing and describe best practices and guidelines for sizing.

# How Veeam Indexing Works

Veeam indexing creates a separate catalog file for each restore point. These index files are then used by Veeam Enterprise Manager to support file browsing and search without a need to mount the restore point to the Veeam backup server. Users can quickly search for files across multiple restore points and view required files history when looking for a specific version of a document. They can also select a specific VM and browse the file system to restore guest files.

Veeam Backup Enterprise Manager also allows for file-level restore functions to be delegated to a subset of users by leveraging the role-based access control.

During the VM backup job run, the following operations are performed:

4. Veeam accesses the guest OS (using credentials specified in the job settings) and injects a small run-time process to collect the list of files.

    - For Microsoft Windows-based VMs, the process gathers file metadata by reading the MFT data of the supported file system (NTFS and ReFS).
    - For Linux-based VMs, the process leverages the existing "locate" database that is commonly installed on most Linux distributions.

    These operations take place in parallel with the backup and do not increase the duration of the process. For more details on the indexing process, refer to the Veeam Backup Enterprise Manager User Guide.

1. Veeam Backup & Replication creates a catalog (index) of the VM guest OS files and stores index files on the Veeam backup server in the *C:/VBRCatalog/Index/Machines/{vm_name}* folder. Creation of index is extremely fast and has minimal impact on network and VMware environment.

**Important!**

To search within the index catalog, it is necessary to deploy Veeam Backup Enterprise Manager, as this component is in charge of catalog data replication and retention (see this section of the User Guide for more details). If you enable indexing without configuring Enterprise Manager, the catalog files in the *VBRCatalog* folder of the backup server will never be collected and deleted, and may fill up the disk drive.

1. Once the index is created and stored on Veeam backup servers, the indexing service on Veeam Backup Enterprise Manager performs index replication — it aggregates index data for all VM image backups from multiple backup servers. This consolidated index is stored on the Veeam Backup Enterprise Manager server in the *C:/VBRCatalog/Index/catalog* and is used for search queries.

## When to Use Indexing?

File-level indexing should be enabled only if you plan to utilize advanced file search and one-click file level restore capabilities of Veeam Backup Enterprise Manager (including delegated restore). While indexing is a job-level setting, you can use filters to index only a subset of files; it is also possible to exclude specific VMs from indexing, as described, for example, in this section of the Veeam Backup Enterprise Manager User Guide.

## Sizing Veeam Catalog

Estimated raw space of the final index file is approximately 2 MB per 1,000,000 files for a single VM restore point on the Enterprise Manager server, in the backup files and temporary folders on the Veeam backup server own catalog. During the indexing process, indexing information is temporarily stored on the local VM guest, requiring additional free space on the system drive; estimated free space is about 10 MB per 1,000,000 files.

The Veeam Catalog Service is responsible for maintaining index data. When running on the Veeam backup server, this catalog service will maintain index data for all jobs that run on that specific server as long as the backup data remains on disk. When running on the Enterprise Manager server, the service will replicate index data from all managed Veeam backup servers into the local catalog, so it should be sized appropriately to hold all data from the remote Veeam servers.

- When using a *Standard* license, Enterprise Manager will only keep index data for restore points still in repositories.
- For *Enterprise* and *Enterprise Plus* licenses, you can configure Enterprise Manager to keep indexes even longer, with the default being 3 months. This can significantly increase the amount of space required for the catalog.

### Example

There are two backup jobs configured to process 2 VMs, with 10 000 000 files per each VM. Backup jobs run two times a day, producing 60 restore points a month. The default Enterprise Manager setting is used for catalog retention (3 months).

Space required on the first drive in the VM (C:\ drive) can be calculated as follows:

- For Windows VMs: 100 MB per one million files and directories of all saved restore points with indexing enabled.

**Note**

This was tested with one million randomly named 20-character-long filenames in one directory. Depending on the saved metadata and folder structure of the files, the value can be lower or higher.

- For Linux VMs: 50 MB per one million files and directories of all saved restore points with indexing enabled. Linux indexes require round about 50% less space because **mlocate** does not index any metadata (such as timestamps and ownership information).

Space required on Enterprise Manager server in the catalog folder can be calculated as follows:

- 50 MB * 10 million files * 60 restore points per month * 3 months (for default Enterprise Manager retention). A total of 90 GB per indexed VM with 10,000,000 files.

## Recommended Settings

Follow these recommendations when setting up Veeam indexing:

- Place the catalog on a dedicated volume of good performance disk. To change the default Veeam Catalog folder location, refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1453.
- Consider that by default, Veeam enables NTFS compression on the catalog folder. This can reduce the space requirements by well over 50%. However, for very large catalogs (with 100s of VMs and 10's of millions of files) it can be more beneficial to use a Windows 2012 R2 volume with Data Deduplication enabled. This volume should be dedicated to index files and configured to run deduplication functions outside of the normal backup window.
- It is recommended to enable indexing only on VMs where the advanced search capabilities are necessary. Use filters to exclude unnecessary files from indexing (Windows system folder, Program Files and other system directories are excluded by default). For the Linux systems to be indexed, make sure they have **mlocate** or another compatible **locate** package installed.

**Note**

If you have SUSE Linux Enterprise Server (SLES) 11, consider that indexing will not work, due to no native mlocate package. It is recommended upgrading to SLES 12, which does have native mlocate package, however you may use this openSUSE package http://software.opensuse.org/package/mlocate

- Configure index retention in Veeam Backup Enterprise Manager to the minimum necessary to meet the IT policy requirements. Index retention setting is available in the Enterprise Manager web console under **Configuration > Settings > Guest File System Catalog**.

**Note**

Veeam Backup Enterprise Manager SQL database (*VeeamBackupReporting*) will not grow much while using indexing functions, as this database will only store the corresponding metadata.

- To enhance search performance, SSDs can be used. If you plan to index a very large number of VMs, it is recommended to limit the search scope at restore to a single VM before you click the search button – this will bring faster results.

# Using Veeam Backup Search (Optional Component)

In its early versions, Veeam did not have its own indexing engine; instead, it used a connector named Veeam Backup Search to connect to the Microsoft Search Server 2010 that provided search capabilities. Since then, Veeam's own indexing engine was developed for this purpose.

Currently, Veeam Backup Search is not necessary, as Veeam indexing engine is likely to perform better.

If you need, however, to use that Veeam Backup Search component (and, respectively, Microsoft Search Server) for indexing, consider the following:

- Microsoft Search Server Express Edition can be used, as it has no limitations for the number of indexed files.
- Other editions of Microsoft Search Server deliver higher scalability, because Search Server components can be separately installed on multiple servers. Actually, if you are using Veeam Backup Enterprise Manager, consider that it can spread the load between multiple Microsoft Search Servers Express automatically.

- Microsoft Search Server functionality is used to scan content in the shared *VBRCatalog* folder on the Veeam Backup Enterprise Manager server and to create a content index on the Search Server; this content index is used to process search queries. For more details, refer to the Veeam Backup Search section of the User Guide.

**Note**

Though using content index streamlines the search process, the index itself can require significant space on disk in *C:\VBRCatalog\Journal\[YYYY_MM]\[search-server]*.

- Search Server requires an SQL database for its operation. Consider that Microsoft SQL Server Express Edition leverages only one CPU, which limits the Search Server performance. Besides, the database size supported by this edition is also limited (in particular, 10 GB for Microsoft SQL Server 2008 R2 Express Edition or later).

# Proxy Server and Transport Modes

## Veeam Backup Proxy Overview

With backup proxies, you can easily scale Veeam backup infrastructure based on the organization demands:

- In a basic installation (simple deployment scenario for smaller environments or Proof of Concepts), the backup proxy is automatically installed on the Veeam backup server as part of the Veeam Backup & Replication installation process.

- In other deployment scenarios, the backup proxy role is usually assigned to several Windows servers (64-bit). This approach allows for offloading the Veeam backup server, achieving better performance and a minimized backup window. Backup proxies can be deployed both in the primary site and in remote sites on any managed Microsoft Windows server in the infrastructures. Depending on the data transport mode you plan to use, a backup proxy can be installed on a physical server or on a VM, as explained later in this section.

A backup proxy handles data traffic between the VMware vSphere infrastructure and Veeam Backup & Replication during backup, replication, VM copy, VM migration jobs or VM restore.

Backup proxy operations include the following:

- Retrieving VM data from production storage, compressing and sending it to the backup repository (for a backup job) or another backup proxy (for a replication job).
- A backup proxy is also part of Veeam's deduplication engine; plus, it performs encryption if the corresponding option is selected in the data transportation settings.

Technically, a backup proxy runs a light-weight transport service that takes a few seconds to deploy. When you add a Windows-based server to Veeam backup management console and assign proxy role to it, Veeam Backup & Replication installs the necessary components and starts the required services on that server. When a job is started, the Veeam backup server becomes the "point of control" for dispatching tasks to proxy servers, using its built-in load balancing algorithm.

# Advanced Veeam Technologies for Data Processing

To specify the threshold for proxy load, an administrator uses the **Max concurrent tasks** proxy setting (where a task stands for a single VM disk), and Veeam Backup & Replication uses a unique load balancing algorithm to automatically spread the load across multiple proxies. This feature allows you to increase backup performance, minimize backup time window and optimize data flow. By default, Veeam Backup & Replication analyzes the backup proxy configuration, determines the datastores it can access, and automatically selects the best transport mode depending on the type of connection between the backup proxy and datastore:

First, Veeam Backup & Replication checks if data processing can be assigned to a backup proxy with the Direct SAN Access, then it checks whether a Hot-Add proxy can be used, and then looks for a Network (NBD) proxy. For more details, see the "Transport Modes" section of this guide.

After the algorithm identifies all existing backup proxies, it spreads the load across them in an optimal way:

a. It discovers the number of tasks being processed at the moment by each proxy, and looks for the server with the lowest load and the best connection.

b. All tasks are standing in a "VM to process" queue, and when a proxy's task slot becomes free, Veeam Backup & Replication will automatically fill it up with the next task.

c. Note that priority goes to the disk that belongs to an already processed VM; also, short-term scheduled jobs take priority over long-term scheduled jobs (like daily or weekly jobs).

**Tip**

Default recommended value is **1** task per Core/vCPU. To optimize the backup window, you can slightly overbook the **Max concurrent tasks** count.

Starting with v7, Veeam Backup & Replication supports parallel processing of VMs/VM disks:

- It can process multiple VMs (each with a single disk) simultaneously, increasing data processing efficiency.
- If a VM was created with multiple disks, Veeam will try to process these disks simultaneously to reduce VM snapshot lifetime.

**Note**

Parallel processing is a global setting that is turned on by default.

# Backup Proxy Services and Components

Veeam backup proxy uses the following services and components:

- **Veeam Installer Service** - a service that is installed and started on the Windows server once it is added to the list of managed servers in the Veeam Backup & Replication console. This service analyses the system, installs and upgrades necessary components and services.
- **Veeam Transport Service** – a service responsible for deploying and coordinating executable modules that act as "data movers" and perform main job activities on behalf of Veeam Backup & Replication (communicating with VMware Tools, copying VM files, performing data deduplication and compression, and so on).
- **VeeamAgent.exe process** - a data mover, which can be started multiple times (on demand) for each data stream. These processes work in read-and-write pairs, being started for reading data on a source backup proxy, and for writing data - on a target backup proxy (replication), gateway proxy (for a CIFS repository), or on a repository itself. At restore, the dataflow is in the opposite direction.

# Data Transport Modes for Backup Proxies

Job efficiency and time required for its completion are highly dependent on the data transport mode. Transport mode is a method used by the Veeam proxy to retrieve VM data from the source host and write VM data to the target destination.

The primary role of the backup proxy is to provide an optimal path for backup traffic and to enable efficient data transfer. Therefore, to deploy a backup proxy, you need to analyze the connection between the backup proxy server and storage it will be accessing. Depending on the type of connection, the backup proxy can be configured in one of the following modes (preferred mode is Direct SAN, if possible; all modes will be explained later in more details):

- **Direct SAN mode:** in this mode, backup proxy server has direct access to the storage on which VMs reside. In this case, the backup proxy will retrieve data directly from the storage, bypassing the ESXi infrastructure. Depending on the connection, the proxy can be deployed as follows:
    - On a physical server(FC/SAS/FCoE/iSCSI and, in cases, NFS on NetApp - when Backup from Storage Snapshot feature is enabled and licensed.)
    - On a VM (iSCSI and, for the special case, NFS on NetApp when Backup from Storage - Snapshot feature is enabled and licensed).
- **Network mode:** another option is to use the VMKernel Interfaces of VMware to read and write VM data. As it needs no special configuration and works flawlessly, this mode is also used as a default fail-over option for other modes. This universal transport mode can be used with all storage protocols between primary storage and ESXi host , as it uses the default VMware Storage Stack for read and write through the VMKernel Interfaces by NFC protocol (via TCP port 902). This proxy can be deployed as a physical server or a VM.

**Note**

This backup mode is also known as NBD.

- **Virtual Appliance mode** - to work in this mode, the backup proxy should be deployed as a VM. For smaller deployments (e.g., several branch offices with a single ESXi host per each office), you can deploy a virtual backup proxy on a ESXi host that has access to all required datastores. When backup or replica-

tion takes place and a VM snapshot is processed, the snapshotted disks are mapped to the proxy to read data (at backup) and write data (at restore/replication); later they are unmapped.

**Note**

As the disks are hot-added, you can find the mode's name referred to as **HotAdd** in documentation and logs.

The following sections explain transport modes in detail.

## Network Mode (NBD)

Network mode is by far the easiest backup mode to implement, as it requires no additional configuration. Veeam always uses this mode at least to back up and restore VMware configuration files and to read Change Block Tracking information.

When in this mode, the backup proxy will connect to ESXi hosts on VMkernel interfaces by DNS name resolution and use this connection to transport data, utilizing Veeam's file copy technology (also known as FastSCP). Remember that the backup proxy requires several ports to be open, as described in the User Guide:

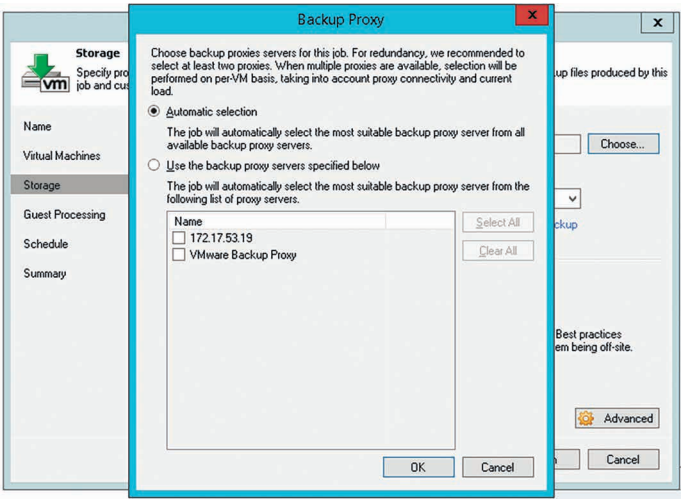
http://helpcenter.veeam.com/backup/80/vsphere/used_ports.html.

**Note**

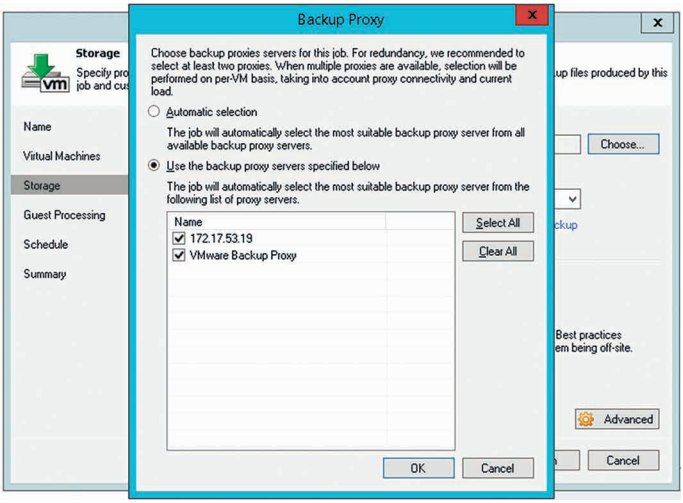
It is highly recommended to maintain a good network connection between the VMware ESXi VMKernel port and Veeam Backup & Replication, as it will be used by many other features like Instant VM Recovery, Virtual Lab and SureBackup, Linux FLR appliance, etc.

For load balancing, Veeam uses a selection of proxy servers based on the network subnet:

- Backup proxies in the same subnets as the VMKernel interfaces (DNS Name of ESXi hosts) are selected if you have the **Automatic Selection** proxy setting configured in the backup jobs.




- If proxy servers do not run in same subnets as the VMKernel interfaces of the ESXi hosts, you will have to manually select the proxies that will process your backup jobs, otherwise it is possible that proxies from other sites will be used to transport data. You can select all proxies from the same site to enable load balancing in that case.




- In case you work with several branches or datacenter environments, it is also recommended that you manually choose the proxies (per site) in the job settings - to streamline and speed up the load balancing.

## Pros:

- Network mode can be used for both backup and restore.
- Can work with both physical and virtual backup proxies.
- Being the most mature of all transport modes, it supports all types of storages.
- Is recommended for use in virtual deployments with NFS-based storage systems, as it helps to minimize VM stunning. See also the "Considerations for NFS Datastores" section of this guide.
- Performance on 10 Gb Ethernet is highly positive.
- As data transfers initiate very quickly, the Network mode is preferable for processing incremental backups on relatively static servers (that is, VMs with small amount of changes).
- Can be helpful when you have plenty of clusters with individual storage configurations (e.g., at hosting providers). In such deployments, using the Network mode for data transfer can help to reduce Veeam footprint and costs, as well as to increase the security (if compared to other modes and storage configuration).

## Cons:

- Typically, Network mode uses only 40% of the physical available bandwidth of the external VMKernel Interface connection due to throttling mechanisms implemented on the management interfaces of VMware vSphere 4.x-5.x.
- Is rather slow on 1 Gb Ethernet (about 10-20 MB/s) also due to throttling mechanisms, so restores via the Network mode will take quite a long time.

**Tip**

You can influence the usage of the specific VMKernel interface by modifying the DNS name resolution for all Veeam backup servers (for example, by adding entries in the *hosts* file or by using special DNS configuration).

## Recommendations

- As the Network mode is relatively inefficient in bandwidth usage, consider setting up at least one virtual backup proxy for hot-add restores – then it will be possible to achieve higher throughput and thus lower RTO.

- You can also restore to a thin disk format and later use standard VMware methods to change the disk format to thick disk, if needed.
- Another way to overcome this limitation is to use Instant VM Recovery with Storage vMotion (if licensed on the VMware vSphere side).
- As there is no overhead (like SCSI disk Hot-Add, or search for the right volumes in Direct SAN) on backup proxies, the Network mode can be recommended for scenarios with high-frequency backups or replication jobs, as well for environments with very low overall data and change rate (VDI).

## Direct SAN Access Mode

The Direct SAN Access mode uses a direct data path (a Fiber Channel (FC) or iSCSI connection) between the VMFS datastore and the backup proxy for data transfer.

## Pros:

- Direct SAN Access mode provides the fastest, the most reliable and predictable backup performance (typically, using 8 Gb Fiber Channel or 10 GbE for iSCSI).
- Produces zero impact on VMware vSphere hosts and production networks.
- Starting with Veeam Backup & Replication v8, it is also possible to perform full VM restore using Direct SAN Access mode: if backup proxies that can use the Direct SAN Access mode are available in the backup infrastructure, and the VM disks are thick-provisioned, this mode will be used automatically.
- Can be used at replica target for the initial replication (with thick-provisioned disk).
- This is the fastest method to restore thick-provisioned disk files.

## Cons:

- This mode requires FC, iSCSI, Infiniband or sharedSAS-based storage. NFS-based storage is currently only supported with NetApp storage systems when Backup from Storage Snapshot feature is enabled and licensed.
- Typically, Direct SAN Access requires a physical server (in-guest iSCSI mapping is an exception to this rule, but it rarely makes sense).

- Direct SAN-based restore is possible only for thick-provisioned VM disks.

- This mode is the most difficult backup mode to configure, as it involves reconfiguring storage and SANs (FibreChannel Zoning, LUN masking, or reconfiguration of iSCSI targets) to provide the physical proxy server(s) with direct access to the production VMFS datastores. When such configuration has been implemented, it is extremely important to ensure that HBAs and NIC drivers and firmware are up-to-date, and that multipathing driver software (e.g. MPIO) is properly configured.

For more information about configuring Direct SAN Access, refer to FAQ at Veeam Community Forums: Direct SAN Access Mode

## Example

If operating datastores and vRDMs on a block storage that can share the LUNs to different systems (SharedStorage) like FibreChannel, FCoE, iSCSI and SharedSAS, you can add a backup proxy as a member to that shared storage using LUN masking. This will allow for accessing the storage system for backup and restore.

Remember to ensure that a connection between the storage and backup proxy can be established – for that, verify FC HBAs, zoning, multipath driver software, and iSCSI configuration including network changes, or SAS HBA and its configuration.

## Recommendations

- Use the multipath driver software of the storage vendors choice (preferred integration into Microsoft MPIO) to avoid disk or cluster failovers at storage level. This will also prevent the whole storage system from being affected by possible failovers if wrong data paths are used.

- If you attach a great number of volumes to the backup proxy, consider that logging and search for the correct volume at the job run can require extra processing time per VM disk (as well as for overall volume count). To avoid that, you can disable logging by setting the registry value to **VDDKLogLevel = 0 (DWORD)** in **the HKLM\SOFTWARE\Veeam\Veeam Backup and Replication** registry key.

**Note**

As this reduces the amount of information in debug logs, remember to enable it again when working with Veeam support (to facilitate debugging of the Direct SAN Access-related challenges).

- To achieve the best performance vs. costs, use fewer proxies with more CPU cores available. This will help to fully utilize the HBA or NIC capacity of each proxy server.

**Tip**

It is highly recommended to contact the storage vendor for optimal settings.

## Security Considerations

While the datastores and vRDMs are mapped to the backup proxy server, during deployment of the proxy role to a Windows VM, Veeam Backup & Replication uses the following security mechanisms to protect them:

- Changes the Windows SAN Policy to "Offline (shared)". This prevents Windows from bringing the attached volumes online, and also prevents Windows write operations to the volumes.
- Deploys VMware's VDDK Kit into the backup proxy server – in most cases, this VDDK Kit coordinates 'read' and possible 'write' process (Direct SAN restore) with VMware vSphere.

If necessary, you can take additional measures, as follows:

- Disable Disk Management snap-in with **Group Policy User Configuration > Administrative Templates > Window Components > Microsoft Management Console > Restricted/ Permitted snap-ins > Disk Management**.
- Avoid providing excessive administrative access to the account used to run Veeam proxy servers.
- Present LUNs as read-only to the backup proxy server. This capability is supported by most modern storages. For that, implement read-only LUN masking on the storage system, or use ReadOnly Zoning on the FibreChannel Switches (possible on most SAN directors) for the Veeam proxy servers.

**Note**

Consider that read-only LUNs will prevent Direct SAN restore.

If a VMFS data store was brought online, including Windows volume re-signaturing, you can contact VMware Support. For more information on Windows re-signaturing process and VMware datastores, please refer to VMware KB1002168: Unable to access the VMware virtual machine file system datastore when the partition is missing or is not set to type fb .

## Virtual Appliance Mode (Hot-Add)

Veeam's Virtual Appliance (or VMware's Hot-add) mode has become quite widespread, as it is the default setting with Veeam Backup & Replication all-in-one deployment in a virtual machine (for details, see the Deployment Scenarios section of the User Guide). It is often used when Veeam is deployed in branch office configurations (ROBO).

This mode supports a 100% virtual deployment, using direct storage access through VMware ESXi storage I/O stack, thus having a very little overhead. For example, in the case of a backup, the disk files from a VMware snapshot are mapped by VMware Disk Hot-Add to a virtual backup proxy server and later unmapped after backup/restore.

**Note**

For more information on how it works, refer to the section "Data Retrieval and Restore in Virtual Appliance Mode" of Veeam User Guide.

Virtual Appliance mode can be recommended for proxies in highly dynamic environments where it can be difficult to maintain access to newly created datastores for Direct SAN Access. In such scenarios, using Virtual Appliance mode for data transport will significantly reduce administrative overhead due to leveraging VMware Hot-Add.

Virtual Appliance mode should be used when it is impossible to leverage Direct SAN, for example, in the case of local datastores or NFS shares - then using VMware Hot-Add will likely provide the optimal throughput.

This mode is also a good choice for VMware VSAN configurations, as Veeam's built-in intelligent load balancing will have VM disk placement awareness.

**Note**

For details on using this mode for NFS datastores, see the "Interaction with vSphere: Considerations for NFS Datastores " section of this guide.

- If using the Virtual Appliance data transport mode with shared storage, it is necessary to deploy at least one proxy for each environment where you add a specific datastore. In most cases, datastores are added on a per-vSphere cluster base, so deploy at least one Hot-Add proxy there.
- When backing up from local datastores, it is necessary to deploy one proxy per host (local disks in the ESXi are used for VM data), otherwise the proxy servers will fail back to Network mode (VMware's NBD).

When planning for the Virtual Appliance mode for a backup proxy, consider the time required for actual hot-add operations (such as adding and removing VM disks from the source virtual machine) – they can add up to 1-2 minutes per VM. Therefore, for a backup job containing 1000 virtual machines, this could result in more than two hours of adding and removing disks with no actual data processing. To mitigate the issue, enable parallel processing and process multiple disks from the same virtual machine simultaneously (using this transport mode).

**Tip**

It is recommended to benchmark how such operations affect the backup window - by monitoring a test job in the vSphere console.

## Pros:

- Using the Virtual Appliance mode for proxy servers enables a fully virtual deployment.
- As the proxy will perform source-side data deduplication and compression, this mode will provide very good performance in environments running 1 GbE configurations.

## Cons:

- If working in this mode, backup proxy will occupy the virtual infrastructure resources, impacting consolidation ratio. This could ultimately require additional physical ESXi hosts and licensing.
- This mode requires additional planning and configuration in the enterprise environments because of the additional large disk Hot-Add processes in VMware vSphere.

## Considerations and Limitations

Additional load is put on the vCenter Server and ESXi hosts as each disk is mapped and unmapped (disk hot-add) at the backup proxies.

**Note**

For more information, see vCenter Server connection overview in the «Veeam Backup & Replication Server» section of this guide.

It may occur that VMware API reports that unmap and snapshot commit were done correctly, but a snapshot file still remains on disk. These "orphaned snapshots" will grow over time and can fill up the datastore, leading to downtimes (such a situation is most possible on NFS-based storage). To overcome this issue, Veeam offers the following methods:

- Veeam Snapshot Hunter – this feature introduced in v8 supports automatic clean-up of orphaned snapshots after each backup process. It can also repair the orphaned snapshots that were left over by other software (including VMware).

- Bypassing VDDK processing to overcome some limitations and performance challenges, in particular:

    - Veeam can back up multiple disks of VM in parallel on same proxy (default number is 4).

    - Typical "I/O bursts" do not happen at Hot-Add restore or replica target Hot-Add processing.

- To avoid some VMware issues related to NFS datastore and Hot-Add processing (described at http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2010953 ), enable a specific setting that will process VM backups only on backup proxies that run on the same host. For details, see http://www.veeam.com/kb1681.

**Note**

For additional tips, refer to the "Impact of Snapshot Operations" section of this guide.

## Recommendations

- You will need at least one type of SCSI controller on each proxy in the infrastructure.

- Add an extra SCSI controller to allow for more VM disks processing in parallel (check the corresponding Veeam proxy settings – default value is 4).

- Specific client OS limitations for Hot-Add processing documented in Veeam Backup & Replication Release Notes at http://www.veeam.com/veeam_backup_8_0_release_notes_rn.pdf and in the KB article at http://www.veeam.com/kb1054 . To test whether Hot-Add data transport mode is possible, refer to http://www.veeam.com/kb1184.

- Disk Hot-Add is supported in all VMware vSphere editions, starting from a standalone ESXi host with a license and including vSphere Standard and vSphere Essentials (unlike the CPU and RAM Hot-Add Feature that is only available in higher VMware vSphere editions ).

- When deploying hot-add backup proxies, try to avoid cloning existing VMs, as this may lead to identical UUIDs in the vSphere environments and cause hot-add operations to fail.

- You may re-use any existing Windows server VM (to save on licensing). The Veeam data mover process runs with 'below normal' priority by default.

**Note**

Changed block tracking (CBT) will be disabled for these hot-add proxies. Consider that it may impact the backup window in case the said virtual machines should be included in backup or replication jobs.

# Veeam Storage Integration and Processing Modes for Backing Up from Storage Snapshots

Storage integration implemented with Veeam Backup & Replication supports HP StoreVirtual, HP StoreServ and NetApp Data ONTAP storage types. Backup proxy servers are utilized for storage discovery, reading and writing data and for leveraging storage snapshots, which minimizes the VMware Snapshot commit load.

**Note**

To read more about storage integration, refer to Veeam Backup & Replication User Guide and corresponding sections of this guide.

Veeam Backup & Replication uses storage snapshots to minimize VMware VM snapshot lifetime and to reduce load on VMware ESXi hosts and storage systems.

The Storage integration feature utilizes proxy servers to read data directly from these snapshots (without adding them to a proxy ESXi host), to transport data and to scan the storage systems for possible restore points in storage snapshots. Based on the supported storage protocol, this feature requires direct network connections or SAN connections (like iSCSI/Fiber Channel).

**Note**

For more information, refer to the corresponding sections of this guide and Veeam Backup & Replication User guide.

# Choosing Transport Mode

## Very small sites (single host with local disks as primary datastores / branch office)

- Virtual Appliance (Hot-Add) mode is the recommended option, as it gives you the best performance.
- NBD over 10GbE VMKernel Interfaces link will provide a very stable and good performing solution
- NBD over 1GbE VMKernel Interfaces can be used for failover.
- Direct SAN Access mode or Veeams storage integration based backups cannot be implemented, as the disks are local and thus cannot be mounted to a physical proxy server.

## Small sites (2-4 hosts with shared storage / branch office/ datacenter)

- If storage integration is possible, the corresponding transport modes are preferred.
- If the primary datastore runs on block protocols like Fiber Channel or iSCSI, then Direct SAN Access mode will provide the best performance (if all corresponding requirements are met).
- For example, you can place a physical server with access to FC data stores to the local site and perform backups to a local repository (scenario with a single Veeam backup server). If you use thin-provisioned disks for the VMs, configuring a dedicated backup proxy for restore in the Virtual Appliance (Hot-Add) mode can help to increase restore performance.
- The Virtual Appliance (Hot-Add) mode is a recommended option if the above modes cannot be used.

**Tip**

If you are using NFS primary data stores, consider setting up a backup proxy on same host (as described in the previous scenario).

- NBD over 10 GbE VMKernel Interfaces link will provide a very stable and good performing solution.

- NBD over 1 GbE VMKernel Interfaces can be used for failover and for situations where you do not have to transport much data, or if you cannot use the Direct SAN Access mode but look for the best reliability.

## Medium sites (4-20 hosts with shared storage / branch office / datacenter)

- Use storage integration for all VMs with a high change rate. Veeam ONE "VM Change Rate Estimation" report can help to identify them. For VMs with low change rate, you can decrease the overhead of storage snapshot processing (faster as less data needs to be moved).

- If the primary datastore is run on block protocols like Fiber Channel or iSCSi, then Direct SAN Access mode will provide the best performance (if all corresponding requirements are met).

- For example, you can place one or multiple physical servers with access to the Fibre Channel data store to the local site and perform backups to a local or centralized repository (scenario with multiple proxies backup to single repository server). If you use thin-provisioned disks for the VMs, configuring several proxy servers for restore in the Virtual Appliance (Hot-Add) mode can help to increase restore performance in that scenario.

- NBD over 10 GbE VMKernel Interfaces link will provide a very stable and good performing solution; it can be recommended for virtual only or NFS data stores.

- The Virtual Appliance (Hot-Add) mode is a recommended option if the above modes cannot be used. A Proof of Concept is recommended after thorough planning to implement this mode.

- NBD over 1GbE VMKernel Interfaces can be used for failover.

## Large sites (21-100 hosts with shared storage / branch office / datacenter)

- Use storage integration for all VMs with a high change rate. Veeam ONE Change Rate Estimation report can help to identify them (faster as less data needs to be moved).

- If the primary datastore is run on block protocols like Fiber Channel or iSCSi, then Direct SAN Access mode will provide the best performance (if all corresponding requirements are met). For example, you can place one or multiple physical servers with access to the Fiber Channel datastore at the local site and perform backups to multiple repositories. If you use thin-provisioned disks for the VMs, configuring several proxy servers for restore in the Virtual Appliance (Hot-Add) mode – per cluster - can help to increase restore performance in that scenario.

- NBD over 10 GbE VMKernel Interfaces link will provide a very stable and good performing solution; it can be recommended for virtual only or NFS data stores.

- The Virtual Appliance (Hot-Add) mode implementation should be planned carefully, regarding vCenter connection count and general overhead on VMware side for disk mount and un-mount processes.

- NBD over 1 GbE VMKernel Interfaces can be used for failover for some VMs; however a 10GbE network on VMKernel interface is highly recommended.

### Enterprise or large datacener (above 100 hosts)

- Use the recommendations for Large Sites scenario above.

- In large scale environments, the deployment of Veeam components, configuration and job creation is typically automated using the Veeam PowerShell integration.

- To balance the management load, it is recommended to use multiple Veeam backup servers for at least every 5000 VMs, together with a central reporting and administration by Veeam Backup Enterprise Manager and Veeam ONE.

- If you run a central Veeam backup server and multiple branches with it, a separate Veeam Backup & Replication management instance is recommended for at least every 200 branches. Consider using Veeam Backup Enterprise Manager for central administration.

## Sizing a Backup Proxy

### Processing Resources

As per system requirements, a proxy requires 2 GB RAM plus 200MB for each concurrent task (VM disk to be processed).

As described above, you can define the max concurrent tasks value in the backup proxy settings. On average, a task consumes 1 Core or 1 vCPU Core (for compression and encryption). Depending on the job setup and other settings, these numbers can vary. A PoC in the specific environment can help to estimate resource usage, especially if you need to fit into an existing budget.

- If you double the proxy task count, that will, in general, decrease backup time window up to 50%.
- It is recommended to plan for some additional resources – for further growth and possible new features: for example, the RAM usage for one proxy processing task should not be lower than 2 GB - to be prepared for upcoming features.

As a rule of thumb, a proxy will need 1 CPU+2 GB RAM for each 30 VMs (with average change rate of 2-3 % at the block level) to fit into a 8-hour backup window.

## Example: Calculating Overall Task Count

Sample infrastructure has the following configuration:

- 480 VMs
- 48 TB used data
- Backup window - 8 hours
- Daily change rate - 3%

For that, the following calculation can be used as a starting point:

30 VMs per CPU core (physical core or vCPU core) for a proxy. => 1 proxy task slot (1 core, 2 GB RAM). => 16 cores for 480 VMs => 1 server with 2x 8 core 32 GB RAM

If you need to achieve a 4 hour backup window, then double the resources => 2 server with 2x 8 cores 32GB RAM.

The same counts if you have two times bigger amount of data (with 3% change rate).

**Note**

Overall performance largely depends on the underlying storage and network infrastructure.

Required processing resources may seem too high if compared with traditional agent-based solutions. However, consider that instead of using all VMs as processing power for all backup operations (including data transport, source deduplication and compression), Veeam Backup & Replication uses its central proxy and repository resources. Overall, required CPU and RAM resources are normally below 5% (and in many cases below 3%) of all virtualization resources utilized by backup and replication jobs.

## How Many VMs per Job?

Best practice is to add 20-50 VMs to a job (30 VMs at about 1000-2000 VMs; 50 VMs at 5000 VMs).

Also, remember that the number of running backup jobs should not exceed 100 jobs concurrently running (not overall). Veeam can handle more, but a "sweet spot" for database load, load balancing and overall processing is about 80-100 concurrently running jobs.

## How Many Tasks per Proxy?

Typically, in a virtual environment, proxy servers use 4, 6 or 8 vCPUs, while in physical environments you can use a server with a single quad core CPU for small sites, while more powerful systems (dual 16 core CPU) are typically deployed at the main datacenter with the Direct SAN Access mode processing.

**Note**

Parallel processing may also be limited by max concurrent tasks at the repository level.

So, in a virtual-only environment you will have slightly more proxies with less proxy task slot count, while in physical infrastructure with good storage connection you will have a very high parallel proxy task count per proxy.

The "sweet spot" in a physical environment is about 20 processing tasks with 2x 16 Gbps FC cards for read.

Depending on the primary storage system and backup target storage system, any of the following methods can be recommended to reach the best backup performance:

- Running fewer proxy tasks with a higher throughput per current proxy task
- Running higher proxy task count with less throughput per task

As performance depends on multiple factors like storage load, connection, firmware level, raid configuration, access methods and others, it is recommended to do a Proof of Concept to define optimal configuration and the best possible processing mode.

### Considerations and Limitations

Remember that several factors can negatively affect backup resource consumption and speed:

- **Compression level**: It is not recommended to set it up to *High* (as it needs 2 CPU Cores per proxy task) or to *Extreme* (which needs much CPU power but provides only 2-10% additional space saving).
- **Block Size**: the smaller the blocks size is, the more RAM is needed for deduplication. For example, you will see a RAM increase when using LAN mode if compared to Local target, and even greater (2-4 times) when using WAN. Best practice for most environments is to use default job settings (*Local* for backup jobs and *LAN* for replication jobs) where another is not mentioned in the documentation or this guide for specific cases.
- **Antivirus Scanner** - see the corresponding section of this document.
- **3rd party applications** – it is not recommended to use an application server as a backup proxy.

## Repository Server

Before you start planning for the repository, go through Veeam Backup & Replication online documentation at http://www.veeam.com/documentation-guides-datasheets.html to get basic understanding of repositories.

In brief, backup repository is a storage location used by Veeam Backup & Replication jobs to store backup files, copies of VMs and metadata for replicated VMs. Technically, a backup repository is a server that hosts Veeam Transport Service and provides a destination folder on the backup storage. Each job can use only one repository as its destination storage, but one repository can be used by multiple jobs in parallel. You can balance the load across the backup infrastructure by setting up several repositories in the environment and limiting the number of parallel jobs for each one.

# Choosing Repository Type

Being storage-agnostic, Veeam Backup & Replication supports a wide range of repository types. When deciding on repository storage, consider the following factors:

- Capacity
- Write performance
- Read performance
- Data density
- Files security
- Backup file utilization

As a basic guideline, a repository should be highly resilient, since it is hosting the essential copy of data. It also needs to be scalable, allowing the backup to grow as needed.

Organization's policies may require different storages for backups with different retention. If so, you can configure, for instance,two repositories:

- A high-performance repository hosting only several recent retention points for instant restore
- A repository with more capacity, probably using a cheaper and slower storage, storing long-term retention points - you can set up a backup copy job with that repository as the destination.

**Important!**

Remember that repository design is not only about storage footprint, it is also about IOPS, depending on the intended use.

# Server-Based Repository: DAS, SAN, or NFS?

## Direct-Attached Storage

This is a cheap, easy-to-use solution that can be very efficient in terms of performance; however, it is less manageable due to non-transportable volumes, capacity growth, and so on.

- With a proper RAID configuration, the physical layers between the server and storage are minimal; since the DAS storage can be fully dedicated to backup, this type of repository is considered to offer a good balance between the "performance" and "cost" factors.

- A strong benefit of a JBOD repository is that it supports the features offered by Veeam Backup & Replication in a very flexible way. In particular, is provides good read and write performance, sufficient for vPower-based features (such as Instant VM Recovery, SureBackup, and others). As it typically provides good random I/O performance, it will be the optimal solution when using I/O intensive backup modes such as reverse incremental or forever forward incremental (also used in backup copy job).

However, consider that though DAS is a valuable option in many cases, its scalability may not meet organization's requirements. When using DAS, it is important to ensure proper proactive monitoring to avoid hardware failures (such as multiple RAID drive failures or other issues).

**Tip**

To achieve optimal performance, it is often required to install a battery module to the server's RAID to enable write-back mode for the flash cache.

| Pros | Cons |
|---|---|
| Cost efficiency | Manageability |
| Performance | Single point of failure |
| Simplicity | Monolithic |

## SAN Storage

This is a more modern, manageable solution that offers the same advantages that DAS, and adds more refinements e.g. higher availability and resiliency.

The volume size and quantity are easily scalable, offering scalable capacity. Volumes can be easily moved from one server to another in case of a Veeam Repository Server failure.

**Tip**

You can configure multiple backup repositories on the SAN storage to increase repository throughput to the storage system.

| Pros | Cons |
|---|---|
| Reliability | Complexity |
| Performance | Cost |
| Technical possibilities | |

## Windows or Linux?

Technically, the main difference between Windows and Linux is the way they handle NAS shares – this can be summarized as a choice between NFS and SMB. Generally, a Linux-based repository can handle a higher throughput than a Windows-based repository with same CPU/RAM/Disk resources. However, if you deploy Veeam in a small-sized infrastructure, you may want to keep the configuration "all-in-one" on a single Windows server, so deploying a Linux server as a repository could add extra complexity to the solution. Other possible concerns relate to cost and administrative burden.

## Physical or Virtual?

You can use a virtual machine as a repository server, however, keep in mind that the storage and associated transport media will be heavily occupied.

If you are using a SAN storage, it can be accessed through software iSCSI initiators, or directly (as a VMDK bound to the VM).

Best practice is to avoid using any storage technology that can be a single point of failure leading into data loss together with virtualization infrastructure. For example, you can add the backup target storage by physical Raw Device Mapping (pRDM) into the VM, so if VMFS has a failure, it will not affect the backup target storage.

## CIFS Repository

While a CIFS (or SMB) repository is often considered to provide less performance than direct attached storage, it still can provide very good results as a repository due to leveraging Veeam's load-balancing technology for write operations, as explained in the next sections.

# CIFS Gateway Server

When you set up an SMB share as a repository, the following options are available:

- Automatic selection of the server as the CIFS gateway proxy (that is, the server that will host the target-side transport and thus perform the role of "data writer")
- Specifying a unique server (from Windows servers added in Veeam Backup management console) as a CIFS gateway proxy

The second option is very helpful in case the SMB share is located remotely — if so, it is recommended to use a CIFS gateway server as close as possible to the SMB storage. Remember, however, to keep the data flow under control.

To get the best performance out of CIFS storage, you can set up as many connections as possible, and also leverage the automatic CIFS gateway proxy load balancing by selecting **Automatic** from the drop-down list.

**Note**

Automatic selection will disable metadata caching (see Release Notes for Veeam Backup & Replication v8 Update 2b at http://www.veeam.com/kb2024 ).

## Load Balancing (with Automatic Proxy Selection)

A Veeam backup job features out-of-the-box load balancing across proxies in charge of reading data from the datastores. This means the following for any data transport mode: each proxy server (eligible for the certain job) will read data from multiple VMDKs off the job — but only one proxy (called "gateway proxy") per job will be used to write data to the CIFS share. In the **Automatic** mode, the first selected proxy in the running job will also start this gateway proxy instance.

In addition to traditional load balancing, this method — together with an Active/Active or Active/Passive (with automatic failover) CIFS target storage — will provide High Availability on the proxy and repository side.

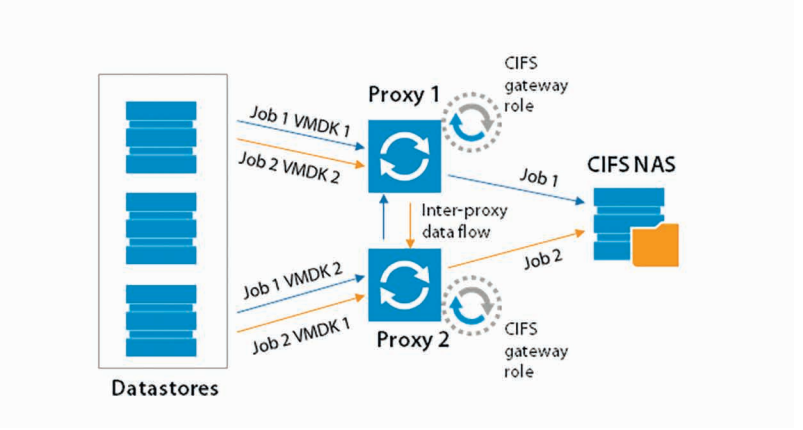
Here are some recommendations for managing Veeam backup proxies:

- The inter-proxy networking should be fast enough to allow seamless load balancing.
- For **Automatic** proxy gateway selection, proxies should have enough resources to host multiple of the gateway proxy instances (see resource requirements for traditional repositores).

**Note**

Consider that increasing the number of jobs also increases the number of threads to the NAS storage.

- As the first proxy of a job is used as a gateway proxy, it may happen that all gateway proxy instances are started on the same proxy and require, consequently, enough CPU and RAM resources. Thus, be sure to monitor RAM utilization of all proxies so that you can make configuration changes when needed.




Scaling out using this approach will allow you to process bigger amounts of data and optimize the ingest rate of the NAS shares. Best practice for large-scale environments is to use at least a mid-range or enterprise NAS storage system that provides good I/O performance. Low-end NAS devices often implement non-ideal changes to the SMB protocol that will improve performance test results, but may corrupt backup files. For low-end NAS devices, it is recommended to use a block-based protocol, such as iSCSI.

**Tip**

Check with the storage vendor for the best practices for optimizing the storage system for Veeam workloads.

**Important!**

To prevent possible network issues, make sure that network traffic is under control and that automatic load balancing is used properly, as described above.

# Optimizing for Deduplication Storage

If you plan to use a deduplicating storage appliance as a repository, consider the following options (available on **Repository** page of the repository wizard):

- For storage systems using fixed block size deduplication, you may want to enable the **Align backup file data blocks** option — then Veeam Backup &Replication will align VM data (saved to a backup file) to a 4 KB block boundary. This option provides better deduplication across backup files, but it can lead to greater amount of unused space on the storage device; if used with a traditional storage system without deduplication, it may result in higher level of fragmentation.

**Tip**

It is recommended to disable this option for deduplicating storage that uses variable block size deduplication, such as HP StoreOnce and EMC DataDomain.

- When you enable compression for a backup job, VM data is compressed at the source side before it is transmitted to the target. You can use the **Decompress backup data blocks before storing** option on backup repositories — then, if data compression is enabled for a job, Veeam backup proxy will compress VM data, transmit it efficiently over the network, decompress data on the target side and write uncompressed VM data to the storage device. In this case, you should pay attention to the additional load generated on the repository or CIFS gateway server.

As for any other deployment and configuration option, this is all about finding the right balance for the environment – here between better performance and higher dedupe ratio. Veeam Backup & Replication offers a flexible set of configuration options, so you can make the optimal choice between the computing power of the proxies and repositories and the available network bandwidth. For example, if the proxy and CIFS gateway/repository are communicating through a low bandwidth link, then it can be reasonable to first compress data, then send it and decompress on target – for that, you will need more powerful computing resources on both sides.

### Deduplication File Systems

Generally, a deduplication repository reduces footprint but makes the recovery process longer, as it will need to "rehydrate" data so you can access the VMs. This process is quite sophisticated if the backup job is configured to perform transformations on the backup files (like forever forward incremental, or reverse incremental, or synthetic full). Thus, deduplication repository should only be used with an active full backup job and with the backup window that is long enough to perform a periodic full backup.

Another option is to use deduplication storage systems (from those supported by Veeam) that enable syntetic full processing with these devices. For example, this refers to DDBoost feature of EMC Data Domain storage enhancements (see the corresponding section in the Veeam Backup & Replication User Guide for this).

Veeam users also utilize repositories hosted on Windows Server 2012 deduplication file system. This solution can be implemented to reduce footprint (consider Read rate as a counterpart). For more details, settings, performance and limitations, see http://forums.veeam.com/veeam-backup-replication-f2/best-practice-for-ms-server-2012-dedup-repo-t14002-135.html

## Sizing Computing Resources

In mid-sized or enterprise customer environments, the recommended CPU and RAM for a repository is 1 vCPU or core per concurrent job that process data on a repository server. At least 2 vCPU or cores allow for the OS to be more responsive.

It is recommended to configure 4 GB RAM per vCPU or core. The same amount of resources are needed for CIFS/SMB gateway servers. Also, consider that VM recovery processes (Instant Recovery, FLR and others) require sufficient resources (as described here).

### Estimating Repository Capacity

When estimating the amount of disk space required, you should know the following:

- Total size of VMs being backed up
- Frequency of backups

- Retention period for backups
- Will jobs use forward chains or reverse incremental

Also, make assumptions on compression and deduplication ratios, change rates, and other factors. The following figures are typical for most sites (however, it is important to understand the environment if there are exceptions):

- Compression and deduplication savings 2:1 or more; typical is 3:1 or better, but you should always be conservative when estimating required space.
- Typical change rate of 2-5% per day overall at a mid-size or enterprise environment; this can greatly vary among servers; some servers show much higher values.
- Include additional space for one-off full backups.
- Include additional space for backup chain transformation (forward forever incremental, reverse icremental) – at least the size of a full backup multiplied by 1.25x.

**Note**

When using deduplication storage, please consult the vendor for sizing guidelines.

Using the numbers above, you can estimate required disk space for any job. Besides, always leave plenty of extra headroom for future growth, additional full backups, moving VMs, restoring VMs from tape.

**Tip**

With Veeam Availability Suite v8, you can use the "VM Change Rate Estimation" report from the "Infrastructure Assessment" report pack as an indicative pre-deployment assessment of the potential amount of space that should be available on the backup repositories. This report is built measuring the number of VM virtual disk write operations supplied by VMware vSphere.

It is also recommended to periodically run the "Capacity Planning for Backup Repositories" report from the "Veeam Backup & Replication Reports" pack to analyze the amount of free space on backup repositories and estimate the projected date when the repository will run out of available storage capacity. The report provides recommendations on how to adjust the allocated storage resources in order to meet the future demand for backup storage. Furthermore, it calculates the amount of additional space that needs to be provisioned to accommodate the necessary restore points.

For more information on Veeam Availability Suite v8, please refer to its Reviewer's Guide at http://www.veeam.com/documentation-guides-datasheets.html

**Tip**

A repository sizing tool that can be used for estimation is available at http://vee.am/rps.

## Examples

The examples below explain the impact of backup method and retention policy on the estimated repository size, assuming the environment is the same in all three cases.

Environment: 10 VMs, 100GB each, 80GB avg/used

2:1 Estimated Compression/Deduplication, 5% daily change

**Example 1**

Backup: Reverse Incremental, Daily Backup, 30 Day Retention

- Estimated Full Backup Size: 10 * 80GB (Used space) * 50% (2:1 Compression) = 400GB
- Estimated Reverse Incremental Size: 10 * 80GB * 50% (2:1 Comp) * 5% (Change Rate) * 29 (reverse incremental restore points) = 580GB
- Spare : 500 GB
- Estimated total Backup Size: 400GB + 580GB + 500 = 1480 GB

**Example 2**

Backup: Forward Incremental, Daily Backup, 30 Day Retention, Weekly Full

- Estimated Full Backup Size: 10 * 80GB (Used space) * 50% (2:1 Compression) = 400GB
- Estimated space for 6 Weekly Fulls (Max required for 30 Day Retention): 400GB * 6 = 2400GB
- Estimated Forward Incremental Size Max: 10 * 80GB * 50% * 5% * 32 = 640GB
- Estimated total Backup Size: 2400GB + 640GB = 3,040GB (~3TB)

**Example 3**

Backup: Forward Incremental, Daily Backup, 30 Day Retention, Monthly Full

- Estimated Full Backup Size: 10 * 80GB (Used space) * 50% (2:1 Compression) = 400GB
- Estimated space for 3 Monthly Fulls (Max req for 30 Day Retention): 400GB * 3 = 1200GB
- Estimated Forward Incremental Size Max: 10 * 80GB * 50% * 5% * 60 = 1200GB
- Estimated total Backup Size: 1200GB + 1200GB = 2,400GB (~2.4TB)

To summarize, when estimating the size of the repositories, use the following best practices:

- Be conservative when estimating compression and deduplication ratios if actual ratios and disk content are unknown.
- Use higher estimates for change rate if a significant number of servers are transactional such as Microsoft SQL and Microsoft Exchange.
- Include enough free space to take at least one and a quarter extra full backup for each transformation job.

# Configuration Guidelines

## Parallel Processing

A repository can be configured to manage a certain amount of parallel tasks at a time; here a *task* is one VMDK handled by the proxy, or one backup copy job (sequential by nature) in progress. When enabling parallel processing, please note if there are many parallel tasks on the proxy side for only few tasks on the backup repository, this will lead the Veeam scheduler service to wait for available resources on the repository. To prevent such situation, you can figure out on which side the bottleneck will be (proxy or repository) and then set the overall amount of parallel tasks on the proxies equal to the total amount of parallel tasks on the repositories.

**Note**

Consider tasks for read operations on backup repositories (like backup copy jobs).

## Blocks Alignment

As a part of the backup process, data blocks are processed in chunks and stored in backup files in the backup repository. You can customize the block size using the **Storage Optimization** setting of the backup job. By default, block size will be set to **Local target**, which is 1 MB before compression for backup and LAN (512 KB) for replication.

Whilst compression ratio is very often around 2x, with this block size Veeam will write about 512 KB or less to the repository. Though this is not a key figure in case of a mid-range or enterprise-class repository storage, this estimation can help speeding up backups on a low-range storage system. On this kind of storage, you may want to keep transferred block size as close as possible to the stripe size in order to mitigate the *write* "penalty" caused by RAID parity calculation. Configuring the stripe size as close to the effective backup block size will enhance overall I/O rate of the repository storage. For some storage systems, performance can be increased using a 2x higher block size. Check recommendations from the storage vendors for detailed information.

**Tip**

As can be seen from the field, optimal value for the stripe size is often between 256 KB and 512 KB; however. It is highly recommended to test this prior to deployment whenever possible.

For more information, refer to this blog post: http://www.virtualtothecore.com/en/veeam-backups-slow-check-stripe-size/

## File System Formats

In addition to the storage stripe size alignment, the file system may also benefit from the best cluster size (or Allocation Unit Size). By default, Allocation Unit Size is set to 4KB; to mitigate fragmentation issues, set it to 64 KB.

It is also recommended to use journaled file systems (this makes exFAT a less reliable option than NTFS).

## Using "Large File" Switch for NTFS

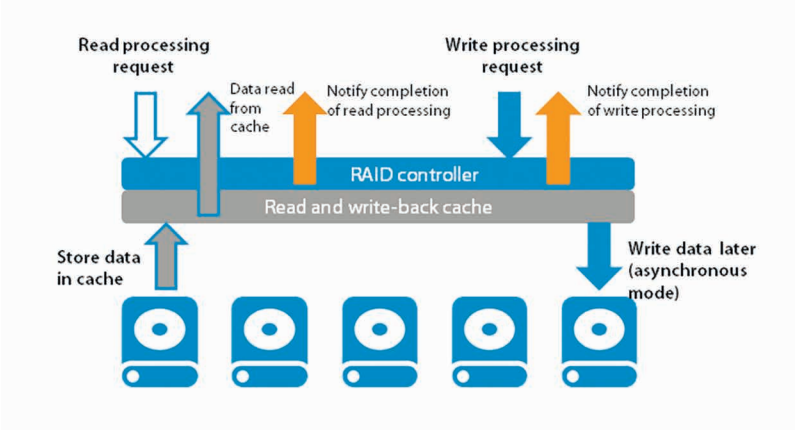
A file size limitation can be occasionally reached on NTFS (especially on Windows 2012 R2 with deduplication). This happens due to a hard limit reached on the file records size – in particular, on deduplication NTFS – because of high file fragmentation. To mitigate the issue, best practices recommend to format Windows NTFS repositories with the "**/L**" (large files) option.

## Keeping File Size Under Control

Try to avoid the backup chains growing too much. Remember that very big objects can become unmanageable. Since Veeam allows a backup chain to be moved from one repository to another with nothing more than a copy/paste operation on the files themselves, it is recommended to keep backup chain size (of a single full and relevant incrementals) under 10 TB per job (~16TB of source data). This will allow for a smooth, simple and effortless repository storage migration.

## Synthetic Backup and Caching

To get the best out of a synthetic backup and enhance the performance, it is recommended to use a write-back cache. Read and write request processing with write-back cache utilization is shown in the figure below.




## Example: Using Windows 2012 R2 Server with Built-In Deduplication as a Veeam Repository

**Important!**

If you plan to use a Windows 2012 or Windows 2012 R2 server as a backup repository, it is strongly recommended to deploy Veeam backup server on the machine with the same Windows version. That is, you should either have both the repository and Veeam backup server running on Windows 2012 machines, or both running on Windows 2012 R2 machines.

Follow the recommendations provided in the configuration guidelines above; here is the summary:

1. Use **Windows 2012 R2** and apply all patches (some roll-ups contain improvements to deduplication).
2. Format the disk using the command line "**/L**" option (for "large size file records") and **64KB** cluster size.
3. Turn Veeam compression **OFF** and use the **LAN** block size (Veeam deduplication can stay ON) for better overall space savings.
4. Modify garbage collection schedule to run daily rather than weekly.
5. Use backup jobs configured to perform Active full with incrementals.
6. If possible, spread active full backups over the entire week.
7. Try to keep the .VBK files **below 1TB** in size (there is no official support from Microsoft for files bigger than this; see https://msdn.microsoft.com/en-us/library/hh769303(v=vs.85).aspx). Large files take a long time to deduplicate and will have to be fully reprocessed if the process is interrupted.
8. Where possible, use multiple volumes. Windows deduplication can process multiple volumes using multi-core CPU – one CPU core per volume; see http://blogs.technet.com/b/filecab/archive/2014/12/04/sizing-volumes-for-data-deduplication-in-windows-server.aspx for details.)
9. Configure deduplication process to run once a day, and for as long as possible.

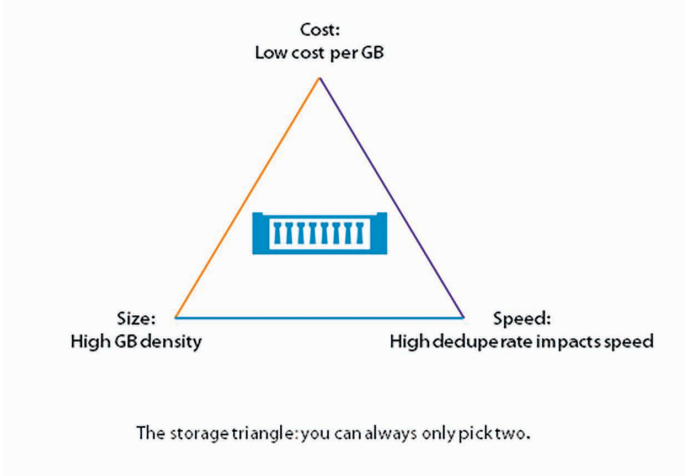
More information can be found here:

http://forums.veeam.com/veeam-backup-replication-f2/best-practice-for-ms-server-2012-dedup-repo-t14002-135.html.

# Deduplication Appliances

## Overview

Essentially, a deduplication storage is aimed at reducing footprint on long-term retention storage at the cost of an overall performance reduction due to data digest calculation.




The storage triangle: you can always only pick two.

**Note**

Basic facts about deduplicated repositories can be found in the "Repository Server" section of this document.

Deduplicated storage systems are often optimized for writing and can offer rather high ingests rates. However, any non-sequential read I/O pattern will suffer a lot from compute resource required by data "re-hydration", so please consider that this kind of storage should be used with care, thorough planning, and for their intended use only, which means lots of writes for long-term retention.

## Using a Deduplication Appliance

As a storage-agnostic product, Veeam Backup & Replication will let you use a deduplication appliance as any repository storage, that is, for the following purposes, each described in the separate section below: primary backup repository, backup copy repository, and VTL container.

## Deduplication Appliance as a Primary Backup Repository

Unless you are using DDBoost on DataDomain storage, you should configure the jobs for active full backups plus incremental backups - since jobs with transformation will require block "de-hydration" and then "re-hydration" on the storage. These operations might require significant time.

**Note**

"Re-hydration" means retrieving/transferring the original blocks in a non-deduplicated form, an operation that consumes hardware resources of the appliance. During backup files transformation, the same blocks are read and then written back to the appliance where they are de-hydrated (deduplicated) again. This two-step process can generate significant load on the appliance, slowing down operations.

Also, consider that Instant VM Recovery might not be as fast as expected – unless the deduplication appliance offers a fast non-deduplicating landing zone for the latest restore points (such as ExaGrid). In this case, Instant VM Recovery will work as fast as expected from a traditional storage – but only for the latest restore points (residing in the landing zone). If you have the latest increment in the landing zone, and full in dedupe zone, restore process will be almost as slow as if you had no landing zone at all. These considerations also apply to SureBackup jobs.

**Note**

For performance comparisons, review Veeam Community Forum topic at Version 7 and Deduplication devices.

Thus, if you plan to use this storage as a primary repository, configure backup jobs for active full and incremental backups and consider that instant restores may take some time. This approach will allow you to achieve high ingest rates and significantly reduce footprint. On the other side, you will need to transport 100% of data on a weekly/monthly base into backup. This can lead to long snapshot commit, so this mode needs to be planned wisely. As an alternative, a deduplication storage system supported by Veeam can be used, with processing functionality similar to DDBoost (that performs synthetic operations on the storage device).

With proper configuration, it is possible to achieve more than 1 GB/s on 10 GbE links and CIFS access for a single high-end appliance.

## Using Deduplication Appliance as a Backup Copy Repository

By design, a backup copy job applies transformations to the backup chain. This will lead to this "de-hydration"/"re-hydration" overhead at the end of the backup copy job cycle (due to synthetic full or transformation). Keep this in mind when looking for a balance between job overall duration and backup copy scheduling.

For example, if you plan to use DataDomain with DDBoost, consider that synthetic operations will be performed on the storage device, so they

will require minimal additional time and produce no impact on production infrastructure. On the other hand, if backup copy jobs have GFS enabled, using DataDomain will help you to save storage space.

**Note**

Refer to this blog article about the issue.

## Using Deduplication Appliance as a Virtual Tape Library Container

If a deduplication appliance can be used as a Virtual Tape Library (VTL), this might be very helpful as long as the backup job is configured with the compression level set to **None** or **Dedupe-friendly**.

Also, before you start using it, make sure that the appliance meets Veeam tape requirements described in the User Guide.

**Important!**

If backup jobs are configured with compression rate other than **None** or **Dedupe-friendly**, it is not recommended to use the deduplication appliance as a VTL because you will then write compressed data (deduplication will be less efficient).

These backup to tape jobs will not produce any transformations on the appliance, letting you benefit from the footprint reduction for long-term retention of backup data. However, overall backup management will require more attention in this case.

As these jobs are tape jobs, they do not support direct SureBackup, Instant VM Recovery, direct VM restore or direct guest object recovery – these operations are not performed directly from tape, but involve a repository as interim staging storage. See the User Guide at http://helpcenter.veeam.com/backup/80/vsphere/index.html?tape_restoring_vm_from_tape.html for details.

## Recommendations and Best Practices

### Parallel Write Operations

You can benefit from using multiple parallel write operations to a deduplication appliance, as follows:

- If sufficient throughput is not achieved, try increasing the number of jobs to generate more write streams (as one job will generate one write stream)

- Leverage multiple CIFS gateway servers to generate more TCP-IP connections.

**Note**

There is generally a maximum supported number of data streams depending on the specific appliance and the stream type. Contact appliance vendor to obtain the most efficient number of write streams for each specific model.

- Matrix for EMC DD OS 5.4: https://community.emc.com/docs/DOC-45240
- Matrix for EMC DD OS 5.5: https://community.emc.com/docs/DOC-45301

See also: http://forums.veeam.com/veeam-backup-replication-f2/write-streams-t21574.html.

This feature is especially useful in case you have LACP (or any equivalent networking load balancing algorithm) and utilize the appliance in CIFS mode. The more gateways are involved, the more jobs can be processed and more overall throughput can be reached (in some cases, until you reach physical limitations of the network bandwidth).

### Veeam File-Level Recovery and Veeam Explorer Tools

By design, Veeam Explorer tools (should they concern file or application restore) will perform Read operations intensively on the repository side. This may lead to long waiting time while the management interface is launched. Thus, consider that with a deduplication appliance it can be possible to restore a full VM faster than a single file, generating a continuous, sequential read flow from the storage.

When performing file-level restore, it is recommended to use Veeam file indexing for the VM guest OS, as the file selection can be performed via Veeam Backup Enterprise Manager, browsing the catalog of indexed files. While this may not enhance restore performance, it will provide a better user experience, as waiting time is significantly reduced.

### Job Configuration

Several parameters can be used to optimize a deduplicated repository, so when planning for backup job configuration, consider the following settings:

- **Compression level** – can be set to **Optimal**.

**Tip**

You can have **Optimal** compression to save traffic between source and target data movers, and use **Decompress on target** option in the **Advanced** properties of repository to save data in decompressed format. (For example, if the network bandwidth between the backup proxy and repository is low, you can leave compression enabled and allow the repository server to decompress data before storing.)

- **Veeam Inline Deduplication** – Veeam deduplication needs to read metadata from the backup files (in a random manner), so you can increase the performance by disabling this setting on the job. On the other hand, if this option is disabled and CBT fails, Veeam will have to stream all of the data. Thus, if existing performance is satisfactory, it is recommended to leave this option selected.

- **Block size** – Veeam block size is configured per job, also on the **Storage** tab of the job's **Advanced** settings for repository storage. You may want, for example, to select the **Local target (16 TB + backup files)** option and use 8 Mbyte blocks. This will limit the amount of generated metadata and optimize sequential writes on the storage.

**Important!**

Remember that large blocks result in faster backups but slower restores (especially with Instant Recovery). Similarly, small blocks result in slower backup but faster restore.

- **Encryption** – deduplication rates will be heavily affected by backup file encryption, so it is not recommended to activate encryption for backup files residing on such storage appliances.

**Note**

See also the "Encryption" section in the "Job Configuration" chapter of this document for more information.

For example, if *ESXi cluster1* is backed up using *key1*, and *ESXi cluster2* is backed up using *key2*, both using the same deduplication storage, then blocks that could be common between both backups chains will be different - because of encryption-generated differences. (In fact, even if you have 2 exactly same blocks in one file, in that case they will be encrypted into two totally different blocks.)

### Backup File Size

General recommendation is to keep multiple smaller backup chains on deduplication appliances to reduce the amount of metadata that needs to be read, before a restore process can be initiated. As a starting point, try keeping the backup chain size (of a single full and relevant incrementals) under 10 TB per job (~16 TB of source data). This figure can be adjusted.

## Recommendations for EMC Data Domain Appliances

### Using DDBoost

When deciding on whether to use DDBoost, consider that its usage will allow for the following capabilities :

- High deduplication between the Veeam gateway server and Data Domain appliance since deduplication will be processed at the source side. This will reduce the amount of data sent over the network to the appliance, and thus reduce datacenter footprint.

**Important!**

DDBoost over WAN is not supported as current version of DDBoost libraries assumes low latency and reliable links.

- Better LAN parallelization, since DDBoost manages its own network load balancing algorithms which are considered more efficient than standard network links aggregation.
- Seamless Veeam files transformations like synthetic full or forever forward incremental.
- DDBoost can be used through Fiber Channel SAN, providing a totally LAN-free backup solution.

For more details, refer to the DDBoost configuration guide by Rick Vanover at https://www.veeam.com/wp-configuring-emc-data-domain-boost-with-veeam-availability-suite-v8.html.

### Chain Length Limitation

Consider that Data Domain can support only up to 59 incremental restore points for a single full backup. For details, refer to the Veeam Backup & Replication User Guide.

## Recommendations for ExaGrid Appliance

Starting with Veeam Backup & Replication v8, the ExaGrid appliance can integrate with the Veeam Data Mover, so a repository server can be directly integrated with the appliance. In this case, you will not need an additional Windows or Linux server to act as a repository gateway server.

After you select an ExaGrid appliance as repository type, all jobs that use this repository will be preconfigured to use it aaccording to ExaGrid and Veeam recommendations (see http://go.veeam.com/rs/veeam/images/Best_Practices_and_Deployment_Veeam_and_ExaGrid.pdf ).

For example, ExaGrid recommends to use 1 job per repository. Thus, if you want to have parallel processing, create several repositories and set up 1 job per repository.

As a rule of thumb, the "landing zone"(which is the zone that will hold most recent set of data waiting to be deduplicated) should be at least 125% of a full backup so that each backup can fully be written there and processed. This will ensure that SureBackup, Instant VM Recovery and application items restore will be usable for the latest restore point without the read rehydration overhead.

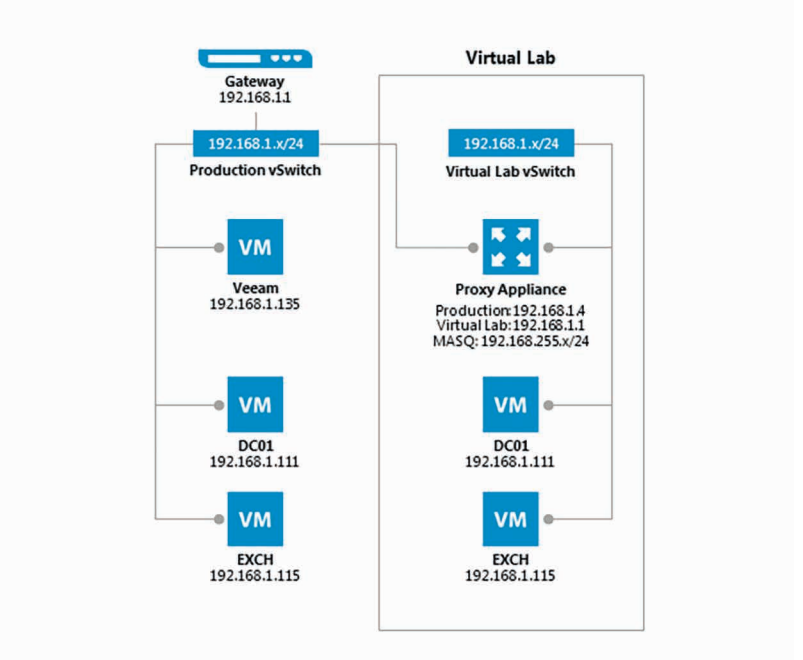
## Recommendations for HP StoreOnce

Selecting StoreOnce appliance as a repository will automatically change the jobs and repository settings to the recommended values. For more information, refer to storage vendor guidelines.

# vPower NFS and Virtual Lab

## Virtual Lab Appliance Overview

The Virtual Lab appliance operates as a gateway for the layer of network address translation (NAT) between the Veeam backup server and the isolated virtual machines in the Virtual Lab . It can also be used to provide access to other clients in the network, using static mapping. If VMs running in the isolated network need Internet access, the Virtual Lab appliance can act as an HTTP proxy server.




While a SureBackup job is running, static routes for the masquerade networks are automatically added (temporarily) to the routing table on the Veeam backup server. To review the routing table, open a command prompt on the Veeam backup server and enter:

```
route print -4
```

You may run this command before and after starting the SureBackup job to compare the differences.

Usually, the routes are added just after the Virtual Lab appliance has booted and has been initialized by the Veeam backup server. As static routes are added, this will ensure the Virtual Lab appliance is the first hop for all packets to the masquerade networks.

**Tip**

To skip extra network configuration, place the backup server and the Virtual Lab appliance in the same network subnet.

Check Veeam Backup & Replication documentation for configuration details:

- http://www.veeam.com/vmware_vpower_eval_guide_8_pg.pdf
- http://helpcenter.veeam.com/backup/80/vsphere/

# How SureBackup Job Works

This section describes the processes in a SureBackup job.

First, standard vSwitches and port groups, or port groups on existing Distributed vSwitches are created during the Virtual Lab creation wizard; then the process goes as described below.

## Booting Virtual Lab Appliance

1. Virtual Lab appliance configuration file is built and mapped to the Virtual Lab appliance as an ISO.
2. Virtual Lab appliance network interfaces are reconfigured for appropriate isolated networks.
3. The Virtual Lab appliance is powered on.
4. The SureBackup job waits for IP configuration to be published and stabilized through VMware Tools.
5. A static route for the configured masquerade networks is added dynamically to the routing table of the Veeam backup server. Those static routes are pointing to the IP address of the Virtual Lab appliance.

**Tip**

Earlier versions of Veeam Backup & Replication use VMware hardware version 4 for the virtual lab appliance. This limits the usable lab networks to 3. To add more networks/NICs, update Veeam to actual version and launch the Virtual Lab wizard again; this will re-create the Virtual Lab VM with the virtual hardware version defined in the VMX file: *%ProgramFiles%\Veeam\Backup and Replication\Backup\LiveCD\drv_va.vmx* Beginning with Veeam Backup & Replication 8 patch 1, the VMX file was updated to use virtual hardware v7. Another option is (for older Veeam versions) to update the Virtual Lab VM by VMware vSphere (Web) client to at least virtual hardware v7. After upgrading the virtual hardware version (by Veeam wizard or manually) you can start the wizard again and create up to 9 virtual lab networks.

## Booting Virtual Machines

1. Veeam publishes and registers VMs using Veeam vPower NFS from the repository containing the backup file.

**Note**

If the VMs to be booted in the Virtual Lab in an Application Group are replicas, this step will be skipped.

2. NICs are connected to port groups that are not configured on the virtual lab configuration, they will be disconnected automatically.

3. Veeam creates a snapshot for virtual machine to redirect write operations to the production datastore selected during the Virtual Lab configuration.

4. If the domain controller role is selected, registry settings are injected to ensure the NETLOGON service will not shutdown due to missing peer communication.

5. VMs are powered on.

6. During boot, VMware Tools announce IP configuration. The SureBackup job waits for this information to stabilize.

**Note**

If VMware Tools are not installed on the virtual machine, the job will wait for the duration of **Maximum allowed boot time** configured for such VMs. This will slow down SureBackup jobs significantly. Therefore, it is always recommended to install VMware Tools on a verified VM.

7. VM testing begins (if enabled for the Application Group):

    a) **VMware Tools heartbeat** is used for verifying that the VM OS is successfully started.
    b) **PING** tests are initiated based on the masquerading network configuration. The ping is sent from the Veeam backup server. Since the masquerade network is not part of the Veeam backup server's own broadcast domain, the packet is sent to the first hop matching this network. As the static route to the masquerade networking was added after the boot of the Virtual Lab appliance, this appliance will act as gateway between the two components: Veeam backup server and isolated virtual machine.
    c) **Application-specific testing** using scripts is enabled based on the roles assigned to a VM in the application group configuration. The built-in roles will check corresponding TCP ports for a given service, while additional testing is available for the SQL Server (see the next section). TCP requests are sent from the Veeam backup server, and the routing to the virtual machine is handled by the Virtual Lab proxy appliance.
    d) **CRC verification** is optionally available and is disabled by default. If enabled, it will ensure all content of the backup file is consistent with the hash values at the time they were written. This consistency check is using the CRC algorithm for hashing. Warning, this feature will read 100% of the data from the backup file.

8. Once all virtual machines within an application group have been successfully booted and verified, VMs from linked jobs may boot.

## Checking SQL Server Database Availability

With Veeam Backup & Replication v8, a new Visual Basic script has been shipped with the product to allow for testing whether all databases on a given instance are available. This script is available in the Veeam installation folder as the **Veeam.Backup.SqlChecker.vbs** file.

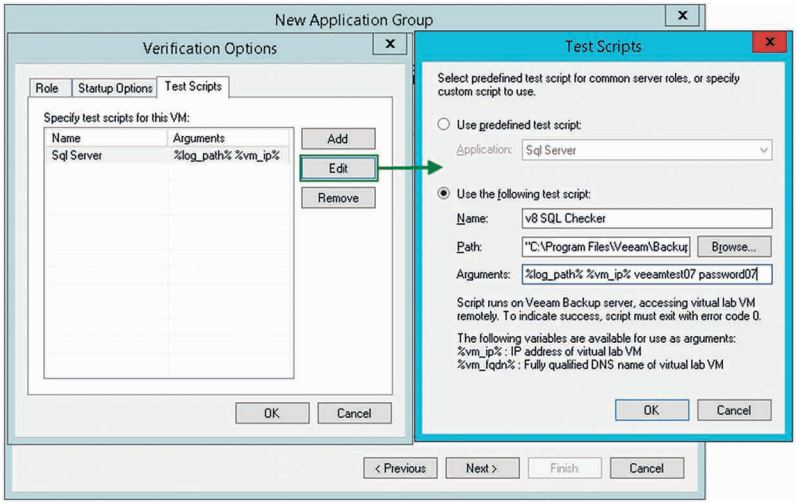
By default, the script tries to retrieve and check all instances (you can optionally specify the necessary instance). It enumerates all databases and checks if these databases are available, using the USE statement.

By default, the script will impersonate the service account under which the Veeam Backup Service is running (default in v8 is SYSTEM). The script also accepts two additional parameters to use SQL authentication instead.

**Important!**

To ensure proper authentication and access, it is required for the said user to have 'public' access to all databases.

The script includes detailed comments that describe the exact behavior.




## Creating Custom Roles

Though there are a number of built-in tests intended for application-level testing, you may need to develop additional scripts for testing proprietary applications. For that, carry out the following:

1. Open the Veeam installation folder and look through the *SbRoles* folder- all roles are defined in the XML files available in this folder.

2. To create custom roles, duplicate one of the above mentioned files and modify the **<Id>** tag using a UUID generator (such as https://www.uuidgenerator.net). Use this configuration file to specify the GUI settings.

When creating custom roles for Linux-based applications, you may need to execute code locally within the VMs. For that, use **plink.exe** shipped with the product and located in the *Putty* subfolder of the Veeam Backup & Replication installation directory.

When executing bash scripts locally on a Linux virtual machine using **plink.exe**, the exit codes are passed to the SureBackup job, enabling correct error reporting. If using **plink.exe** in combination with a SSH private key, you should connect manually (one time) to the VM via SSH using **putty.exe** to accept the target VM SSH fingerprint (otherwise, the SureBackup job will wait for this input and ultimately timeout).

**Note**

You can use **puttygen.exe** to create a private key.

Another option for testing service availability with *Veeam.Backup.ConnectionTester.exe* is decribed in http://www.veeam.com/kb1312.

## Troubleshooting Mode

If you need to troubleshoot Virtual Lab, it is recommended to start sessions in the Troubleshooting Mode. For that:

1. Open up **Statistics** for a SureBackup job.

2. Right-click any VM.

3. Select **Start**.

The SureBackup lab will now start in the troubleshooting mode, which means that errors will not cause the Virtual Lab to shut down immediately.

This opportunity is especially helpful during an implementation phase while measuring application boot times via vPower NFS, or implementing custom verification scripts. When you have finished troubleshooting, you can stop the SureBackup session manually.

**Tip**

ICMP traffic is blocked on all network interfaces of the Virtual Lab appliance, but you can PING the VMs by masquerade IPs. ICMP will be fully supported in one of the forthcoming Backup & Replication versions . See also http://forums.veeam.com/veeam-backup-replication-f2/sure-backup-ping-virtual-appliance-gateway-t26850-15.html#p149069.

Windows servers may change their network and firewall profiles to "Public". This may lead to application testing scripts not responding to TCP socket connections.

## Virtual Lab in Complex Environments

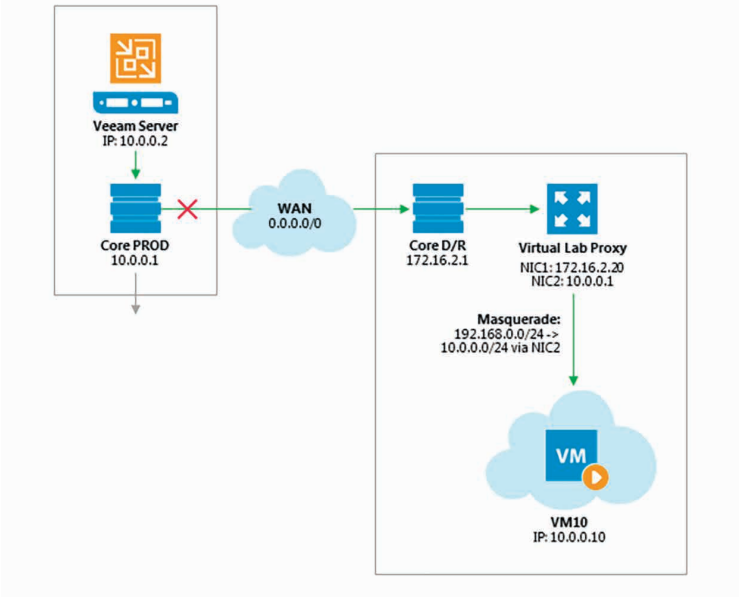
When using standard vSwitches in the VMware vSphere infrastructure, the Virtual Lab proxy appliance and the isolated networks will run on the same ESXi host. The reason is that standard vSwitches and their port groups are bound to one single host. Since the Virtual Lab port groups are isolated by nature, these should not be known to the core network in terms of vLAN tagging or routing.

When a distributed vSwitch (dvSwitch) is available, port groups can span multiple ESXi hosts. Distributed vSwitches are typically required when using Virtual Lab for replicas (SureReplica), as replicas will often span multiple hosts (vSphere Distributed Resource Scheduler (DRS) may also distribute VMs across multiple hosts within a cluster when they are booted).

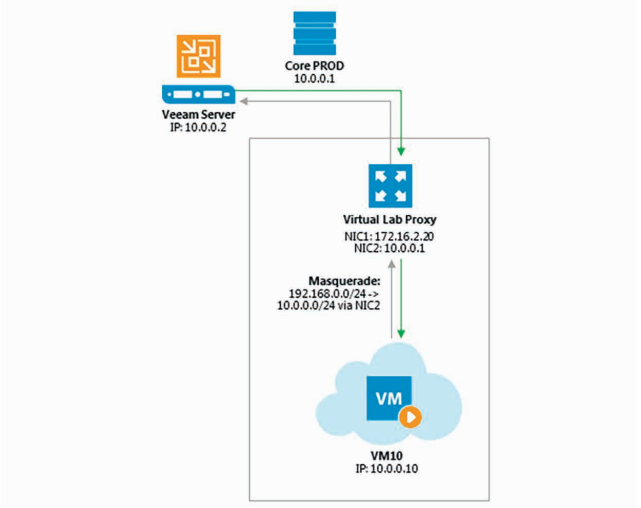
**Important!**

Please check the following help article and the links at the bottom of the webpage before you configure Virtual Labs together with distributed switches:http://helpcenter.veeam.com/backup/80/vsphere/index.html?surebackup_advanced_vlab.html.

Even in the environments where distributed vSwitch is available, make sure that the Veeam backup server and the Virtual Lab proxy appliance are placed in the same port group to prevent network packages (sent to the masquerading IP subnets) from being routed.




Most DR datacenters are configured with separate IP networks to allow for "active-active" configurations. In such cases, layer 3 (IP) and the requirement for routing will be introduced to establish communication between the production site and the DR site. For those scenarios, it is recommended to deploy a Veeam backup server in the DR site. This will help to get the Virtual Lab working and ensure correct one-click failover and failback handling if the production site becomes unavailable. Alternatively, you can specify an unused network segment so that it does not overlap with any segment used in production.



# WAN Acceleration

WAN acceleration was introduced as a key component of backup copy jobs in Veeam Backup & Replication v7. By combining multiple technologies such as network compression, multi-threading, dynamic TCP window size, variable block size deduplication and global caching, WAN acceleration provides sufficient capability whilst the required network bandwidth is dramatically reduced when performing backup copy and replication jobs. To determine whether WAN acceleration is necessary in the environment, it is important to understand what particular savings can be achieved.

## Determining Required Bandwidth

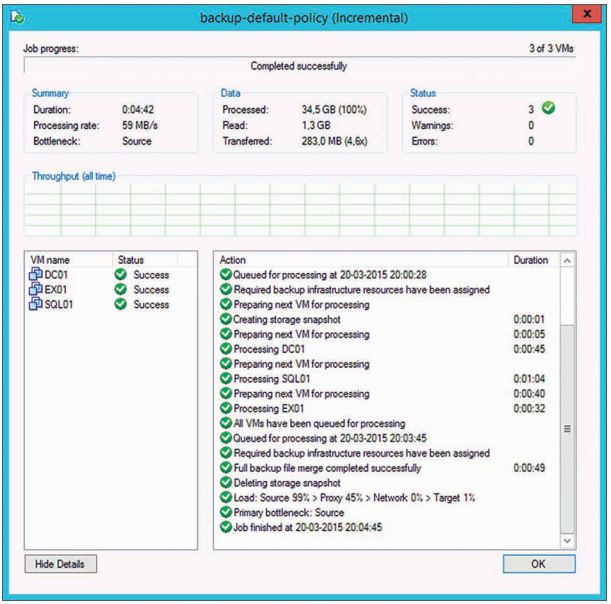
When using WAN acceleration on links with very poor bandwidth, you may have to manually seed the initial copy to the target. For more information, refer to the WAN Acceleration section of the Veeam Backup & Replication User Guide.

The WAN accelerator calculates its own digests based on the hashes of the blocks inside the storage which means that it reads data from the backup files while re-hydrating them on the fly.The WAN accelerator component will then re-process data blocks with much more efficient data deduplication and compression algorithms. This is the reason why the WAN accelerator consumes significant amounts of CPU and RAM resources.

To determine how much data has to be transferred over the WAN link with and without WAN acceleration enabled in backup copy job, you can compare the daily changes of the primary backup job statistics (as the same data is transported in a standard backup copy job without WAN acceleration) with the WAN accelerator-enabled backup copy job log and statistics.

### Analyzing Backup Job

During both full and incremental job sessions, three metrics are displayed in the session data: **Processed**, **Read** and **Transferred**. To better understand the difference between direct data transfer and WAN accelerated mode, examine the **Read** and **Transferred** values:

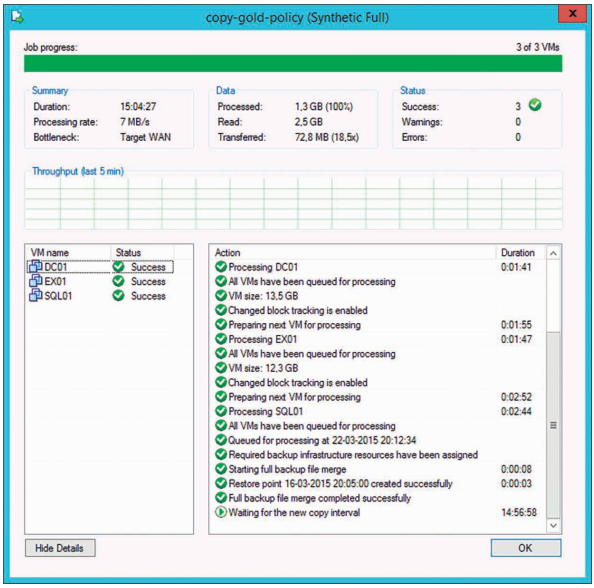



- **Read —** amount of data read from the production storage prior to applying any compression and deduplication. This is the amount of data that will be optimized by the WAN accelerator.

- **Transferred** — amount of data written to the backup repository after applying compression and deduplication. This is the amount of data that will be processed by the backup copy job running in Direct Transfer mode (without WAN acceleration), assuming all VMs from the backup job are included in the backup copy job.

## Analyzing Backup Copy Job

When analyzing a backup copy job, you can see the same metrics in the job session data: **Processed**, **Read** and **Transferred**. Comparing the backup copy job with WAN acceleration enabled and the backup job, it is possible to correlate the information in both outputs.




- The amount of **Processed** blocks in the backup copy job session is equal to the amount of **Read** blocks in the backup job session. This is the most important metric, as it is the amount of data that has to be processed by the WAN accelerator.

- The number of **Read** blocks for the backup copy job is typically higher than the amount of **Processed** - this is due to the backup copy job using a differing fingerprinting algorithm that work on a differing block size to the fingerprinting algorithm and block size used by backup jobs that created the original backup file. For that reason, this metric can be ignored.

- The amount of **Transferred** data is the amount of data actually transferred over the WAN link.

## Comparing Direct Mode with WAN Accelerated Mode

Consider that the savings rate (18.5x) displayed in the GUI will be based on **Processed** data ("re-hydrated" data blocks). In the example above, 283 MB would have been transferred over the WAN link in Direct Transfer mode, while only 72.8 MB were transferred after enabling WAN acceleration. The actual savings rate equals 3.9x in this relatively static demo infrastructure, whilst it would typically be significantly higher in real-life scenarios.

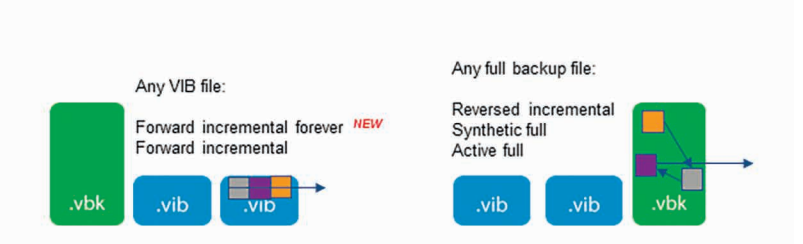
**Note**

Approximate savings ratio can be assumed as of 10x.

To calculate possible savings and needed bandwidth you may use the following calculator http://vee.am/bandwidth.

# Backup Mode Effect

When planning for WAN acceleration, review the backup mode used on the primary backup job. Some backup methods produce a random I/O workload on the source repository (as opposed to sequential I/O patterns in other backup modes). The methods of reading from source is illustrated by the figure below:




For example, forward incremental and forever forward incremental method will make backup copy jobs work much faster, as read operations will be sequential rather than random.

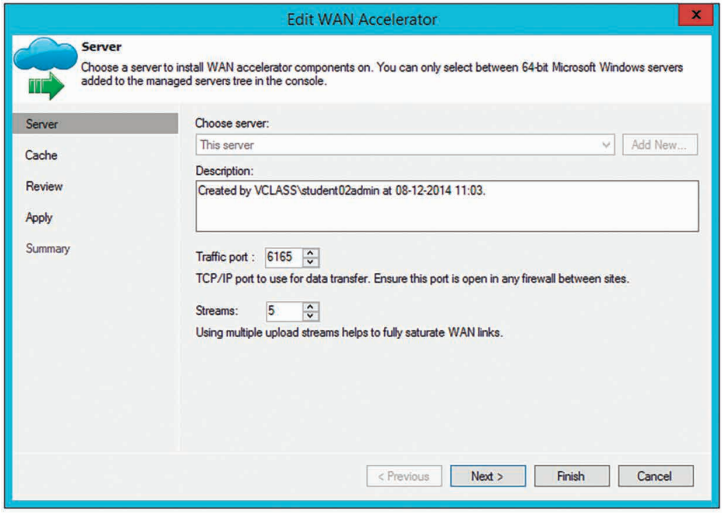
Though a workload penalty may not be significant, it can be a good idea to monitor the storage latency on the backup repository, especially if the reported bottleneck is *Source*. If the storage latency on the backup repository is high, it is recommended that you change the backup mode in order to increase the throughput of one pair of WAN accelerators.

# Configuration

## Source WAN Accelerator

At the first step of the WAN accelerator configuration wizard, you can change the default setting of five TCP threads. This setting applies to the source WAN accelerator, and is automatically accepted by the target WAN accelerator.

Veeam can support up to 100 simultaneous threads as throughput optimization and compensation for high latency or packet loss.



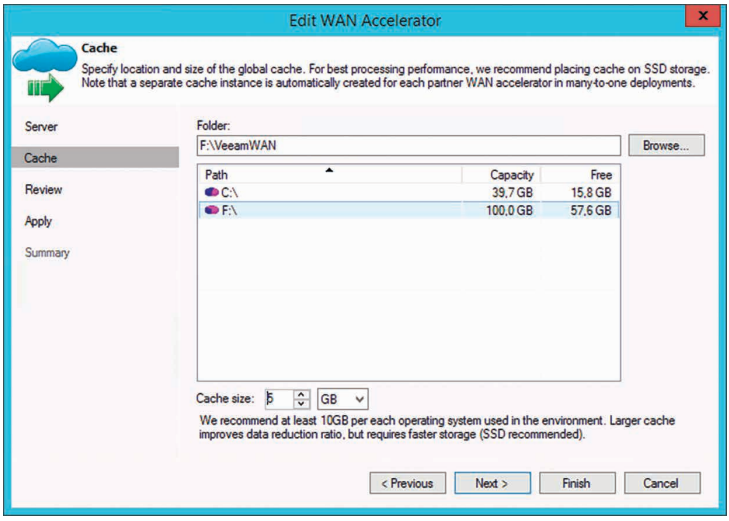

If the link has low latency and high bandwidth, the default setting (5 streams) may be enough to fully saturate it.

**Important!**

If the link is still not saturated, you can increase the number of streams accordingly.

**Tip**

To test different scenarios in the lab before deploying WAN acceleration, you can use a free WAN emulator (such as WANem). For more information, see this link.



When configuring the cache location for the source WAN accelerator, consider that the actual cache size is irrelevant on the source, as it is used only for digest files.

- If a WAN accelerator will be used for bi-directional acceleration (act as both source and target), follow the guidelines provided in the "Target WAN Accelerator" section below.
- To understand the storage requirements for storing digests, refer to the "Sizing" section below.

## Target WAN Accelerator

The following recommendations apply to configuring a target WAN accelerator:

- The cache size setting configured on the target WAN accelerator will be applied to the pair of WAN accelerators. This should be taken into account when sizing for many-to-one scenarios, as configuring 100 GB cache size will result in 100 GB multiplied by the number of pairs configured for each target WAN accelerator.

**Note**

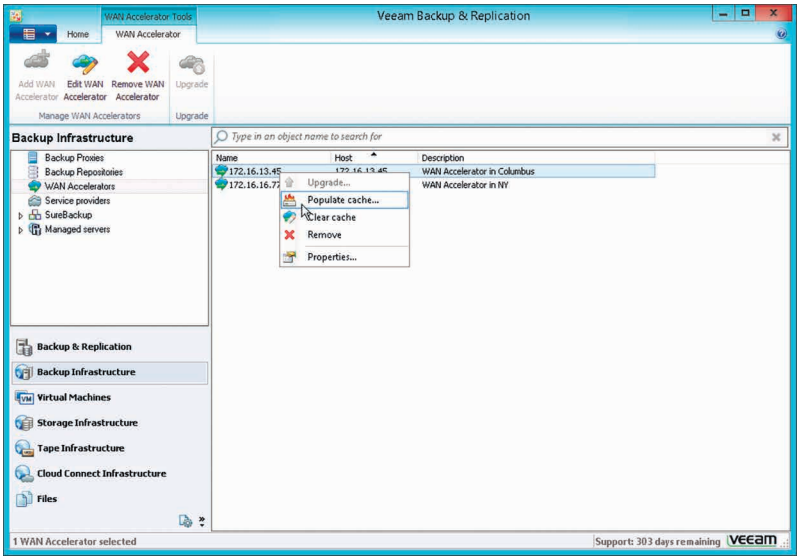
A pair of WAN accelerators means any source WAN accelerator paired with the target WAN accelerator.

- It is recommended to configure the cache size at 20 GB for each operating system processed by the WAN accelerator.

**Note**

All Linux operating systems are considered as one in terms of WAN accelerator sizing.

- Once the target WAN accelerator is deployed, the optimal option is to leverage the cache population feature implemented in Veeam Backup & Replication v8 (see this section of the User Guide for details). When using this feature, the WAN accelerator service will scan through selected repositories for protected operating system types.




- It is also possible to seed the initial copy of data to the target repository to further reduce the amount of data that needs to be transferred during the first run.

# Sizing

## Source WAN Accelerator

When configuring the WAN accelerator on the source side, consider that all VM disk data blocks are already in the source backup repository and they can simply be re-read from the source repository when needed. However, there are other files residing in the source WAN accelerator folder, and the file structure will be described in the next sections.

The I/O requirements for the source WAN accelerator are high: it is strongly recommended to deploy it on the fastest possible disk, as most of I/O will take place there.

The source accelerator will also consume a lot of CPU resources whilst applying heavy compression (also for decompressing before recompressing).

The optimal configuration is with 4 vCPU and 8 GB RAM.

**VeeamWAN\GlobalCache\src**

There is only a *data.veeamdrf* file located in the *\VeeamWAN\GlobalCache\src* folder. This file will be synchronized with the target WAN accelerator during the very first job run (or if the cache was manually cleared) to understand what data blocks are already cached in the target WAN accelerator. The size of this file is typically up to 2% of the configured target cache size; thus, it may take some time for the initial data transfer to begin.

**VeeamWAN\Digests**

On the source WAN accelerator, there are the VM disk digests that take up disk space. For each processed VM disk, a disk digest file is created and placed in *\VeeamWAN\Digests\<Job ID>_<VM ID>_<Disk ID>*.

**Note**

Traffic throttling rules should be created in both directions. See Network Traffic Throttling and Multithreaded Data Transfer: Veeam Backup Guide for VMware for more information.

## Calculating Source Side Disk size for WAN Accelerator

Consider that each digest file is up to 2% of its source VM disk size. This means, for example, that a 1.5 TB VM disk file can produce a digests file up to 30 GB in size.

**Note**

As the cache size on the source WAN accelerator will always be ignored, the digests file will be produced regardless of cache setting been configured. They may consume considerable disk space.

Additionally, plan for 10 GB of working space for payloads and other temporary files.

The disk I/O pattern on the source WAN accelerator is high and should be taken into account when planning for storage; it is recommended to deploy source WAN accelerators on the fastest storage available on that side.

## Target WAN Accelerator

Although a target WAN accelerator will consume less CPU resources than the source, the I/O requirements for the target side are high. For each processed data block, the WAN accelerator will update the cache file (if required), or it may retrieve the data block from the target repository (if possible). Tests show that for one-to-one deployments, there are no significant performance differences using SAS10k drives as storage for

the target WAN accelerator cache rather than flash storage. With multiple caches being updated in many-to-one deployments, it is recommended to use SSD or flash storage in such scenarios.

**VeeamWAN\GlobalCache\trg**

For each pair, there will be a subfolder in the *trg* directory, with a UUID describing which source WAN accelerator the cache is attached to. In each of those subfolders, the *data.blob* file containing the cache will be located. That file size corresponds to the setting configured in the management console.

**Note**

The *data.blob* file will exist for all connected source WAN accelerators.

**VeeamWAN\GlobalCache\temp**

When connecting a new source WAN accelerator, the *temp* folder will temporarily contain the *data.veeamdrf* file that is later transferred to the source containing the cache manifest.

## Calculating Target Side Disk Size for WAN Accelerator

Ensure that enough space has been allocated for global cache on the target WAN accelerator.

1. At least 10 GB per each different OS that is backed up. That is, if you plan to backup VMs running under Windows 8, Windows 2008 R2, Windows 2012 and RHEL 6 (four different operating systems), you will need at least **10 GB * 4 = 40 GB**

2. Plan for additional **20 GB** of working space for cache population, payload and other temporary files.

3. If the cache is pre-populated, an additional cache is created. This pre-populated cache will be used as template for all other connected sources.

**Example:**

Number of sources: 2
Configured cache size for four (4) operating systems
(see point 1 above): 40 GB
Suppose that you have a pre-populated cache of 40 GB -therefore, two sources will use this value as a template:

- Source A = 40 GB
- Source B = 40 GB

Total estimated space for target: 40 GB for pre-populated cache + 2*40 GB for two connected sources + 20 GB working space = 140 GB

## How Many WAN Accelerators to Deploy?

As the source WAN accelerator can only process one task at a time (one VM disk in a backup copy job or replication job), you may need to deploy multiple WAN accelerator pairs to meet the performance demands.

As the target WAN accelerator can handle multiple incoming streams (as described in the Many-to-One WAN Acceleration section of the User Guide), it is recommended to maintain a 4:1 ratio between the number of source WAN accelerators per target WAN accelerator.

This guideline is very much dependent on the WAN link speed. Many source sites with low bandwidth will create little pressure on the target WAN accelerator. So, for instance, in multiple ROBO configurations a 10:1 ratio can be considered.

If there are sites with very high bandwidth (such as datacenter-to-datacenter replication), they will produce a much more significant load on both the target WAN accelerator and the target repository due to the second data block lookup (for more information, refer to the User Guide).

**Note**

The secondary data block lookup is used, when a data block is not available in the WAN accelerator cache. When there is a WAN cache "miss", the secondary lookup for the same data block is performed on the target repository. If it is found here, it is read back to the WAN accelerator instead of re-transmitting over WAN.

Performance of the WAN accelerator was improved significantly in Veeam Backup & Replication v8. Assuming the source and target repositories can deliver the throughput required for the optimal processing rate, use the guidelines that follow.

**Note**

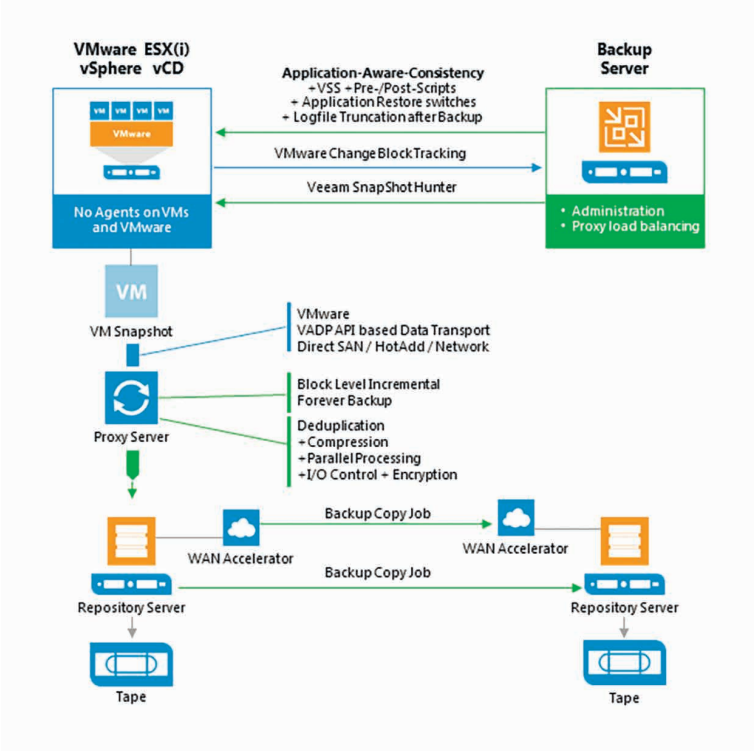
The numbers below are processing rates. The WAN link usage is dependent on the achieved data reduction ratio.

- Average throughput per target WAN accelerator: 500 Mbit/s (62.5 MB/s)
- Depending on the achieved data reduction rate (typically 10x), the transfer rate over the WAN link will vary.
    - If the processing rate is 62.5 MB/s, and the data reduction rate is 10x, then it is possible to sustain 6.25 MB/s (50 Mbit/s) over the WAN link.
    - If the WAN link has high bandwidth (above 100Mbps consider using backup copy jobs without WAN Acceleration. However, if you use WAN accelerators in that scenario, it may require deployment of multiple WAN accelerator pairs to fully saturate the WAN link.

# Tape Support

## Overview

The diagram below illustrates the main components and processes within the backup infrastructure when tape support is implemented within Veeam Backup & Replication deployment:



# Tape Device Connection

All connection types require driver installation.

**Note**

You can use generic drivers from Microsoft Windows, but they may not provide as high performance as the vendor's.

Separate drivers for tape drives and for tape media libraries should be installed.

| Connection Type | Compatibility |
|---|---|
| FC/SAS/SCSI/FCoE/InfiniBand/ other block technology to physical tape server | Works (as long as Windows drivers are available) |
| FC/SAS redirect to VM | Not supported by VMware and Veeam |
| iSCSI | Works (limited for physical tape libraries, mostly used for VTL) |
| 1 GbE FC/SAS to iSCSI converter | Works but the throughput is limited to 1GbE |
| 10 GbE FC/SAS to ISCSI converter | Works but very expensive |
| StarWind Tape Redirector | Works but not needed with v8, as Veeam introduced the tape server role |

# What Tapes are Supported?

**Supported:**

- LTO- 3 and higher.
- Unofficial tape compatibility list: http://forums.veeam.com/tape-f29/unofficial-tape-library-compatibility-list-t17488.html
- For VTLs, see the unofficial list above to check compatibility. In most cases VTLs can also be added as disk repositories (Linux or share) to Veeam Backup & Replication (which is likely more efficient and compatible with all VTLs).

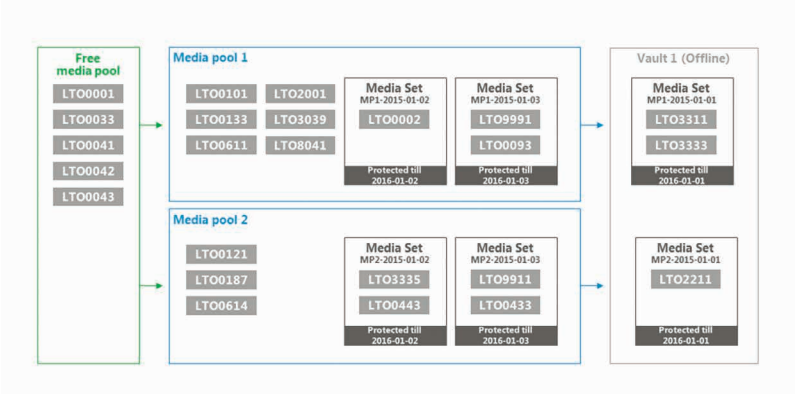
**Not supported:**

- IBM "Jaguar" Enterprise tape drive
- DLT or AIT tape drives

# Special Settings for Drivers

- **IBM drivers**: use "non-exclusive" driver setup and start it with adminstrative rights.
- **HP drivers**: these are not installable with **.exe** file on a VM (for example, to use with VTL). As a solution, run **.exe** and choose **Extract. Use Device Manager –> Update driver** and select the drivers for tape drives and (if you use HP/HP emulation tape library) for media changer.

# Media Pool – MediaSet - Vault




## Media Pool

A media pool simply defines a group of tapes managed by Veeam Veeam Backup & Replication. Media pools are of two kinds:

- **Service media pools** that are created and managed automatically. You cannot create them manually or modify their settings.
    - Empty media starts out in the **Free pool** indicating it's available for use in other pools.
    - Unknown media will be placed to the **Unrecognized pool** so that it is not overwritten.
    - Media with some content is placed into the **Imported pool**. You can view the contents and place them into the **Free** pool (overwrite) or leave in the **Imported pool** and use the data.
    - Exhausted or broken tapes are placed into the **Retired pool** and are not used further.

- **Custom media pools** that are created by the user. These are groups of media to which backup data can be written.
    - You can create as many custom media pools as needed.
    - Media can be assigned to a pool manually, or configured to be automatically assigned from the free pool.
    - Configure each pool settings according to the purpose of the pool, such as the overwrite protection period that is applied to all media within the pool.

### Media Set

A media set is a subset of a media pool that contains at least one backup. A new media set can be created for every backup, or on a time based schedule (i.e. weekly). If a media set contains at least one full backup, it is a self-sufficient restore point. It means that if you have all tapes from the media set at hand, you can be sure that restore will be successful.

### Media Vault

A media vault is used to organize offline media. For example, you have a service organization that transports the tapes to a safe at a bunker. You can name the vault accordingly and add some useful information in the description (phone number, place, etc.). When you need to transport physical tapes to the safe, add these tapes to the vault manually or set automatic export of offline tapes to a vault in the tape jobs or media pools properties.

## Veeam Backup Job and Tape

When you archive forward incremental or reverse incremental backups, a backup-to-tape job does not transform backups but transports them to tape as a copy.

- If you archive reverse incremental backups, the tape job will always copy the full backup.
- If you archive forward incremental backups, with or without active or synthetic full scheduled, the backup chain on tape will be simply a copy of the backup chain on disk.
- To archive forever forward incremental backup chains, schedule virtual full in the tape job. The tape job creates the full backup on the fly from a full and subsequent increment backups on disk.

When creating a backup-to-tape job, you will be asked what you want to copy to tape during the first run: only the last backup chain or all existing restore points.

In small environments, you are more likely to create a single media pool and a single tape job that uses the whole repository as source. Refer to Veeam Backup & Replication User Guide and Evaluators Guide for more examples.

## Sizing in Bigger Environments

Install Windows 2012 R2 on the tape server for best performance. Use the latest Veeam version and patch level as they often contain tape speed optimizations.

Perform a PoC to test speed (tape/disk). If you have no opportunity to test speed, assume that the lowest speed for backup to tape jobs with LTO5/6 is 40MB/s.

For each running backup-to-tape or file-to-tape job, one tape drive is used. You can use this for planning multitask tape scenarios, as such jobs can run simultaneously.

- To utilize multiple drive installations most efficiently, create one tape job per each drive. The tape job must use the backup repository as the source. Set up as many repositories as tape jobs/drives.
- Size repository so that the tape jobs have enough time to archive the new data within the backup window.

## Backup Copy Job and Tape

When the backup to tape job archives backup copy chains, it extracts data blocks and builds new backup files on tape.

The tape server implemented in v8 allows you to archive data to remote tape devices. If you locate a tape server on the backup copy job target side, you will be able to transport the data locally to tape. However keep in mind that the backup copy job and the tape job must be defined and run from the same Veeam backup server.

If you use a separate Veeam backup server to transport the data to tape, set up a file-to-tape job together with a PowerShell script, as described in these forum topics:

- http://forums.veeam.com/tape-f29/v7-stop-bcj-start-file2tape-startbcj-t23048.html?hilit=Tape
- http://forums.veeam.com/vmware-vsphere-f24/hp-data-protector-and-veeam-only-backup-latest-vbk-file-t17204.html

## 3rd Party Backup-to-Tape Software

As Veeam Backup & Replication tracks and orchestrates all backups written to tape, Veeam can recommend use of the native Veeam tape features (backup-to-tape and file-to-tape jobs).

However, in some situations you may want to use an existing library with non-LTO tapes, or you need to integrate Veeam Backup & Replication into an existing backup-to-tape software. Veeam backup files contain all information needed for restore (e.g. deduplication information, VM metadata, etc.), and you can use the existing backup-to-tape solution to bring the Veeam backup files on tape.

Veeam can import and read tapes that use the native Microsoft Tape Format.

## Tape Encryption

Veeam uses hardware encryption if it is provided by the tape device and enabled in Veeam Backup & Replication. Tape library should work in the application-managed encryption mode.

If the hardware encryption is not supported by the tape device, then 256 AES software encryption is used.

When archiving data, Veeam generates a user key which is stored with data on tape. If you restore data using another Veeam backup server, provide the password or utilize the password loss prevention functionality supported by Veeam Backup Enterprise Manager. See the Veeam Backup & Replication User Guide for more information.

If the hardware encryption option is used, and you archive to tape Veeam backups that are already encrypted on disk, they will be encrypted twice. If you restore such backups with double encryption on the same Veeam backup server they will be decrypted automatically. To decrypt on another Veeam backup server, you will need to enter the two passwords accordingly.

# Veeam Explorers

For recommendations on resource planning for Veeam Explorers – the tools that support application item-level restore – please refer to the corresponding User Guide (in particular, to "Before You Begin" sections).

**Note**

Some recommendations for the staging system (if required for Veeam Explorer tool) can be also found in the «Veeam Backup Server» and «Veeam Backup & Replication Database» sections of this document.

If you have special features/enhancements/configuration settings on the production Microsoft SQL and/or Microsoft SharePoint server to be protected with Veeam, these custom settings should be implemented on the staging SQL Server, too.

For Microsoft Active Directory, also check the tombstone lifetime settings, as described in Veeam Explorers User Guide at Veeam Help Center (http://helpcenter.veeam.com/backup/80/explorers/index.html?vead_recommendations.html).

# Interaction with vSphere

Veeam Backup & Replication interacts heavily with the vSphere infrastructure. Much of the implementation success depends on the performance and stability of this environment. In this section, we will discuss those interactions and note the items that should be considered for a successful implementation.

While it is possible to connect a Veeam Backup & Replication server directly to ESX(i) hosts, this section assumes a vSphere environment with at least one vCenter Server, and that the backup server is integrated at the vCenter Server level, as this is the best practice configuration in almost all use cases.

## vCenter Server

One of the most critical components of any vSphere environment is the vCenter Server. This server provides a single view of the entire virtual environment, and a central point of management. Veeam Backup & Replication communicates with the vCenter Server in many operations. For this reason, fast and stable communication between Veeam Backup & Replication and the vCenter Server is critical to achieving a stable backup environment.

Consider some important factors:

- Problems with connectivity to the vCenter Server is one of the top reasons for failed Veeam jobs. Having a well-performing vCenter Server with reliable connectivity will mitigate this issue and provide a strong backbone for a reliable backup infrastructure.
- The vCenter Server must be reliable and always available when backup jobs are running. It must be able to answer queries and perform actions in a reasonable amount of time. If the vCenter Server performs poorly during normal operations, this should be corrected prior to implementing Veeam Backup & Replication.
- For larger environments, with many concurrent jobs, especially jobs that run at short intervals, such as near-CDP, the load on the vCenter Server can be significant. The vCenter Server must be able to handle increased transactional workload to prevent random job failures due to command timeouts.
- The backup server must have reliable network connectivity to the vCenter Server. It is generally suggested that the backup server is placed in close logical proximity to the vCenter Server, but this is not always the best deployment option. In cases where the backup server and vCenter Server must be deployed across a distance, the only real requirement is that this connection is consistent and reliable.
- When maintenance is being performed on the vCenter Server, best practice would dictate that all Veeam Backup & Replication jobs should be idle, and the Veeam Backup Service should be stopped. This includes applying Windows updates, vCenter Server patches and upgrades, or any maintenance that would require the vCenter service to be restarted or the system rebooted.

# Impact of Snapshot Operations

To create VM backups, Veeam Backup & Replication leverages the VMware vSphere snapshot functionality. When Veeam Backup & Replication begins the backup of a VM, it communicates with vSphere to request a snapshot of the VM, and after the backup of the VM is complete, Veeam requests that vSphere remove the snapshot (with the exception of backup jobs leveraging Backup from Storage Snapshots). The creation and removal of snapshots in vSphere creates a significant impact on the environment that must be taken into account. This section will describe various factors that should be considered regarding this process, and offer several techniques to minimize the impact of snapshot operations.

As a concept, VMware vSphere snapshots are a simple technology. A VM generally contains at least one virtual disk, which is represented by a VMDK file. When a snapshot is taken, VMware vSphere continues to read blocks from the file as normal. However, for any new blocks that are written to the disk, these writes are redirected to a new "thin" VMDK file called the delta file.

Since the original VMDK file is only being used for reads, it provides a consistent view of the blocks that made up the VM at the time the snapshot was taken. This allows Veeam Backup & Replication to read this base disk as a consistent image for backup and replication functions. When the snapshot is removed, the blocks that were written to the delta file are read and written back into the original VMDK, and finally the delta file is discarded.

As Veeam Backup & Replication leverages the snapshot technology for performing backups, you should ensure it is possible to snapshot the virtual machine disks, since there are certain configurations that do not support snapshots. To identify VMs that do not support snapshots, see VMware KB article 1025279 ; you can also use Veeam ONE assessment reports to automatically detect them before starting Veeam Availability project.

As with many things in technology, although the concept is simple, the actual implementation is a little more complex. The following section is a quick look at the impact of various operations on the VM and underlying infrastructure.

## Snapshot Creation

The actual operation of creating a snapshot generally has only a minor impact: the snapshot file has to be created, and there is a very short "stun" of the VM. This "stun" is generally short enough (typically, less than 1 sec), so it is rarely an issue except for the most time-sensitive applications.

**Note**

Veeam Backup & Replication leverages a standard VM snapshot for the backup process. These VMware snapshots have a single file size (including snapshot size) limitations. For normal snapshot operations with non-upgraded VMFS versions, try to keep the size of the VMDK disk under 1.98 TB, otherwise snapshot creation may fail due to known vSphere limitations. However, actual VMFS versions are able to snapshot bigger VMDK disks, but keep in mind that the overall usable size is the sum of data + snapshot data. For details, see VMware KB article 1012384.

The default number of concurrently open snapshots per datastore in Veeam Backup & Replication v8 is 4. This behavior can be changed by creating the following registry key:

- Path: **HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication**
- Key: **MaxSnapshotsPerDatastore**
- Type: **REG_DWORD**
- Default value: **4**

## Snapshot Open

Simply having a snapshot open for a running VM involves some performance penalty on the VM, the ESX(i) host and the underlying storage. The host has to track the I/O, split writes to the snapshot file and update the snapshot file metadata. This overhead, in turn, impacts the guest (primarily, with slower I/O).

This is generally most notable for VMs with significant write load, and has less impact on read performance.

From the storage perspective, VMs running with an open snapshot require additional space to store the snapshot data, and additional I/O load on the datastore. This is generally more noted on systems with significant write I/O load.

**Note**

Refer to VMware Knowledge Base article at www.kb.vmware.com/kb/1035550 for information on vMotion and Storage vMotion processes performed with open snapshots.

## Snapshot Removal

Snapshot removal is the step with the highest impact from the performance perspective. I/O load increases significantly, due to the extra R/W operations required to commit the snapshot blocks back into the original VMDK. This eventually leads to the VM "stun" required to commit the final bits of the snapshot. The "stun" is typically a short pause usually only a few seconds or less, when the VM is unresponsive ("lost ping"), while the very last bits of the snapshot file are committed.

VMware vSphere uses the "rolling snapshot" method to minimize the impact and duration of the stun, as described below:

1. The host takes a second, "helper", snapshot to hold new writes.
2. The host reads the blocks from the original snapshot and commits them to the original VMDK file.
3. The host checks the size of the "helper" snapshot. If the size is over the threshold, step 1 is repeated.
4. Once all helper snapshots are determined to be under the threshold size, vSphere "stuns" the VM and commits the last bits of the snapshot.

This "stun" period can be less than one second for small VMs with light load, or several seconds for larger VMs with significant load. To external clients, this small stun looks like the server is busy and thus might delay a response for a few seconds. However, applications that are very sensitive to delays may experience issues with this short period of unresponsiveness.

For explanation of snapshot removal issues, see VMware KB article 1002836.

## How to Mitigate?

To mitigate the impact of snapshots, consider the following recommendations:

- **Minimize the number of open snapshots per datastore**. Multiple open snapshots on the same datastore are sometimes unavoidable, but the cumulative effect can be bad. Keep this in mind when designing datastores, deploying VMs and creating backup and replication schedules. Leveraging backup by datastore can be useful in this scenario.
- **Consider snapshot impact during job scheduling.** When possible, schedule backups and replication job during periods of low activity. Leveraging the Backup Window functionality (see the corresponding setting on the **Schedule** tab of the job wizard) can keep long-running jobs from running during production.
- **Use the vStorage APIs for Array Integration (VAAI) where available.** VAAI can offer significant benefits:
    - Hardware Lock Assist improves the granularity of locking required during snapshot growth operations, as well as other metadata operations, thus lowering the overall SAN overhead when snapshots are open.

- VAAI in vSphere 5.x offers native snapshot offload support and should provide significant benefits once vendors release full support.
- VAAI is sometimes also available as an ESXi plugin from the NFS storage vendor.

- **Design datastores with enough IOPS to support snapshots.** Snapshots create additional I/O load and thus require enough I/O headroom to support the added load of snapshots. This is especially important for VMs with moderate to heavy transactional workloads. Creating snapshots in VMware vSphere will cause the snapshot files to be placed on the same VMFS volumes as the individual VM disks. This means that a large VM, with multiple VMDKs on multiple datastores, will spread the snapshot I/O load across those datastores. However, it actually limits the ability to design and size a dedicated datastore for snapshots, so this has to be factored in the overall design.

**Note**

This is the default behavior that can be changed, as explained in the VMware Knowledge Base: http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1002929

- **Allocate enough space for snapshots.**

  VMware vSphere 5.0 puts the snapshot VMDK on the same datastore with the parent VMDK. If a VM has virtual disks on multiple datastores, each datastore must have enough space to hold the snapshots for their volume. Take into consideration the possibility of running multiple snapshots on a single datastore. According to the best practices, it is strongly recommended to have 10% free space within a datastore for a general use VM, and at least 20% free space within a datastore for a VM with high change rate (SQL server, Exchange server, and others).

**Note**

This is the default behavior that can be changed, as explained in the VMware Knowledge Base: http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1002929

- **Watch for low disk space warnings.**

  Veeam Backup & Replication warns you when there is not enough space for snapshots. The default threshold value for production datastores is 10 GB. Keep in mind that you must increase this value significantly if using very big datastores (vSphere 5 and later). You can increase the warning threshold in the backup server options, in the Veeam Backup & Replication UI.

You can also create a registry key to prevent Veeam Backup & Replication from taking additional snapshots if the threshold is breached:

- Path: **HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication**
- Key: **BlockSnapshotThreshold**
- Type: **DWORD**
- Default value (in GB): **2**

This threshold will be added in Veeam Backup & Replication options in one of the next product updates.

**Tip**

Use the "Veeam ONE Configuration Assessment Report" to detect datastores with less than 10% of free disk space available for snapshot processing.

- **Enable Veeam parallel processing.**

  Parallel processing tries to back up all VM disks that belong to a single VM at the same time. This reduces snapshot lifetime to the minimum. Since v7, this option is enabled by default. If you upgraded from v6.5 or earlier versions, you have to enable this option explicitly in the backup server options.

- **Tune heartbeat thresholds in failover clusters.**

  Some application clustering software can detect snapshot commit processes as failure of the cluster member and failover to other cluster members. Coordinate with the application owner and increase the cluster heartbeat thresholds. A good example is Exchange DAG heartbeat. For details, see Veeam KB Article 1744.

## Considerations for NFS Datastores

Backup from NFS datastores involves some additional consideration, as the only available transport modes are the Network mode (NBD) and the Virtual Appliance mode (Hot-Add). Datastores formatted with the VMFS file system have native capabilities to determine which cluster node is the owner of a particular VM, while VMs running on NFS datastores rely on the LCK file that resides within the VM folder.

During hot-add operations, the host on which the hot-add proxy resides will temporarily take ownership of the VM by changing the contents of

the LCK file. This may cause significant additional "stuns" to the VM. Under certain circumstances, the VM may even end up being unresponsive. The issue is recognized by VMware and documented in http://kb.vmware.com/kb/2010953.

**Note**

This issue does not affect Veeam Backup from Storage Snapshots on NetApp NFS datastores.

To mitigate this issue, ensure that proxies running in the Virtual Appliance mode (Hot-Add) are on the same host as the protected VMs.

To give preference to a backup proxy located on the same host as the VMs, you can create the following registry key:

- Path: **HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication**
- Key: **EnableSameHostHotAddMode**
- Type: **DWORD**
- Default value: **0** (disabled)

  **1** – when proxy A is available on the same host, Veeam Backup & Replication will leverage it. If proxy A is busy, Veeam Backup & Replication will wait for its availability; if it becomes unreachable for some reason, another Hot-Add proxy (proxy B) will be used. **2** - when proxy A is available on the same host, Veeam Backup & Replication will leverage it. If proxy A is busy, Veeam Backup & Replication will wait for its availability; if it becomes unreachable for some reason, Veeam Backup & Replication will switch to NBD mode.

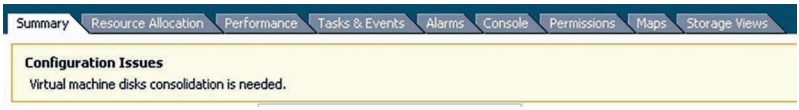
This solution will typically result in deploying a significant number of proxy servers, and may not be preferred in some environments. For such environments, it is recommended switching to Network mode (NBD).

## Snapshot Hunter

One of the new features in Veeam Backup & Replication v8 is Snapshot Hunter. At Veeam Support, one of the most commonly raised support cases was for orphaned snapshots. Orphaned snapshots were caused by VMware's own failed snapshot commit operations due to unreleased VMDK file locks during VDDK operations. Veeam uses the VMware standard VM snapshot processing for backup and replication processes, so although Veeam was not the origin of the orphaned snapshots, as Veeam uses VMware snapshots, Veeam is often seen as a root cause as this

issue was only discovered when a backup failed.

If not monitored appropriately, VMware orphaned snapshots can cause many unexpected problems. The most common problems are overfilled VM datastores, or snapshots growing so large they are impossible to commit. This is a well-known VMware vSphere issue described in VMware KB article 1007814. The only way to manually remediate this issue is cloning the VM and performing a new full VM backup.




Veeam Snapshot Hunter automatically detects VMs with the configuration issue "Virtual machine disks consolidation needed". Prior to performing backup of such VMs, Veeam Backup & Replication will trigger disk consolidation (provided that the datastore performance threshold specified in the Backup I/O Control settings is not exceeded).

Snapshot Hunter will attempt consolidation eight (8) times. If consolidation fails after all retries, Veeam Backup & Replication will send an e-mail with a warning.

You can view information on the Snapshot Hunter sessions on the **History > System** view in VeeamBackup & Replication console.

**Note**

Currently, the default behavior of Snapshot Hunter cannot be changed. As Snapshot Hunter will automatically retry consolidation up to eight times, it may be inappropriate for some VMs that require planned downtime to consolidate the snapshot manually. Such VMs should be excluded from backup or replication jobs until the orphaned snapshots are manually removed.

If you are evaluating Veeam Backup & Replication, use the Infrastructure Assessment Reports included in Veeam Availability Suite to identify VMs with snapshots that can be affected by automatic snapshot consolidation.

## Backup I/O Control

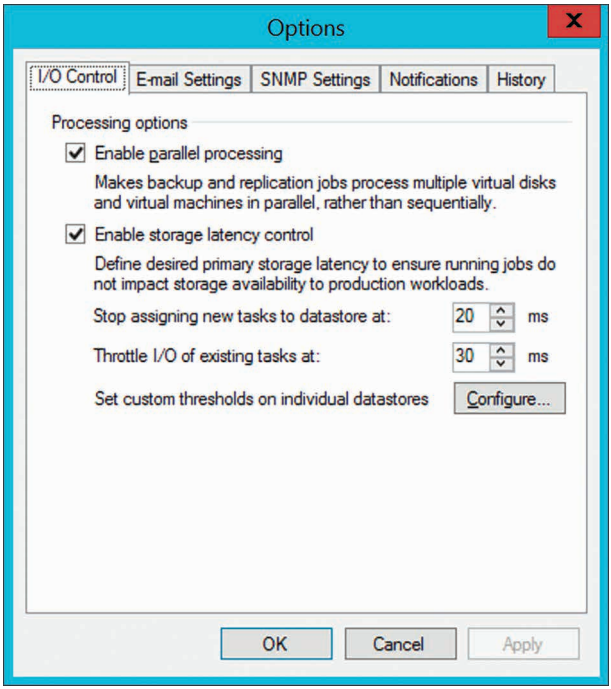
Backup I/O Control is a new feature of Veeam Backup & Replication v8. One question that often arises during the development of a solid availability design is how many proxy servers should be deployed. There must be a balance between the production infrastructure performance (as you must avoid overloading production storage), and completing backup jobs in time.

Modern CPUs have many physical cores and can run many tasks simultaneously. The impact of having many proxy servers reading data blocks from the production storage at a very high throughput may be negative. With this in mind, many businesses avoided running backup or replication jobs during business hours to ensure good response time for their end users. Backup I/O Control was implemented to help avoid this issue.

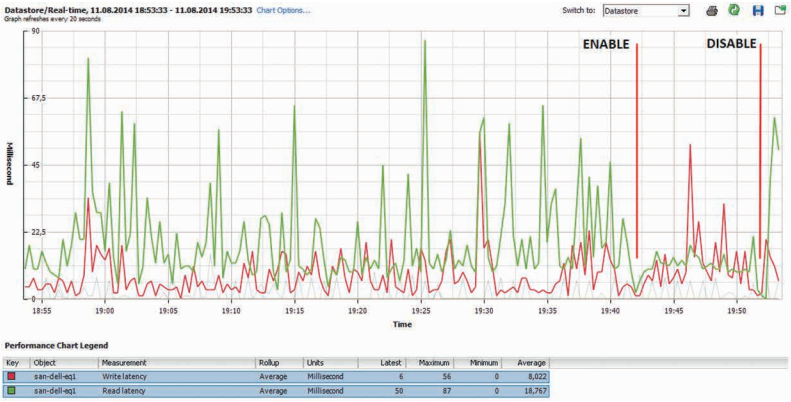
When Backup I/O Control is enabled, it monitors the storage read latency on the production datastores using real-time metrics from the hypervisor. By default, metrics from the hypervisor are collected every 20 seconds. These settings are inherited from vSphere.

The first Backup I/O Control threshold Stop assigning new tasks to datastore at puts a limitation on assigning new tasks (one task equals one VM disk). If the latency for a particular datastore is exceeded, no more proxy tasks will be assigned to it, until the latency drops below the threshold.

If limiting the number of tasks assigned to the datastore is not sufficient, Backup I/O Control will throttle the throughput for existing tasks according to the second threshold Throttle I/O of existing tasks at.



The results of enabling Backup I/O Control are very easy to review using the vSphere Client.




## When to Use?

Backup I/O Control provides a smart way to extend backup windows or even eliminate backup windows, and run data protection operations during production hours.

When Backup I/O Control is enabled, Veeam Backup & Replication measures the storage latency before processing each VM disk (and also during processing, if **Throttle I/O of existing tasks at** setting is enabled). Furthermore, if the storage latency for a given datastore is already above the threshold, committing VM snapshots can be delayed. In some environments, enabling Backup I/O Control will reduce the overall throughput, as latency increases during the backup window.

However, in most environments having this feature enabled will provide better availability to production workloads during backup and replication. Thus, if you observe performance issues during backup and replication, it is recommended to enable Backup I/O Control.

Backup I/O Control is available in Enterprise and Enterprise Plus editions. The Enterprise Plus customers are offered better granularity, as they can adjust latency thresholds individually for each datastore. This can be really helpful in infrastructures where some datastores contain VMs with latency-sensitive applications, while latency thresholds for datastores containing non-critical systems can be increased to avoid throttling.

# vCenter Server Connection Count

If you attempt to start a large number of parallel Veeam backup jobs (typically, more than 100, with some thousand VMs in them) leveraging the VMware VADP backup API, you may face two kinds of limitations:

- Limitation on vCenter SOAP connections
- Limitation on NFC buffer size on the ESXi side

All backup vendors that use VMware VADP implement the VMware VDDK kit in their solutions. This kit provides standard API calls for the backup vendor, and helps to read and write data. During backup operations, all vendors have to deal with two types of connections: the VDDK connections to vCenter Server and ESXi, and vendor's own connections. The number of VDDK connections may vary for different VDDK versions.

If you try to back up thousands of VMs in a very short time frame, you can run into the SOAP session count limitation. For example, in vSphere 5.1 the default maximum number of sessions is 500. If you hit this limitation, you can increase the vCenter Server SOAP connection limit from 500 to 1000. For details, see http://kb.vmware.com/kb/2004663.

You can also optimize the ESXi network (NBD) performance by increasing the NFC buffer size from 16384 to 32768 MB and reducing the cache flush interval from 30s to 20s. For details, see VMware KB article 2052302.

In the current version, Veeam's scheduling component does not keep track of the connection count. For this reason, it is recommended to periodically check the number of vCenter Server connections within the main backup window to see if you can possibly run into a bottleneck in future, and increase the limit values on demand only.

# Security

When connecting Veeam Backup & Replication to the vCenter Server infrastructure, you must supply credentials that the backup server will use to communicate with the vCenter Server.

The features that Veeam provides, such as backup, restore, replication, and SureBackup, interact with vSphere at the fundamental level. Thus, certain permissions are required to take snapshots, create VMs, datastores, and resource groups. Because of this level of interaction, it is generally recommended that Veeam Backup & Replication uses an account with full administrative permissions.

However, in some environments full administrative permissions is not desirable or permitted. For those environments, Veeam has identified the minimum permissions required for the various software functions. Review the "Required Permissions" document and configure the account used by Veeam Backup & Replication to meet these requirements.

You can also leverage security to restrict the part of the environment that the backup server can "see". This can have multiple benefits beyond security in that it lowers the time required to parse the vCenter Server hierarchy and reduces the memory footprint required to cache this information. However, care must be taken when attempting to use this level of restriction, as some permissions must be provided at the very top of the vCenter Server tree.

For a detailed description of accounts, rights and permissions required for Veeam Backup & Replication operations, see the "Required Permissions" document.

# JOB CONFIGURATION

## Veeam Backup Methods

Veeam Backup & Replication stores backups on disk using a simple, self-contained file based approach. However, there are several methods available for exactly how those files are created and stored on the file system. This section will provide an overview of these methods, their pros and cons, as well as recommendations on use cases for each one.

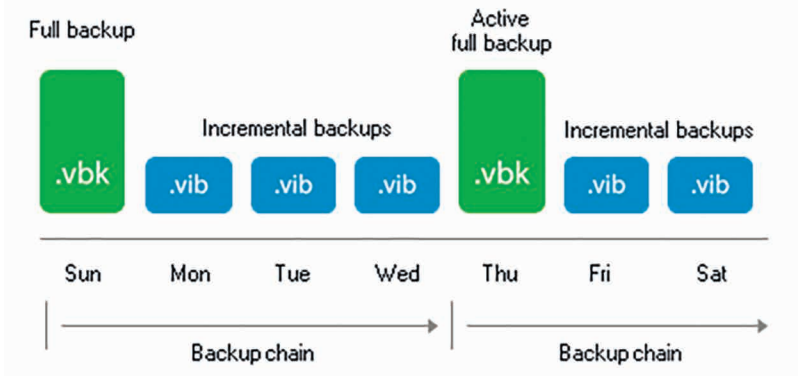
## Forward Incremental

The forward incremental backup method is the simplest and easiest to understand; it generally works well with all storage devices although it requires more storage space than other backup methods due to the fact that it requires the creation of periodic full backups (either using active or synthetic backups), typically scheduled weekly. This is necessary because the incremental backups are dependent on the initial full backup; thus, older backups cannot be removed from retention chain until a newer backup chain is created. When a new full backup arrives, a new chain is started, and the old backups can be removed once the new chain meets the retention requirements.

### Active Full Backups

The first time a job is run it always performs an active full backup. During this process the VM is read in full, and VM data is stored (typically compressed and deduped) into a Veeam's full backup file (.VBK).

Each time an active full is performed (either on schedule or by manual launch of the Active Full command), a new .VBK file is created once again by reading all data from the source VM. Increments are stored in Veeam's incremental backup files (.VIB).



This approach provides for a very reliable backup process, since every time a full is created, each VM is read in its entirety.

### I/O Impact of Active Full

When creating an active full, the I/O load on the backup storage is mainly sequential writes, which generally provides good performance for most storage types. However, all the data (not just the changes) has to be copied from the production ESX hosts, and this will increase the time a VM snapshot remains open (see also the "Impact Snapshot Operation" section of this guide).

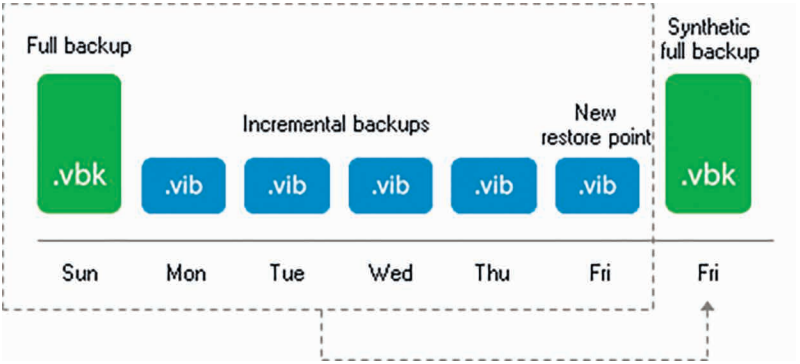
### Recommendations on Usage

Forward incremental backup provides good performance with almost any storage. It can be used in any case where plenty of repository space is available and the backup window allows enough time to support reading the source data in full.

| Use | **D**on't Use |
|---|---|
| Recommended for deduplication appliances that use SMB or NFS protocols. | When backup window does not allow enough time for re-reading all of the source VM data. |
| On storage systems that use software or non-caching RAID hardware such as many low-end NAS devices. | For large or performance sensitive VMs where re-reading the data can have a negative impact on the VMs performance. |
| Provides highest level of backup chain consistency since each new chain is populated by re-reading VM source data. | |

## Synthetic Fulls

Synthetic fulls take all the data in the previous chain (full .VBK and incremental .VIB files) and the incremental changes from production (using CBT if available), and create a new full backup using the existing data on the backup storage.




If a synthetic full is scheduled, when the job runs, it first creates a normal incremental backup – to collect the most recent changes.

After the job completes this phase, a synthetic full process is stared. It reads the most recent version of each block for every VM in the job from the appropriate backup file and writes those blocks to a new VBK file. This is how a new full backup is created.

### I/O Impact of Synthetic Full

Synthetic full creation is an I/O intensive process on the repository. Since the process reads individual blocks from the various files in the chain and writes those blocks to the VBK, it creates a roughly 50/50 read/write mix. The processing speed is limited by the IOPS and latency profile of the repository storage, so it can take significant amount of time. However, there is no impact on the source storage during this time as I/O occurrs only in the repository.

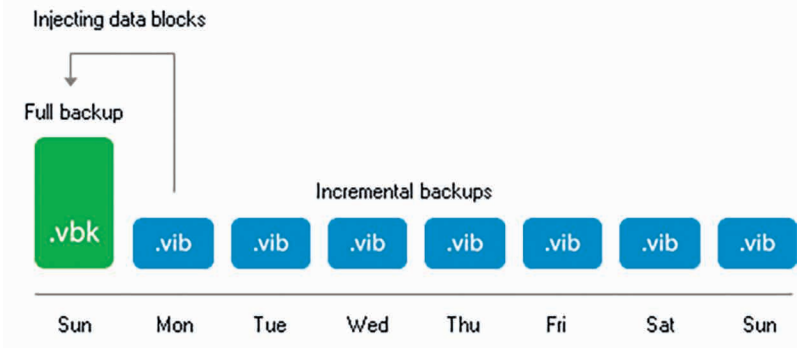
#### Recommendations on Usage

Due to the way this synthetic full process works, having many smaller backups jobs with fewer will perform synthetic full processing faster than having very large backup jobs with many VMs. Keep this in mind when setting up jobs that will use this method.

| Use | Don't Use |
|---|---|
| Recommended for use when repository storage uses fast disk with caching RAID controllers using large stripe sizes. | Small NAS boxes with limited spindles that depend on software RAID. |
| Can be used for deduplication appliances that support synthetic offload (currently only Data Domain with DDBoost). | Deduplication appliances that use SMB or NFS protocols. |

# Forever Forward Incremental

Forever forward incremental method keeps one full backup file (VBK) on disk, and then only incremental backups (VIBs) afterwards. This method allows backups space to be utilized efficiently, as there is only a single full backup on disk due to a merge process. This process is launched when the retention setting is met. It takes the oldest incremental backup and writes those blocks into the VBK, moving the VBK forward.




## I/O Impact of Merge Process

The merging process is performed at the end of the backup job once the retention for the job has been reached. This process will read the blocks from the oldest incremental backups (VIB file) and write those blocks into the VBK file thus it creates a 50/50 read-write mix on the target storage. The time required to perform the merge will be based on the size of the incremental data and the random I/O performance of the underlying storage.

## Recommendations on Usage

The primary advantages of using forever forward incremental backup method are the space savings and very fast, incremental backups. However, the tradeoff is the time required for the merge process. This process can take from minutes to hours depending on the amount of incremental change that the job has to process. However, this merge process impacts only the target storage thus the impact on production is quite low.

Like with synthetic full, it can be recommended to have many smaller jobs with a limited number of VMs (20-30), which can significantly increase the performance of synthetic merge process. Very large jobs with more than100 VMs can experience significant increase in time due to extra metadata processing.

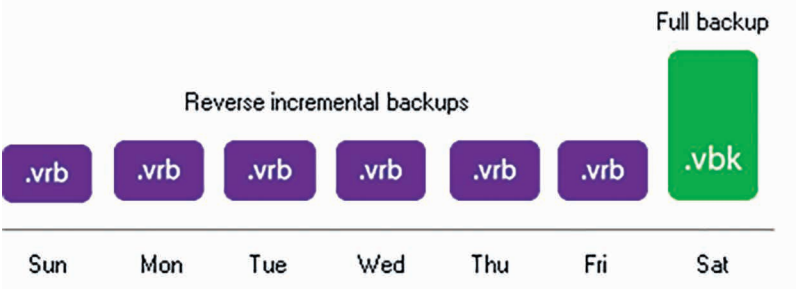
| Use | Don't Use |
|---|---|
| Recommended for use when repository storage uses fast disk with caching RAID controllers using large stripe sizes. | Small NAS boxes with limited spindles that depend on software RAID. |
| Excellent for low change rate VMs, especially large VMs with limited daily change. | Deduplication appliances that use SMB or NFS protocols. |
| | May not be ideal for VMs that create a large amount of change each day as merge times can be significant although this may still be acceptable if the merge finishes in an acceptable time. |

# Reverse Incremental

At its first run, reverse incremental backup creates a full backup file (VBK). All subsequent backups are incremental, that is, only changed data blocks are copied (using CBT if available). During the incremental backup, changed blocks are written directly into the full backup, while replaced blocks are taken out and copied into Veeam's rollback file (.VRB).

This method provides space-efficient backup, as there is only one full VBK to store. It also facilitates granular retention, since removing old points is simply a matter of deleting old VRB files.

The disadvantage is that creation of rollbackup files occurs during the backup process itself, which results in high I/O load on the target storage and can slow the backup process down. This could be a matter of concern especially for the VMs that experience high change rates.




## I/O Impact of Reverse Incremental

During the backup process as changed blocks are read from the source VM they are written directly to the VBK file. If this block replaces an ex-

isting, older block, that old block is read from the VBK and then written to the VRB file. This means that reverse incremental backups create a 66/33 read-write mix on the target storage during the backup process itself. This I/O typically becomes the limiting factor for backup performance of the job.

This can be especially noticeable for VMs with a high random change rate, or when running multiple simultaneous jobs, and is more noticeable on low-end storage or de-duplication appliances.

### Recommendations on Usage

| Use | **D**on't Use |
|---|---|
| Can be for used when repository storage uses fast disk with caching RAID controllers using large stripe sizes. | Small NAS boxes with limited spindles that depend on software RAID. |
| Excellent for low change rate VMs, especially large VMs with limited daily change. | Deduplication appliances that use SMB or NFS protocols. |
| | May not be ideal for VMs that create a large amount of change each day as merge times can be significant although this may still be acceptable if the merge finishes in an acceptable time. |
| | As the rollback is created during the backup process itself, backup throughput can be limited by target storage. This slower performance can lead to VM snapshots open for longer time. |

# Encryption

## Overview

The encryption technology in Veeam Backup & Replication allows you to protect data both while it is in transfer between backup infrastructure components and at rest, when it stored at its final destination. This can be disk, tape or a cloud repository. Customers can use one of encryption methods or a combination of both to protect against unauthorized access to important data through all steps in the data protection chain.

Veeam Backup Enterprise Manager additionally provides the password loss protection option that allows authorized Veeam users to recover data from the backup even if the password is lost. Note that you can recover data only if the backup server on which you decrypt data is connected to the same Veeam Backup Enterprise Manager server as the backup server on which the backup file was encrypted.

The encryption algorithms used are industry standard in all cases, leveraging AES-256 and public key encryption methods. The User Guide provides detailed information on the encryption algorithms and standards used by the product.

The following sections describe encryption options available in the product, what they protect, when they should be used and best practices for their use.

# Backup and Backup Copy Job Encryption

## What does it do?

Backup and backup copy job encryption is designed to protect data at rest. These settings protect data if an authorized user gets access to backup files outside of the backup infrastructure. Authorized users of the Veeam console do not need to know the password to restore data from encrypted backups. Encryption does not prevent authorized Veeam users from being able to access data stored in backups.

An example is the use of rotated drives for an offsite repository. Because these drives are rotated offsite, they are at a higher risk of falling into the hands of unauthorized users. Without encryption enabled, these unauthorized users could install their own copy of Veeam Backup & Replication and gain access to the stored backups easily.

On the other hand, if the backup files are encrypted, unauthorized users cannot access any data in the backups or even learn any critical information about the backup infrastructure as even backup metadata is encrypted. Without the key used for encryption or access to the original Veeam Backup & Replication console itself, the backup files remain secure.

## How does it work?

When Veeam Backup & Replication reads a block from disk, it compresses the block (unless compression is disabled at the job level), encrypts it with a session key generated for that job session, and stores the block into the backup file.

## When to use it?

Backup and backup copy job encryption should be used if backups are transported offsite, or if unauthorized users may easily gain access to backup files in another way than by using the Veeam console. Common scenarios are:

- Offsite backups to a repository using rotated drives
- Offsite backups using unencrypted tapes
- Offsite backups to a Cloud Connect provider
- Regulatory or policy based requirements to store backups encrypted

### Best Practices

- Enable encryption if you plan to store backups in locations outside of the security domain.
- While CPU usage for encryption is minimal for most modern processors, some amount of resources will still be consumed. If Veeam backup proxies are already highly loaded, take it into account prior to enabling job-level encryption.
- Use strong passwords for job encryption and develop a policy for changing them regularly.

**Note**

Veeam Backup & Replication helps with this, as it tracks passwords' age.

- Store passwords in a secure location.
- Obtain Enterprise or a higher-level license for Veeam Backup & Replication, configure Veeam Backup Enterprise Manager and connect backup servers to it to enable data loss prevention.
- Back up the Veeam Backup Enterprise Manager configuration database and create an image-level backup of the Veeam Backup Enterprise Manager server. If these backups are also encrypted, make sure that passwords are not lost as there will be no data loss prevention for these backups.

# Tape Job Encryption

## What does it do?

Similar to backup job encryption, tape job encryption is designed to protect data at rest. These settings protect data if an unauthorized user gains access to tape media outside of the backup infrastructure. Authorized users do not need to know the password to restore data from encrypted tape backups. Encryption does not prevent authorized Veeam users from being able to access data stored in tape backups.

Typical use case is to protect data on tapes when media is shipped to an offsite location or to a 3rd party vendor. Without encryption enabled, a lost tape could easily be accessed, and data stored on tapes could be compromised.

## How does it work?

Data is read from disk, and the session encryption key is used to encrypt data blocks as they are written to tape. Tape encryption can leverage either hardware tape encryption (if present and enabled) or software-based encryption. If the tape drive supports hardware encryption, the session key is sent to the tape device via SCSI commands and the drive itself performs the encryption prior to writing data to tape. This allows encryption to occur with no impact on performance or CPU of the tape server. If the tape hardware does not support encryption, Veeam falls back automatically to using software-based AES 256 data encryption prior to sending it to the tape device.

### When to use it?

Tape job encryption should be used any time you want to protect the data stored on tape from unauthorized access by a 3rd party. Tapes are commonly transported offsite and thus have a higher chance of being lost and turning up in unexpected places. Encrypting tapes can provide an added layer of protection if the physical tapes are lost.

If tape jobs are pushing already encrypted data to tape (for example, Veeam data from backup jobs that already have encryption enabled), you may find it acceptable to not use tape-level encryption. However, be aware that a user who gets access to the tape will be able to restore the backup files. Although this user will not be able to access the backup data in those files, some valuable information, for example, job names used for backup files, may leak.

## Best Practices

- Enable encryption if you plan to store tapes in locations outside of epy security domain.
- Consider the risks/rewards of enabling tape job encryption even if the source data is already encrypted and decide appropriately for level of risk.
- Use strong passwords for tape job encryption and develop a policy for changing them regularly (you can use Veeam Backup & Replication password age tracking capability).
- Store passwords in a secure location.
- Obtain Enterprise or a higher-level license for Veeam Backup & Replication, configure Veeam Backup Enterprise Manager and connect backup servers to it to enable data loss prevention.
- Back up the Veeam Backup Enterprise Manager configuration database and create an image-level backup of the Veeam Backup Enterprise Manager server. If these backups are also encrypted, make sure that passwords are not lost as there will be no data loss prevention for these backups.

# Network Transport Encryption

## What does it do?

Unlike the backup and tape job encryption features, the network transport encryption feature is designed to protect data "in-flight". For example, when the proxy is sending data across the network to the backup repository, the data can be encrypted between these two points even if job-level encryption is not enabled. This is primarily useful when the network between the source and target is not trusted, for example, when sending data across the Internet.

## How does it work?

Network encryption in Veeam Backup & Replication is controlled via the global Network Traffic options.

Whenever two backup infrastructure components need to communicate with each other over the IP network, a dynamic key is generated by the backup server and communicated to each node over a secure channel. The two components then establish an encrypted connection with each

other using this session key, and all communications between these two components for that session are then encrypted with this key. The key is of one-time use and is discarded once the session is complete.

### When to use it?

Network transport encryption should be used if the network between two backup infrastructure components is untrusted or if the user desires to protect Veeam traffic across the network from potential eavesdropping.

By default, Veeam Backup & Replication automatically encrypts communication between two nodes if either one or both has an interface configured (if used or not) that is not within the RFC1918 private address space (10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16, 169.254.0.0/16). Veeam also automatically uses network-level encryption for any connection to a Cloud Connect provider.

### Best Practices

- Enable encryption if a possibility of network-level eavesdropping is a security concern.
- Network-level encryption can use significant CPU resources, especially on the encrypting side (source) of the connection. Make sure that component nodes have enough resources.
- Use network-level encryption only where required. If backup infrastructure components are running on a network that is using non-RFC1918 IP addresses but is still private and secure from eavesdropping, consider using the **DisablePublicIPTrafficEncryption=1** registry key under **HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication** at the Backup & Replication (management) server to disable automatic network-layer encryption.

# Deduplication and Compression

## Storage Optimization Overview

Veeam Backup & Replication takes advantage of multiple techniques for optimizing the size of stored backups, primarily compression and deduplication. The main goal of these techniques is to strike the correct balance between the amount of data read and transferred during backup as well as what is stored on the backup target while providing acceptable

backup and restore performance. Veeam Backup & Replication attempts to use reasonable defaults based on various factors but there can be cases when leveraging settings other than default might be valuable.

# Deduplication

## What does it do?

The primary purpose of deduplication is to reduce the amount of data that has to be stored on disk by detecting redundant data within the backup and storing it only once. Veeam deduplication is based on identifying duplicate blocks across multiple VMs in a job. This is primarily beneficial when VMs are deployed from the same template since the base image is identical, but is less useful for incremental data.

## How does it work?

Deduplication is performed both by the source proxy (only for virtual disk currently being processed) and the target repository (for all virtual disks of all VMs in the job).

Veeam reads data blocks during the backup, calculates a unique hash for those blocks, and stores all identical blocks into the backup file for that session only once. By default, Veeam offers 4 different storage optimization settings that impact the size of read blocks and hash calculation for deduplication:

- **Local** – this is the default setting and is recommended when using a true disk-based repository. With this setting selected, Veeam reads data and calculates hashes in 1 MB chunks.
- **LAN** – this value is recommended when using a file-based repository such as SMB share. With this setting selected, Veeam reads data and calculates hashes in 512 KB chunks.
- **WAN** – this value is recommended when backing up directly over a slow link or for replication as it creates the smallest backups files at the cost of memory and backup performance. With this setting selected, Veeam reads data and calculates hashes in 256 KB chunks.
- **Local (>16MB)** – this setting is recommended for large backup jobs with more than 16 TB of source data in the job. With this setting selected, Veeam reads data hashes and calculates data on 8 MB blocks.

**Note**

See also Veeam forum at http://forums.veeam.com/veeam-backup-replication-f19/veeam-inline-dedupe-t10928.html?hilit=decipher#p47822 (login required to view this forum board).

## When to use it?

Veeam deduplication should be enabled in almost all cases. However, there are a few special cases where a user might consider disabling this option:

- **Deduplicating storage** – if you use deduplicating storage as a repository for storing Veeam backups, it might be desirable to disable Veeam deduplication. Veeam deduplication does not work at that same level as storage-based deduplication and will not actually interfere with the usefulness of such storage appliances. However, enabling Veeam deduplication requires storing deduplication metadata within the backup file and reading this data during backup and restore operations. This can be time-consuming and can lead to overall slower backups and restores. This behavior is worse with large backup jobs (>1 TB) that have long backup chains. Disabling deduplication in such cases may offer significant benefits.
- **Large compressed or deduplicated source VMs** – when backing up VMs, especially large VMs (>1 TB) that contain already compressed data (images, video, Windows deduplicated file servers, etc), it may be beneficial to simply disable Veeam deduplication since it is unlikely to provide much benefit for this type of source data. Note that Veeam deduplication is a job-level setting so VMs of the same type should be grouped and processed with one job.

## When do I change the defaults?

As a rule, the default settings provided by Veeam are designed to provide a good balance of backup size vs. backup and restore performance and resource usage during the backup process. However, given an abundance of resources or other specifics of the environment, it might be useful to change the defaults for a particular job.

For example, transactional servers like Microsoft Exchange and Microsoft SQL commonly make small changes across the disk. If you use the 1 MB blocks setting, this can lead to a great amount of incremental changes each day. Using the WAN optimization with a smaller block size of 256 KB may significantly decrease the size of increments. However, this can have a very significant impact on the speed and the amount of memory needed during the backup process on the repository, especially for large

backup jobs. This could also impact the source proxy at the CPU level.

A 2 TB Microsoft Exchange server may need only 2 GB of RAM on the repository during backup when using default settings of Local/1 MB blocks, but would potentially need 8 GB of RAM on the repository with WAN/256 K blocks. Also, transform operations such as synthetic fulls, forever forward retention/merge and reverse incremental rollback will perform 4x as much I/O, which can significantly increase total backup time. All of this must be taken into consideration prior to changing the defaults.

**Note**

Changing the block size of an existing job will not actually take effect until the next active full backup.

## Best Practices

- Unless you have a really good understanding of the impact that can cause block size changing, stick to the defaults.
- If you want to change the default block size, be sure to test it well and make sure you have planned appropriately for the extra I/O and memory requirements on the repository.
- Maximum recommended job size for WAN block size is 4 TB of source data.
- Maximum recommended job size for LAN block size is 8 TB of source data.
- Maximum recommended job size for Local block size is 16 TB of source data.
- Maximum recommended job size for Local (>16 TB) block size is 64TB of source data.
- When using a block size smaller than the default one for a large server, it is recommended to use a backup mode that does not perform synthetic processing (forward incremental with scheduled active full).

**Note**

Block size changes will only become effective after an active full is created.

# Compression

## What does it do?

Similar to deduplication, the primary purpose of compression is to reduce the amount of data that has to be transferred across the wire and stored on disk. Veeam Backup & Replication leverages several different compression algorithms that provide various balances between compression ratios, throughput and the amount of CPU use on the backup proxy. Compression provides maximum effect on space savings in a backup job, so understanding the tradeoffs in these settings can be very important.

## How does it work?

Veeam Backup & Replication performs compression on a per-block basis, using the block size selected by the storage optimization settings. The proxy reads each block from the source disk and applies the compression algorithm to the block before transferring it to the repository. This saves network bandwidth between the proxy and repository and allows the repository to store the already compressed block as soon as it receives it.

There are multiple compression options available:

- **None** – this option disables compression for the job. The proxy reads blocks and sends them uncompressed to the repository where they are written to disk as is.
- **Dedupe-friendly** – this option uses a very simple compression algorithm that needs very little CPU. It creates somewhat predictable data patterns, which is useful if users want to leverage 3rd party WAN accelerators with Veeam and/or a deduplication appliance (without the 'decompress before storing' setting). This allows the network stream to be moderately compressed while still being effectively cached.
- **Optimal** – this is the default compression used on Veeam jobs that leverages LZ4 compression. It provides typical compression ratios around 2:1 with fairly light CPU overhead. This light CPU overhead allows for excellent throughput with rates up to 150 MB/s per core and even faster decompression rates. This is a most commonly used practice that allows achieving excellent balance between performance and compression savings.

- **High** – this option uses lz3 compression tuned for low to moderate CPU overhead. This setting provides for around 10% higher compression ratios compared to optimal, but uses over 50% more CPU horsepower with rates up to 100 MB/core. If proxies are not otherwise CPU bound, this extra savings may still be very much worth it, especially for larger repositories or if the bandwidth available is less than the 100 MB/s limit (i.e., 1 Gb links or less).

- **Extreme** – this option uses lz3 compression tuned for high CPU overhead. This setting uses even more CPU and lowered through even further- to around 50 MB/core, with typically only around 3-5% additional savings. It is quite rarely used, however, in cases where bandwidth between the proxy and repository is limited, for example, when you backup directly through WAN links and are not able to backup on first side and use backup copy jobs for this.

Best practices is to use **Dedup-friendly** for deduplication storage, and **Optimal** for all other storage types.

## When to use it?

Veeam compression should almost always be enabled. However, when using a deduplicating storage system as a repository for storing Veeam backups, it might be desirable to disable Veeam compression at the job level.

If you configure deduplicating storage as a repository, it is recommended to use the **Decompress backup data blocks before storing** advanced option so that data should be written to the storage in the uncompressed format.

**Note**

The job-level compression settings do not change this, as they control the compression of data between the proxy and repository.

If there are any bandwidth constraints between the proxy and the repository, it may still be useful to enable compression at the job level while allowing the repository to decompress the data prior to writing it to the deduplicating appliance.

## When do I change the defaults?

As a rule, the default settings provided by Veeam are designed to provide a good balance of backup size vs. backup and restore performance and resource usage during the backup process. However, given an abun-

dance of resources or other specifics of the environment, it might be useful to change the defaults in particular circumstances. For example, if you know that CPU resources are plentiful, and backups are unable to make full use of the CPU due to other bottlenecks (disk/network), it might be worth increasing the compression level.

Compression settings can be changed on the job at any time and any new backup sessions will write new blocks with the new compression mode.Old blocks in already stored backups will remain in their existing compression level.

### Best Practices

- Defaults are good, don't change values without understanding the impact.
- Use compression levels above optimal only if you have plenty of CPU and understand that maximum throughput, especially during full backups, will likely be significantly lower, especially if the backup proxy CPUs can't take more load.
- Test various compression levels and see how the impacts the environment but remember the balance. A single backup job with a few concurrent streams may seem fine with **Extreme** compression, but may overload all available proxy CPUs during production run of all jobs.
- Remember that higher compression ratios may also negatively impact restore speeds.

# Backup Job

## Job Layout and Object Selection

Veeam Backup and Replication allows you to flexibly select objects to add to the job. At the **Virtual Machines** step of the job wizard, the **Add Objects** screen offers various "views" into the vCenter architecture that match the views provided by the vSphere client. You can switch between the **Hosts and Clusters**, **VMs and Templates**, **Datastores and VMs** or **Tags** views by pressing the appropriate button on the backup object selection screen.

This screen also provides an advanced object exclusion tool that allows you to select a parent object and then exclude child objects, or even individual disks within a VM.

More guidelines on object selection are listed below.

## Keeping Amount of Data Under Control

It is recommended to use the following settings for deduplication:

- **Local Target** as default for standard backup jobs
- **Local Target 16TB** in case full backup file is bigger than 8TB
- **LAN Target** as default setting for standard replication jobs

These values are based on the backup architecture.

Consider that there are good reasons to keep amount of data processed by a job even smaller: small backup files are easier to manage or move to new storage, they need less time and space when running a full backup, and they require smaller staging areas for restores from tape.

## Increasing Deduplication Rate

VMs created from the same or similar templates will increase deduplication rate, but keep in mind the backup windows and the size of the job for manageability.

## Adding Containers to Jobs

Adding resource pools, folders, datastores, or vSphere Tags (vSphere 5.5 and higher) to backup jobs makes backup management easier. New machines that become members of these groups are automatically included in the backup job.

With this approach, you need to monitor the amount of data in job to make sure there is enough datastore space.

If you add a number of datastores (or a mix of resource pools), make sure that you do not overlap, as VMs disks may reside on multiple datastores: a virtual machine residing on several datastores included in more than one backup job will be entirely backed-up in each job.

Limit the number of exclusions in backup jobs and exclude objects carefully. While exclusions can be very useful, virtual infrastructure is dynamic and changes over time. It's quite possible that a VM gets moved to a folder or resource pool that is excluded which makes it unprotected.

A word about tags:

- Tags might be very convenient for VM selection; however, you need to monitor the number of VM that will be automatically added to the job.
- Using tags, you can classify VMs by service levels, using different backup jobs for different service levels.
- Veeam ONE Business View is a very convenient tool to handle tags, allowing you to create classification rules and write tags on the vCenter - for example, according to CPU number, RAM quantity, names of VMs, folders, resource pools, datastores, and so. Veeam ONE Business View can also import VM/host/datastore descriptions from a CSV file. This feature can be useful when refreshing VMware tags, for example, to update the change management database information.

## Setting Deduplication and Compression Level

Detailed description of these settings and their influence on the backup infrastructure is provided in the "Deduplication and Compression "section above in this document. In almost all cases deduplication should be left enabled. Veeam Backup & Replication uses source side deduplication which decreases the amount of data that must be transferred to the target repository.

**Tip**

If you are using a deduplication storage appliance, you have to disable inline deduplication to get better performance on incremental runs, as Veeams own inline deduplication increases random reads from the target significant and therefore the jobs process longer as needed. It as well optimizes (disable) Metadata Caching accordingly to again increase performance with deduplication appliances without a special Veeam integration (like DDBoost protocol).

## Setting Target Optimization

When choosing target mode, follow the guidelines described in the "Deduplication" and "Compression" sections above.

For very high change rate VMs such as Exchange and SQL servers, choosing WAN target mode (256K block size) may be more efficient even for

local backups. However, WAN target mode takes high CPU load. Tests showed a 50% or more reduction of incremental backup data amount when using WAN target mode for incremental backups of highly transactional servers.

For compression, the **Optimal** level is almost always the best. However, if the target is a hardware deduplication appliance, you will generally achieve the best compression and deduplication ratios by writing the data to the storage uncompressed, or by using **Dedupe-friendly** compression level. Since the target data mover (that is, Veeam Transport service) is available to decompress data prior to writing it to the target, this option is the recommended one for such cases.

You can also use 4K blocks data alignment, which can be helpful for deduplication storages with fixed block size. If necessary, use the **Align backup file data blocks** setting to change the block size.

## Encryption

Detailed description of encryption settings and their influence on the backup infrastructure is provided in the " Encryption "section above in this document.

Job encryption might cause some proxy resources consumption, since Veeam Backup & Replication encrypts data on the proxy and then sends it to the repository. You might need to add a bit more computing resources on the proxy to compensate the additional load.

**Tip**

By default, Veeam Backup & Replication will encrypt data if one of the Proxy/Repository interfaces (if used or not) holds a public IP address. Please refer to the "Encryption" section of this guide for more details.

For general guidelines for encryption configuration, refer to the Veeam User Guide: http://helpcenter.veeam.com/backup/80/vsphere/encryption_keys.html?zoom_highlightsub=encryption.

## Backup Jobs Chaining

Chaining backup jobs is convenient in certain circumstances, but should be used with caution. For example, if a job in such chain fails or stops responding, the whole job chain delivers poor backup success rate.

A most common way to handle multiple jobs is to let Veeam Backup &Replication Scheduler handle the proxy/repository ressources by starting more jobs (overbooking) than available proxy/repository resources. This will allow Veeam Backup & Replication to optimize the backup time window with its load balancing algorithm.

## Load Balancing

When planning jobs schedule, you should think of balancing the load on source and target disks. Too many jobs accessing the same disk will load the storage significantly; this makes the job run slower or may have a negative impact on the VMs performance. To mitigate this problem, you can utilize the Backup I/O Control settings (see the " Backup I/O Control" section above).

Veeam also employs a load balancing method that automatically allocates a proxy, making a choice between all proxies managed by Veeam Backup & Replication that are available at the moment.

For more details on load balancing, refer to the Veeam Backup & Replication User Guide at http://helpcenter.veeam.com/backup/80/vsphere/index.html?resource_scheduling.html.

## Binding Jobs to Specific Proxies

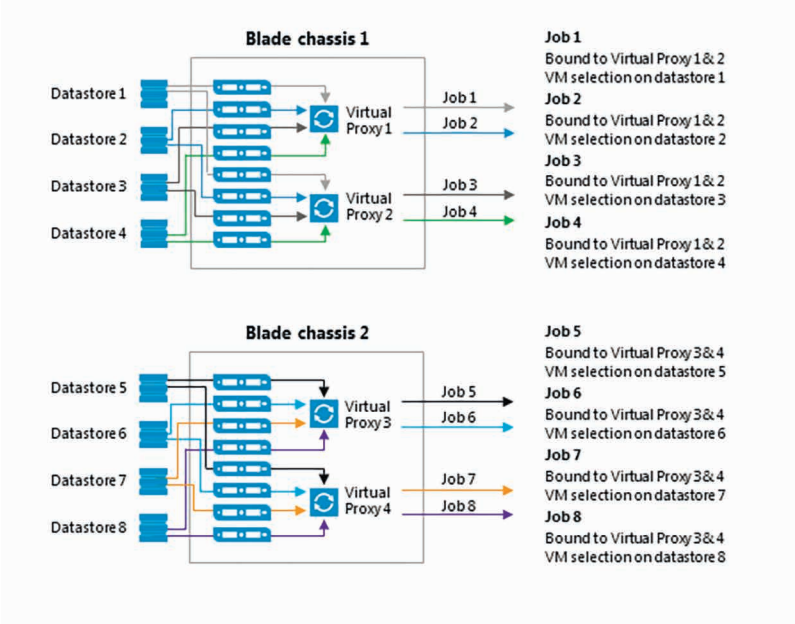
Refer to the User Guide in order to examine the advanced deployment scenario with multiple proxies: http://helpcenter.veeam.com/backup/80/vsphere/advanced.html

While configuring a backup job, you can disable the automatic proxy selection. Instead, you can select particular proxies from the list of proxies managed by Veeam backup server, and appoint them to the job. This is a very good way to manage distributed infrastructures; also it helps you to keep performance under control.

For example, you can back up a cluster residing on multiple blade chassis. In this case, if you use virtual proxies, keep the proxies load well-balanced and optimize the network traffic.

Dedicated proxies can be also very helpful if you use a stretched cluster and do not want proxy traffic to go across inter-switch link.

See the illustration below as a good starting point to reach and keep control on high backup throughput. In this example, administrator wants to keep network traffic as much as possible inside the chassis; only the proxy-to-repository traffic goes via an external link.




**Tip**

To optimize load balancing in a distributed environment where backup proxies are rolled out to multiple sites, it is recommended to select all proxies from the same site in the job.

# Backup Copy Job

Instead of just copying backup files to a second destination, Veeam uses a more intelligent and more secure way of bringing restore points to a second backup target. Backup copy jobs read specific VM restore points from backup files and store them as a new backup file chain on the destionation. The second chain is independent from the first chain and adds therefor an additional level of protection. You can store VMs from multiple jobs at the backup copy job target. As an option, select one or several VMs from a bigger backup job as source (if you do not want to backup all VMs to the backup copy job desitionation).

Every backup copy job creates its own folder on the target backup repository and stores all copied restore points to that location. The folder has the same name as the backup copy job.

**Tip**

Though a backup copy job cannot be targeted at the same repository where the source backup job places its backup files, it is possible to process backup copy jobs between 2 repositories on the same repository server.

Once created, a backup copy job will immediately start processing the latest existing restore point for all VMs included in the job (as long as it has been created less than one synchronization interval before the start of the backup copy job).

By default, Veeam Backup & Replication keeps 7 restore points on the target backup repository in case of simple retention policy (see the "Simple Retention Policy" section of the User Guide for details). If you plan to use Grandfather-Father-Son (GFS) retention, refer to the "GFS Retention Policy" section for details.

Even if a backup copy job processes several VMs, it creates one backup file on the target backup repository and stores to it data for all VMs processed by the job.

If a backup copy job cannot process all requested VMs during its synchronization interval, the job will still create a backup file on the target backup repository, but some VMs will be left inconsistent or unprotected. This might be caused by precedence of the backup task over the backup copy task.

Limitations of backup copy jobs are described in Veeam Backup & Replication User Guide at http://helpcenter.veeam.com/backup/80/vsphere/index.html?backup_copy_select_point.html.

## Backup Copy Job Scheduling

By design, a backup copy job is a continuous process that runs permanently. This process includes several stages.

A backup job restarts itself at the defined **Copy every: …** interval setting (default is 12:00 AM every day). and looking for new restore points of the selected VMs. On the **Scheduler** tab, it is possible to separately schedule the data transfer itself. Veeam global traffic rules can be used to limit bandwidth over all Veeam jobs at a specific connection (IP addresses or ranges).

A single backup copy job can also monitor multiple backup jobs so the concept of the "interval" is to define two things: how often it should be looking for new points, and for a daily interval, what time it should start looking for points. If you set an interval for 1 day, that's telling the backup copy job that once a day, starting at the selected time, it should begin looking for new restore points and, as soon as it finds one, copy it. However, once a single point is copied, another point for that VM will not be copied until the next interval starts.

The idea behind this is that you may run backups locally more often (for example, hourly), but you may only want to copy offsite once a day or weekly, thus you can set the backup copy "interval" independently of the schedule of the backup jobs it is monitoring. Think of it almost like about setting an SLA.

The backup copy job has the following phases:

1. **Data transfer (synchronization) phase** — during this phase, the backup copy job checks for a new restore point on source, creates a file for a new restore point on target and starts copying the most recent restore point of each processed VM to the target repository. The data transfer (synchronization) phase starts at specific time configured in the job properties (see http://helpcenter.veeam.com/backup/80/vsphere/index.html?backup_copy_sync_interval.html). You can define any interval needed in minutes, hours or days. Moreover, you can specify the time slot during which data can and cannot be transferred over the network, thus regulating network usage (see http://helpcenter.veeam.com/backup/80/vsphere/index.html?backup_copy_window.html).

2. **Transform phase** — you can configure Veeam Backup & Replication to perform a number of additional transform operations on the target backup repository: transforming a backup chain, removing deleted VMs from restore points, compacting a full backup file. The transform phase begins after all restore points are copied to target.

**Note**

Consider that the transform process itself puts additional pressure on the target repository. In larger environments with deduplication storage appliances used as backup repositories or with backup copy jobs processing a huge number of VMs or VMs of big size, the transform process can take significant amount of time. In the backup copy job chain, for each block in the first incremental point there will be 1x read and 1x write I/O. For example, for 8 disk NAS in RAID5, 100 GB of daily change in the worst case may take around three hours.

In the properties of the backup copy job, you can select to perform post-job activities, such as execution of custom scripts or sending job results by email. Post-job activities are performed after all transform operations are completed.

3. **Idle phase** — for the most time, the backup copy job remains in the *Idle* state, waiting for a new restore point to appear on the source backup repository. When a new synchronization interval starts, as specified by the "Copy every" setting of the job, steps 1-2 are repeated. For more information, refer to the Backup copy job section of the Veeam Backup & Replication User Guide at http://helpcenter.veeam.com/backup/80/vsphere/index.html?backup_copy_job_task.html.

**Note**

The synchronization interval defines how frequently the backup copy job must copy the VM restore point from the source backup repository to the target backup repository. For instance, if you select to copy a restore point every 7 days, there will be one restore point on target per week created by the backup copy job.

# Job Layout and Object Selection

## Source Object Container

- **Select from infrastructure**: this method of selection allows you to select a specific VM or container from the infrastructure perspective. In this case, the scheduler will look backward its starting time for the most recent restore point containing the VM(s) within the synchronization interval. The scheduler will look for restore points in all backups, no matter which job has generated the restore point. If the restore point is locked (the backup job is not ended yet), the backup copy job will wait for the restore point to be unlocked (the end of the backup job) and then start copying the VM(s) restore point according to its defined schedule.
- **Select from job**: this method of selection is very useful if you have multiple backups protecting the same VMs. In this case, you can bind the backup copy job to a specific restore point you want to copy. This kind of container selection will dynamically protect all VMs in the specified job(s), and only in this job.
- **Select from backup**: this method is equivalent to the **Select from job** method but allows for selecting specific VMs inside any job.

## Backup Copy and Tags

As you can select any VM to be copied from multiple backups, you can consider quality of service-based configuration.

For instance, you may not want to apply GFS retention over some VMs like web servers, DHCP, etc. In this situation, you can use VMware tags to simplify the management of backup copy process. Tags can be easily defined according to the desired backup copy configuration, using VMware vSphere or Veeam ONE Business View.

## Backup Method

Since backup copy process reads VM data from backup files, the RPO and the backup job method will be tied.

For instance, if you create a backup copy job based on a reverse incremental backup job, every day when the backup job initiates its transform phase, the backup copy job will be interrupted (if still running), leaving inconsistent copies on the target backup repository.

**Tip**

If the first backup copy (full) cycle is expected to be very long and if you cannot perform a seeding task (to learn more about seeding, see: http://helpcenter.veeam.com/backup/80/vsphere/backup_copy_mapping_auxiliary.html), consider creating a forward incremental job, which will provide a backup copy job with the requested time to complete. Also, make sure that the backup copy job synchronization interval is long enough. Refer to this section of Veeam Backup User Guide for details about backup copy issues : http://helpcenter.veeam.com/backup/80/vsphere/backup_copy_issues.html

## Common Use: Backup Copy for a Whole Job

Knowing all this, the most common way to configure a backup copy job is to use a daily backup job as a source container and to schedule the backup copy job to start a few minutes after the backup job it is tied to.

Since the backup copy job requires some connections to the vCenter Server, starting it a bit earlier than the backup jobs will lower the number of simultaneous connections against the vCenter Server. As a result, the backup copy will start copying data as soon as the backup job ends and the source backup file on the backup repository gets unlocked.

**Tip**

This is a common RO/BO use case of backup copy job.

# Additional Options

## Restore Point Lookup

By default, after a restart of the backup copy job interval (the **Copy after:** setting) it will analyse the VMs and the last transferred restore points of it. If there was no restore point transferred in the previous run of a VM, the backup copy job starts transferring the last restore point immediately. If you do not want the backup copy job to look backward at the starting time, use the **LookForward** registry key to change the "looking direction" to only look for newer restore points. This option is available since Veeam Backup & Replication 8.0.

The following forum thread provides a very good explanation of the backup copy scheduler and the LookForward registration key: http://forums.veeam.com/veeam-backup-replication-f2/backup-copy-intervals-t24238.html

According to the What's New documentation for Veeam Backup & Replication v8, the registry key **"BackupCopyLookForward (DWORD) = 1"** should be created **in "HKLM\SOFTWARE\Veeam\Veeam Backup and Replication"**.

## Backup Copy from Backup Copy

Since v8, it is possible to use a backup copy job as a source for data transfer and to generate a third copy of VMs. For this, select the VMs from infrastructure and specify the backup repository holding the primary backup copy restore points as the source.

## Job Seeding

Usually, a backup copy is sending data remotely. If it is necessary to send data offsite over a slow WAN link, seed the backup copy job by taking the following steps:

1. Create a "local" backup copy job and target it at a removable device (for example, USB HDD) used as a backup repository. Run the created backup copy job to create a full backup file on this removable device.

2. Once the backup copy job is over, delete the local backup copy job from the Veeam console.

3. Transport the removable device with created backup files to the destination site.

4. Copy backup files to the target backup repository.

5. Create a definitive backup copy job on the Veeam console. On the **Target** step of the **Backup copy job** wizard, click the **Map backup** link and select the copied backup — this backup will be used as a "seed".

If you are using a WAN accelerated transfer, refer to the WAN Accelerator section for proper cache population procedure: http://helpcenter.veeam.com/backup/80/vsphere/index.html?wan_populate_cache.html.

**Tip**

Sometimes VeeamZIP is used to create the actual backup for seeding. VeeamZIP does not create the needed VBM metadata file. Instead of creating a new backup and shipping it to the seeding target side again, contact Veeam support to get a PowerShell script that can create the VBM metadata file.

**Note**

Only the initial first run of a reverse incremental chain can be used with seeding (but any forward incremental chain can be used).

### Target Repository for Forever-Incremental Chain

The backup copy process is forever incremental. This means that only the first transfer of VM data will be a full copy, and every subsequent transfer will be incremental.

This will generate read I/O on the target repository at the end of the job — Veeam Backup & Replication needs to transform the full backup and optionally create GFS full copies of the VMs for archival purposes. Knowing this, you may find a good balance between cost and performance for the backup copy target repository storage. To get more information about the backup repository planning, see the "Repository Server" section above.

**Tip**

Once you are aware of the sequential nature of the backup copy job, try to enhance performance if the target-to-source link is not the bottleneck. For example, you can schedule more backup copy jobs at the same time to get a better parallelization effect.

## Replication Job

Replication jobs are used to replicate specified VMs to another or the same virtual environment (instead of creating deduplicated and compressed backup files at backup run). Veeam can store 28 restore points (VMware).

Like backup, replication is a job-driven process. In many ways, it works similarly to forward incremental backup:

- During the first run of a replication job, Veeam Backup & Replication copies a whole VM image and registers the replicated VM on the target ESXi host.
- During subsequent runs, the replication job copies only incremental changes, and creates restore points for the VM replica — so the VM can be recovered to the necessary state. Every restore point is in fact a usual VMware snapshot.
- When you perform incremental replication, data blocks that have changed since the last replication cycle are written to the snapshot delta file next to the full VM replica. The number of restore points in the chain depends on the retention policy settings.

Replication infrastructure and process are very similar to those used for backup. They include a source host, a target host with associated datastores, one or two proxy servers and a repository. The source host and the target host are the two terminal points between which the replicated data is moved.

Replicated data is collected, transformed and transferred with the help of Veeam data movers. The data movers involved in replication work with the source proxy, the target proxy and the repository. The data mover hosted on the repository processes replica metadata files.

**Important!**

Although the replica data is written to the target datastore, certain replica metadata must be located on a backup repository. This metadata is used by the source proxy and thus should be deployed closer to the source host and therefore no compression/uncompression processing is used.

The replication process involves the following steps:

1. When a new replication session is started, the source-side data mover (proxy task) performs the same operations as in backup process. In addition, in cases when VMware CBT mechanism cannot be used, the source-side data mover interacts with the repository data mover to obtain replica metadata — in order to detect which blocks have changed since the previous job run.
2. The source-side data mover compresses the copied blocks of data and transfers them to the target data mover.

**Note**

In on-site replication scenarios, the source-side transport service and the target-side transport service may run on the same backup proxy.

3. The target-side data mover uncompresses replica data and writes it to the destination datastore.

Veeam Backup & Replication supports a number of replication scenarios that depend on the location of the target host and will be discussed later in this section.

During replication cycles, Veeam Backup & Replication creates the following files for a VM replica:

- A full VM replica (a set of VM configuration files and virtual disks).

During the first replication cycle, Veeam Backup & Replication copies these files to the selected datastore to the *<ReplicaName>* folder, and registers a VM replica on the target host.

- Replica restore points (snapshot delta files). During incremental runs, the replication job creates a snapshot delta file in the same folder, next to a full VM replica.

- Replica metadata where replica checksums are stored. Veeam Backup & Replication uses this file to quickly detect changed blocks of data between two replica states. Metadata files are stored on the backup repository.

During the first run of a replication job, Veeam Backup & Replication creates a replica with empty virtual disks on the target datastore. Disks are then populated with data copied from the source side.

To streamline the replication process, you can deploy the backup proxy on a virtual machine. The virtual backup proxy must be registered on an ESXi host with direct connection to the target datastore. In this case, the backup proxy will be able to use the Virtual Appliance transport mode for writing replica data to target.

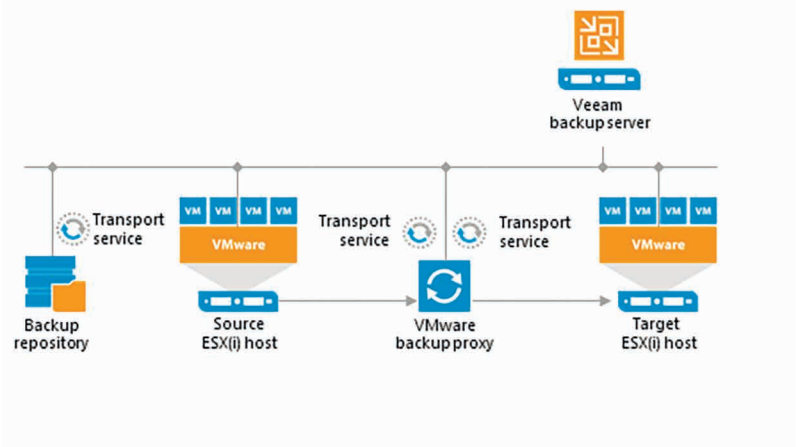
If the Virtual Appliance mode is applicable, replica virtual disks are mounted to the backup proxy and populated through the ESX I/O stack. This results in increased writing speed and fail-safe replication to ESXi targets. For information on Virtual Appliance mode, see http://helpcenter.veeam.com/backup/80/vsphere/index.html?virtual_appliance.html.

If the backup proxy is deployed on a physical server, or the Virtual Appliance mode cannot be used for other reasons, Veeam Backup & Replication will use the Network transport mode to populate replica disk files. For information on the Network mode, see http://helpcenter.veeam.com/backup/80/vsphere/index.html?network_mode.html.

The Direct SAN mode can only be used together with replication targets in case of transferring thick-provisioned VM disks at the first replication run. As replication restore points are based on VMware snapshots, that are thin provisioned by definition, Veeam will failback to Virtual Appliance (HotAdd) mode or Network mode, if configured at proxy transport settings. Direct SAN mode or backup from storage snapshots can be used on the source side in any scenario. See this topic on Veeam forum: http://forums.veeam.com/veeam-backup-replication-f19/direct-san-replication-t23748.html#p122526 (please log in to view this forum board).

## Onsite Replication

If the source and target hosts are located in the same site, you can use one backup proxy for data processing and a backup repository for storing replica metadata. The backup proxy must have access to both source host and target host. In this scenario, the source-side data mover and the target-side data mover will be started on the same backup proxy. Replication data will be transferred between these two data movers and will not be compressed.



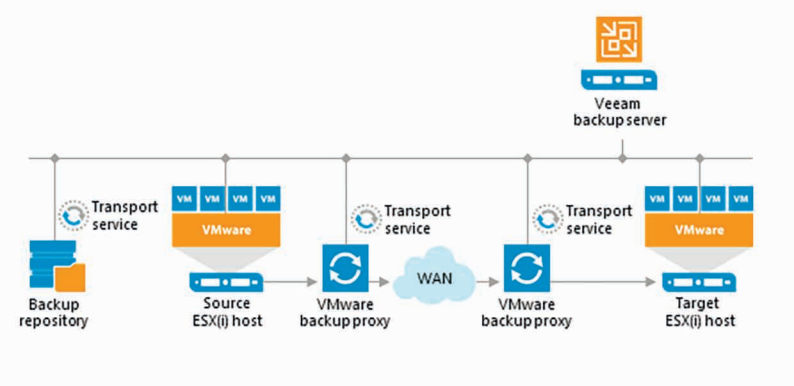

## Offsite Replication

The common requirement for offsite replication is that one Veeam data mover runs in the production site (closer to the source host), and another data mover runs in a remote DR site (closer to the target host). During backup, the data movers maintain a stable connection, which allows for uninterrupted operation over WAN or slow links.

Thus, to replicate across remote sites, deploy at least one local backup proxy in each site:

1. A source backup proxy in the production site.
2. A target backup proxy in the remote DR site.

The backup repository must be deployed in the production site, closer to the source backup proxy.




**Tip**

It is recommended to place a Veeam backup server on the replica target side so that it can perform a failover when the source side is down. When planning off-site replication, consider advanced possibilities — replica seeding, replica mapping and WAN acceleration. These mechanisms reduce the amount of replication traffic while network mapping and re-IP streamline replica configuration.

For offsite replication, open the connections between the Veeam backup components:

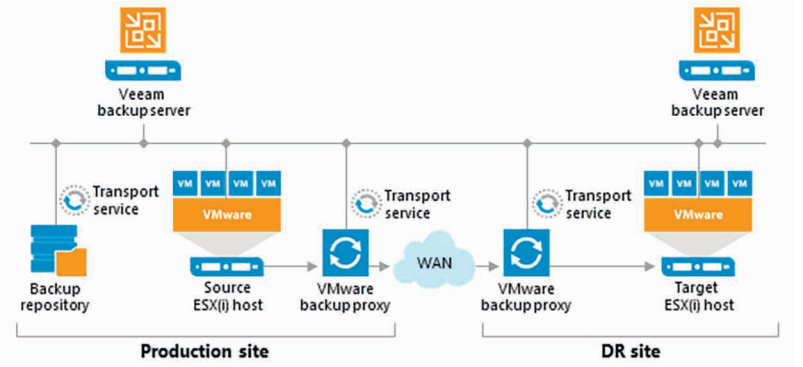
- The Veeam backup server must have access to the vCenter Server, the ESXi hosts, the source backup proxy and the target backup proxy.
- The source backup proxy must have access to the Veeam backup server, the source ESXi host, backup repository holding the replica metadata, the target proxy, and the source vCenter Server.
- The target backup proxy must have access to the Veeam backup server, the source proxy, the target ESXi host, and the target vCenter Server.

The source proxy compresses data and sends it via the WAN to the target proxy, where the data is uncompressed. Note that you also can seed the replica by sending the backup files offsite (using some external media, for example) and then only synchronize it with incremental job runs.

It is also recommended to install an additional Veeam backup server in the DR site. No additional license is required for this, since Veeam is licensed by the physical CPU socket of the source hypervisor hosts (where the protected virtual machines reside), not by the Veeam backup server.

In this scenario:

- The Veeam backup server in the production site will be responsible for backup jobs (and/or local replication).
- The Veeam backup server in the DR site will control replication from the production site to the DR site.




Thus, in disaster situation, all recovery operations (failover, failback and other) will be performed by the Veeam backup server in the DR site. Additionally, it may be worth installing the Veeam Backup Enterprise Manager to have visibility across the two Veeam backup servers so that you only have to license the source virtual environment once (used from both backup servers)

**Tip**

Plan for possible failover carefully. DNS and possibly authentication services (Active Directory, for example) should be implemented redundant across both sides. vCenter Server (and vCD) infrastructure should be as well considered for the failover scenario. In most cases, Veeam do not need a vCenter Server for replica target processing. It can be best practice to add the ESXi hosts directly to Veeam Backup & Replication as managed servers and to perform replication without vCenter Server on the target side.

Replication bandwidth estimation has always been a challenge, because it depends on multiple factors such as the number and size of VMs, change rate (at least daily, per RPO cycle is ideal), RPO target, replication window. Full information about these factors, however, is rarely at hand. You may try to set up a backup job having the same settings as the rep-

lication job, and test the bandwidth (as the backup job will transfer the same amount of data as the replication job).

Also, when replicating VMs to a remote DR site, you can manage network traffic by applying traffic throttling rules or limiting the number of data transfer connections. See Veeam Backup & Replication User Guide for more information: http://helpcenter.veeam.com/backup/80/vsphere/index.html?traffic_throttling.html.

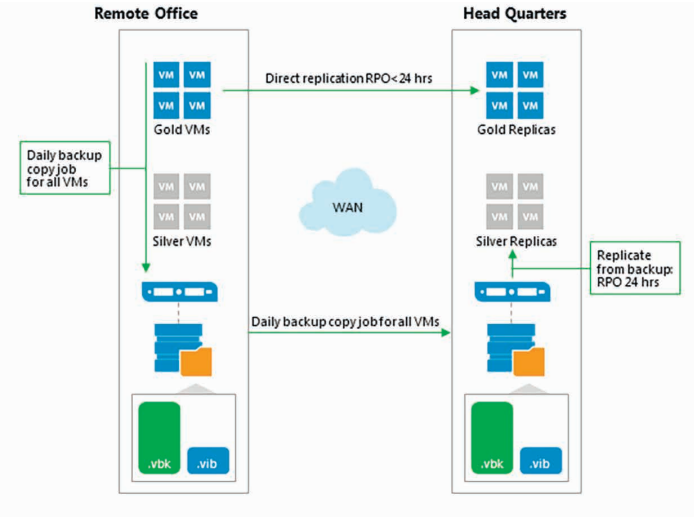
**Tip**

Starting from version 8, replication can leverage WAN acceleration allowing a more effective use of the link between the source and remote sites. For more information, see the User Guide http://helpcenter.veeam.com/backup/80/vsphere/index.html?wan_acceleration.html or the present document (the "WAN Acceleration" section above).

# Replica from Backup

In some circumstances, you can get a better RTO with an RPO greater or equal to 24 hours, using replicas from backup. A common example is a remote office infrastructure, where the link between the remote site and the headquarters provides limited capacity.

In this case, the data communication link should be mostly used for the critical VM replicas synchronization with good RPO. Now, assuming that a backup copy job runs for all VMs every night, some non-critical VMs can be replicated from the daily backup file. This requires only one VM snapshot and only one data transfer.



You can find additional information about replica from backup in the appropriate section of the Veeam Backup & Replication User Guide: http://helpcenter.veeam.com/backup/80/vsphere/replica_from_backup.html

**Tip**

This feature is sometimes named and used as proactive restore. Together with SureReplica, it is a powerful feature for availability.

## Backup from Replica

It may appear an effective solution to create a VM backup from its offsite replica (for example, as a way to offload a production infrastructure); however this scheme is not at all valid because of VMware limitations concerning CBT (you cannot use CBT if the VM was never started). There is a very well documented forum thread about this subject: http://forums.veeam.com/vmware-vsphere-f24/backup-the-replicated-vms-t3703-90.html.

## Replication from Storage Snapshots

Reading replicas from storage snapshots is supported; however it depends on the storage ability to cope with a large number of simultaneous snapshot requests.

Too many snapshots affect overall performance of the production storage. It may happen that a SAN array cannot provide the number of snapshots requested by the replication jobs.

So when configuring the infrastructure, try to group replicated VMs on a minimum number of datastores. In this case, the storage array will take one snapshot containing many replicated VMs.

# Application-Aware Image Processing

When configuring Veeam backup and replication jobs, you can specify how the transactionally-consistent backup images of VMware VMs should be created. Two methods are available for bringing VM file system and applications into consistent state: VMware Tools quiescence and Veeam Application-Aware Image Processing (utilizing Windows VSS). Key features of both methods are illustrated by the following table:

| Feature | VMware Tools Quiescence | Veeam Application-Aware Image Processing |
|---|---|---|
| Support for consistent backup on Windows guest | Yes | Yes |
| **Sync** driver for Linux guest | Yes | No |
| Support for application-aware backup | Limited | Yes |
| Pre-VSS preparation for specific applications (e.g. Oracle) | No | Yes |
| Support for application log truncation (Microsoft SQL Server and Exchange Server) | No | Yes |
| Support for scripts | Yes (need to be placed on VM guest) | Yes (can be centrally distributed) |
| Interaction with user via UI | Not needed | Not needed |
| Error reporting | Within VM guest OS | Centralized, on Veeam backup server |
| Stability | Good | Very good |

## How Veeam Guest OS Processing Works

1. First, Veeam Backup & Replication performs guest OS inventory to find out if there is a VSS-aware application running inside a VM.
2. Veeam Backup & Replication runs pre-freeze script (if any) for the Microsoft Windows/Linux guest OS with applications that utilize other means of VM quiescence.
3. Then VSS quiescence of the VM is performed, including restore awareness settings.
4. VM snapshot is created.
5. VSS unfreeze ("thaw") is performed.
6. Veeam Backup & Replication runs post-thaw script (if any) for the Microsoft Windows/Linux guest OS.
7. Backup data transfer and snapshot commit is performed.
8. Finally, log file truncation is performed with VSS (for Microsoft SQL Server and Exchange Server).

# Selecting Guest Processing Options

When on the **Guest Processing** step of the job wizard, you are presented with the variety of options (as described in detail in the User Guide (http://helpcenter.veeam.com/backup/80/vsphere/index.html?backup_job_vss_vm.html).

Note that you can use pre- and post-job scripting to automate job global settings from the Veeam Backup & Replication server itself. It is recommended to use the VM guest processing options for interaction with VMs.

To select the necessary options, refer to the table below.

| VM guest OS type | Linux (with applications and known user for Guest OS processing) | Windows and VMware VSS-supported applications (without known user for Guest OS processing) | Windows with VSS-aware applications | Windows (no VSS-aware applications) | Linux with applications | Linux server (no applications) |
|---|---|---|---|---|---|---|
| Guest OS processing is applicable | Yes | Yes | Yes | Yes | Yes | No |
| Use VMware Tools quiescence | No | Yes | No | No | No | No |
| VMware Tools quiescence with VMware Script processing | Yes | No | No | No | No | No |
| Use Veeam Application-Aware Processing | No | No | Yes | No | No | No |
| Use Veeam Application-Aware Processing and InGuest Scripts | No | No | No | Yes | No | No |
| Use Veeam InGuest Script Processing (Application Aware Processing disabled) | No | No | No | No | Yes | No |

To coordinate proper VSS and indexing activities, Veeam Backup & Replication deploys a small executable component inside a VM. It is installed only during VSS quiescence procedure and removed immediately after the processing is finished, producing very low impact on VM performance and stability. As for connection method for accessing VM guest OS, Veeam first tries to connect to the VM over network using RPC and then by VMware VIX channel through VMware Tools (for Windows guest only).

**Tip**

If the backup server has no network connection to the VMs (for example, service providers), this order can be changed using the **InverseVssProtocolOrder (REG_DWORD)** registry value under **HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication**:

Value = **1** – try connection through VIX first
Value = **0** – try network connection through \\admin$ first
See Veeam Knowledge Base article at http://www.veeam.com/kb1230

Consider that this is a global setting that will be applied on the Veeam backup server level (to all jobs with application-aware processing configured on it).

Depending on the VM guest OS processing options selected (enabled or disabled application-aware image processing) and on the guest access method, you may need to supply access credentials for the guest OS, as described in the tables below.

## Guest Access Credentials for Windows OS

| Application-Aware Image Processing (AAIP) | Using VMware Tools Quiescence | Using Veeam (guest access by VIX) | Using Veeam (guest access by VIX) | Using Veeam (guest access by Network) | Disabled (crash-consistent backup is created) |
|---|---|---|---|---|---|
| Membership in the local Administrators group | User account not needed | Yes | No | Yes | Not needed |
| Enter username as *<servername>\ Administrator* or *<domain>\Administrator* | No | No | Yes | No | N/A |
| UAC can be enabled | Yes | No | Yes | Yes | Yes |
| VMware Tools must be installed and up to date | Yes | Yes | Yes | Yes | No |

## Guest Access Credentials for Linux OS

| Linux guest OS processing | Using VMware Tools Quiescence | Using Veeam (guest access by Network) | Disabled (crash-consistent backup is created) |
|---|---|---|---|
| Root user account | User account not needed | Yes | Not needed |
| User requires sudoer rights | No | Yes | No |
| Certificate-based authentication is possible | No | Yes | No |
| VMware Tools must be installed and up to date | Yes | Yes | No |

**Tip**

To verify the credentials you supplied on the Guest Processing step of the job wizard, click **Test Now** button.

# Required Ports

The following ports should be open between the Veeam backup server and VM for guest OS processing:

- For Windows VMs - remote RPC ports, including Dynamic Port Range (TCP ports 1025 to 5000 - for Microsoft Windows 2003, 49152-65535 - for Microsoft Windows 2008 and newer); TCP\UDP ports 135, 137-139, 445.
- For Linux VMs – SSH port (default is TCP port 22)

For details, refer to the Veeam Backup & Replication User Guide (http://helpcenter.veeam.com/backup/80/vsphere/index.html?used_ports.html).

# Sizing

Since guest processing produces very low impact on VM performance, no special considerations on sizing are required. If you use VSS processing with VMware Tools quiescence or Veeam in-guest processing, you need free space on each drive of the VM for the software VSS snapshot. Please check Microsoft requirements for more information.

**Note**

If you plan to use snapshot-only backups with NetApp storage system (see http://helpcenter.veeam.com/backup/80/vsphere/index.html?netapp_integration.html for details), consider that default number of VM guest systems processed in parallel is by default set to **20**.

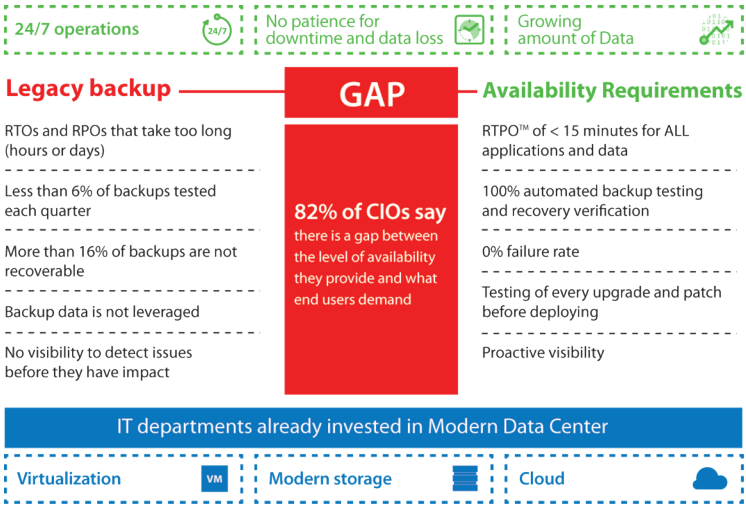
## Bridging the Availability Gap

Organizations are modernizing their data centers in order to provision IT services faster, strengthen security and control, and lower operational costs. While building modern data centers, organizations invest in server virtualization, modern storage applications and cloud-based services. However, businesses are facing new demands from end users including access to data and applications 24/7, no patience for downtime or data loss, and exponential data growth at 30-50% per year.

This opens a gap—an availability gap—between the requirements of the Always-On Business™ and IT's ability to effectively deliver availability. In fact, 82% of CIOs say there is a gap between the level of availability they provide and what end users demand.

Veeam bridges this gap by providing customers a new kind of solution—Availability for the Modern Data Center, which delivers RTPO of < 15 minutes for all applications and data.

Organizations now can leverage their investments in the modern data center to meet new demands of the always-on business.



This section of the document will demonstrate how Veeam solution can be used throughout an entire datacenter availability project, beginning with the first assessment phase to the project implementation from the technical perspective.

**Note**

While these guidelines focus on enterprise customers with more than 100 hosts or thousand virtual machines, Veeam Availability Suite solutions are applicable to any infrastructure size.

# Assessment

Before starting any project, it is very important to understand customers' needs, vision and the IT environment. While the first two are likely to be the outcome of an initial project meeting, the IT environment can be analyzed with Veeam ONE, which is a part of the Veeam Availability Suite.

The following information is very important and can help to streamline the project and proactively prevent situations that impact the environment:

- Veeam ONE Monitor: *Storage Latency* report and data in the *Alert* tab

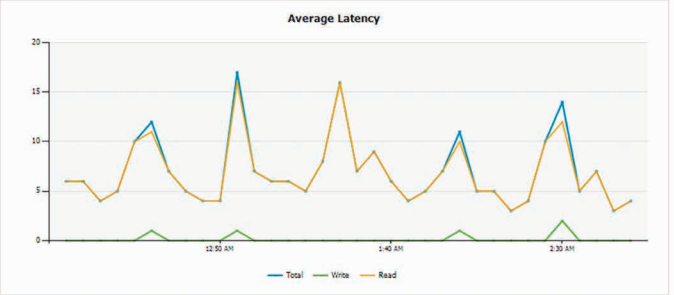
This report will help you identify storage systems that are under heavy pressure or at its maximum load. Let Veeam ONE run at least 24 hours and check if there are high latency situations. Check in the **Alerts** tab of Veeam ONE Monitor if there are specific errors that need to be addressed before you bring extra load to the environment with backup processing that can cause business critical situations.

**Datastore Datacenter01\esx32-ds1 Details**

**Latency**

**Datastore Average Latency (milliseconds)**

| Latency | Average | Minimum | Maximum | Trend |
|---|---|---|---|---|
| Total | 6.97 | 3 | 17 | Decreasing |
| Write | 0.15 | 0 | 2 | Increasing |
| Read | 6.82 | 3 | 16 | Decreasing |

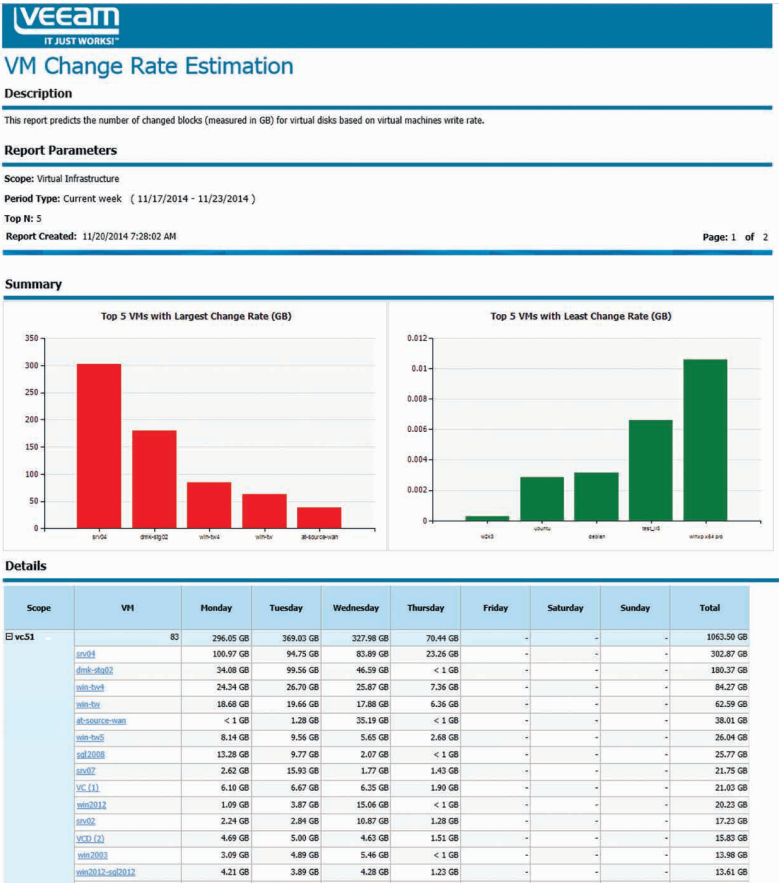



**Total Latency by Host (milliseconds)**

| Latency | Host | Average | Minimum | Maximum | Trend |
|---|---|---|---|---|---|
| Total | | | | | |
| | vcenter02\esx32.veeam.local | 6.97 | 3 | 17 | Decreasing |
| Write | | | | | |
| | vcenter02\esx32.veeam.local | 0.15 | 0 | 2 | Increasing |
| Read | | | | | |
| | vcenter02\esx32.veeam.local | 6.82 | 3 | 16 | Decreasing |

- Veeam ONE report: *Change Rate Estimation*

This report will help you identify VMs with a high change rate at the block level (relevant for incremental backups). You can later configure the backup or replication job to process them at the beginning of the backup window, to address the longer job runtimes. In general, this report will give you numbers for backup target storage planning.



VM Change Rate Estimation

**Description**

This report predicts the number of changed blocks (measured in GB) for virtual disks based on virtual machines write rate.

**Report Parameters**

**Scope:** Virtual Infrastructure

**Period Type:** Current week ( 11/17/2014 - 11/23/2014 )

**Top N:** 5

**Report Created:** 11/20/2014 7:28:02 AM

**Page:** 1 **of** 2

**Summary**

Top 5 VMs with Largest Change Rate (GB)

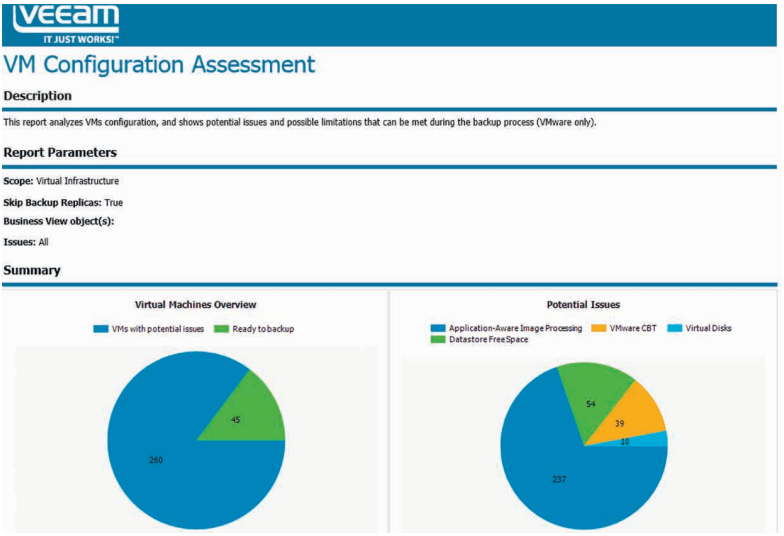
Top 5 VMs with Least Change Rate (GB)

**Details**

| Scope | VM | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | Total |
|---|---|---|---|---|---|---|---|---|---|
| ⊟ vc51 | 83 | 296.05 GB | 369.03 GB | 327.98 GB | 70.44 GB | - | - | - | 1063.50 GB |
| | srv04 | 100.97 GB | 94.75 GB | 83.89 GB | 23.26 GB | - | - | - | 302.87 GB |
| | dmk-stg02 | 34.08 GB | 99.56 GB | 46.59 GB | < 1 GB | - | - | - | 180.37 GB |
| | win-tvv4 | 24.34 GB | 26.70 GB | 25.87 GB | 7.36 GB | - | - | - | 84.27 GB |
| | win-tsv | 18.68 GB | 19.66 GB | 17.88 GB | 6.36 GB | - | - | - | 62.59 GB |
| | at-source-wan | < 1 GB | 1.28 GB | 35.19 GB | < 1 GB | - | - | - | 38.01 GB |
| | win-tvv5 | 8.14 GB | 9.56 GB | 5.65 GB | 2.68 GB | - | - | - | 26.04 GB |
| | sql2008 | 13.28 GB | 9.77 GB | 2.07 GB | < 1 GB | - | - | - | 25.77 GB |
| | srv07 | 2.62 GB | 15.93 GB | 1.77 GB | 1.43 GB | - | - | - | 21.75 GB |
| | VC (1) | 6.10 GB | 6.67 GB | 6.35 GB | 1.90 GB | - | - | - | 21.03 GB |
| | win2012 | 1.09 GB | 3.87 GB | 15.06 GB | < 1 GB | - | - | - | 20.23 GB |
| | srv02 | 2.24 GB | 2.84 GB | 10.87 GB | 1.28 GB | - | - | - | 17.23 GB |
| | VCD (2) | 4.69 GB | 5.00 GB | 4.63 GB | 1.51 GB | - | - | - | 15.83 GB |
| | win2003 | 3.09 GB | 4.89 GB | 5.46 GB | < 1 GB | - | - | - | 13.98 GB |
| | win2012-sql2012 | 4.21 GB | 3.89 GB | 4.28 GB | 1.23 GB | - | - | - | 13.61 GB |

- Veeam ONE report: *VM Configuration Assessment*

This report will help you assess VMs readiness for performing backup with Veeam Backup & Replication. It analyzes configuration of VMs in the virtual environment and shows potential issues and possible limitations that can cause the backup process to fail or prevent VMs from being properly backed up.



VM Configuration Assessment

Description

This report analyzes VMs configuration, and shows potential issues and possible limitations that can be met during the backup process (VMware only).

Report Parameters

**Scope:** Virtual Infrastructure

**Skip Backup Replicas:** True

**Business View object(s):**

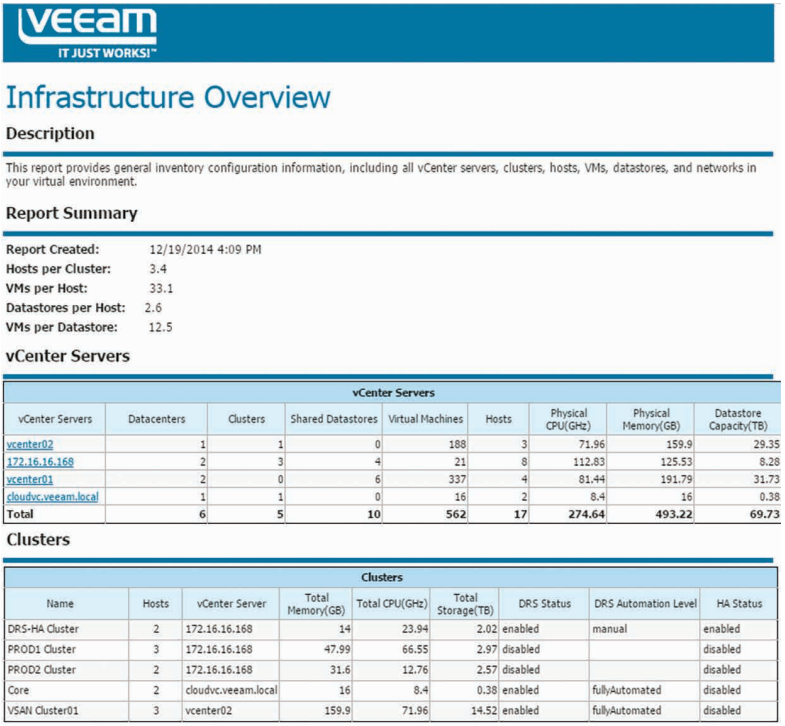
**Issues:** All

Summary



- Veeam ONE report: *Infrastructure Overview*

This report sums up the virtual infrastructure. You can use it to estimate the size of the backup infrastructure, but also to estimate the total amount of required Veeam licenses to cover the environment.



Infrastructure Overview

Description

This report provides general inventory configuration information, including all vCenter servers, clusters, hosts, VMs, datastores, and networks in your virtual environment.

Report Summary

**Report Created:** 12/19/2014 4:09 PM
**Hosts per Cluster:** 3.4
**VMs per Host:** 33.1
**Datastores per Host:** 2.6
**VMs per Datastore:** 12.5

vCenter Servers

| vCenter Servers | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| vCenter Servers | Datacenters | Clusters | Shared Datastores | Virtual Machines | Hosts | Physical CPU(GHz) | Physical Memory(GB) | Datastore Capacity(TB) |
| vcenter02 | 1 | 1 | 0 | 188 | 3 | 71.96 | 159.9 | 29.35 |
| 172.16.16.168 | 2 | 3 | 4 | 21 | 8 | 112.83 | 125.53 | 8.28 |
| vcenter01 | 2 | 0 | 6 | 337 | 4 | 81.44 | 191.79 | 31.73 |
| cloudvc.veeam.local | 1 | 1 | 0 | 16 | 2 | 8.4 | 16 | 0.38 |
| **Total** | **6** | **5** | **10** | **562** | **17** | **274.64** | **493.22** | **69.73** |

Clusters

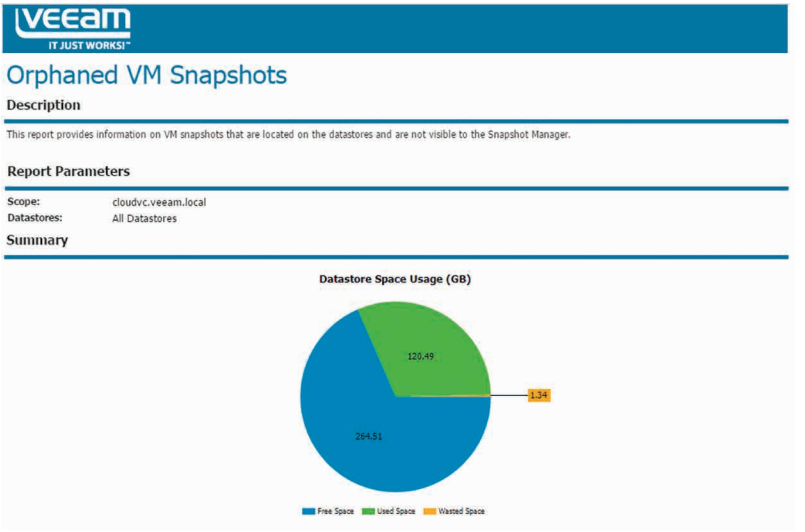
| Clusters | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Name | Hosts | vCenter Server | Total Memory(GB) | Total CPU(GHz) | Total Storage(TB) | DRS Status | DRS Automation Level | HA Status |
| DRS-HA Cluster | 2 | 172.16.16.168 | 14 | 23.94 | 2.02 | enabled | manual | enabled |
| PROD1 Cluster | 3 | 172.16.16.168 | 47.99 | 66.55 | 2.97 | disabled | | disabled |
| PROD2 Cluster | 2 | 172.16.16.168 | 31.6 | 12.76 | 2.57 | disabled | | disabled |
| Core | 2 | cloudvc.veeam.local | 16 | 8.4 | 0.38 | enabled | fullyAutomated | disabled |
| VSAN Cluster01 | 3 | vcenter02 | 159.9 | 71.96 | 14.52 | enabled | fullyAutomated | disabled |

- Veeam ONE report: *Active Snapshots*

VMware snapshots are often done to save a specific state of the VM for some time. While they are created very easily, administrators tend to forget to delete them over time. Together with administrators, you can release all snapshots that are not needed anymore. This will help prevent datastore downtimes because of snapshots filling up the storage.

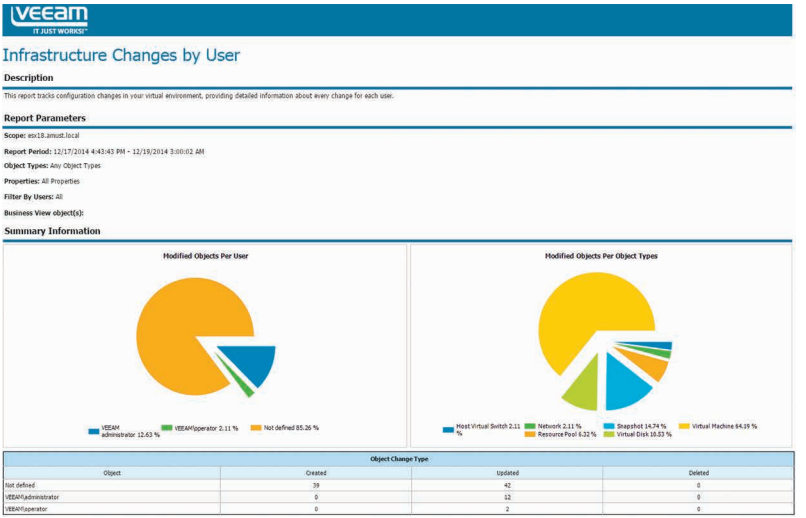

- Veeam ONE report: *Orphaned Snapshots*

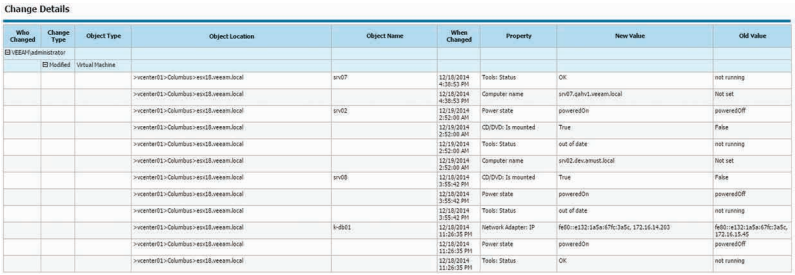
This report detects VM snapshots that are still active on datastores but do not show up in the VMware Snapshot Manager. Veeam Backup & Replication and its Snapshot Hunter will correct this situation by consolidating these snapshots, which can bring extra load at the first backup POC. We strongly recommend that you tune the VMware environment and consolidate all orphaned snapshots before starting the Backup & Replication project.




- Veeam ONE report: *Infrastructure Changes by User*

Prior to the implementation stage, it is recommended to create a separate service account to be used for all operations in the VMware vSphere environment performed by Veeam Backup & Replication. With the Infrastructure Changes by User report, you can track and document all changes done by this user.





Change Details

| Who Changed | Change Type | Object Type | Object Location | Object Name | When Changed | Property | New Value | Old Value |
|---|---|---|---|---|---|---|---|---|
| ⊟ VEEAM\administrator | | | | | | | | |
| | ⊟ Modified | Virtual Machine | | | | | | |
| | | | >vcenter01>Columbus>esxi8.veeam.local | srv07 | 12/18/2014 4:38:53 PM | Tools: Status | OK | not running |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/18/2014 4:38:53 PM | Computer name | srv07.gshv1.veeam.local | Not set |
| | | | >vcenter01>Columbus>esxi8.veeam.local | srv02 | 12/19/2014 2:52:00 AM | Power state | poweredOn | poweredOff |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/19/2014 2:53:08 AM | CD/DVD: Is mounted | True | False |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/19/2014 2:52:00 AM | Tools: Status | out of date | not running |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/19/2014 2:52:00 AM | Computer name | srv02.dev.amust.local | Not set |
| | | | >vcenter01>Columbus>esxi8.veeam.local | srv08 | 12/18/2014 3:55:42 PM | CD/DVD: Is mounted | True | False |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/18/2014 3:55:42 PM | Power state | poweredOn | poweredOff |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/18/2014 3:55:42 PM | Tools: Status | out of date | not running |
| | | | >vcenter01>Columbus>esxi8.veeam.local | k-db01 | 12/18/2014 11:26:35 PM | Network Adapter: IP | fe80::e132:1a5a:67fc:3a5c, 172.16.14.203 | fe80::e132:1a5a:67fc:3a5c, 172.16.15.45 |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/18/2014 11:26:35 PM | Power state | poweredOn | poweredOff |
| | | | >vcenter01>Columbus>esxi8.veeam.local | | 12/18/2014 11:26:35 PM | Tools: Status | OK | not running |

- Veeam ONE report: *Inventory*

This report provides the most complete and up-to-date configuration information on all objects in the virtual environment. It can be used offline at the planning phase to address any virtual infrastructure-related questions.

Optionally, it is recommended to review the VMware Optimization section in the Veeam ONE Dashboards and Reports document. A good example is the Garbage Files Report that can identify possible wasted space on datastores. In some cases, it helped to free up over 10 TB of space on the Tier 1 storage.

# Small POC – Getting Started with the Interface

Many customers decide to do a small scale Proof of Concept (POC) after seeing their first live demonstration and presentation meetings with partners or Veeam System Engineers. The idea is to get started with the interface of Veeam Backup & Replication and to test if everything works as expected/presented within the customer's environment.

As enterprise environments are sometimes very complicated from the firewall and storage perspective, in most cases customers decide to do a POC in very flat test environments. Typically, a test environment includes:

- ESXi hosts, vCenter Server, Veeam Backup & Replication server
- 10-20 VMs running various business applications

TIt is possible to carry out a Veeam Backup & Replication POC in such environment with only a single Veeam backup server on a VM with 4 cores and 6-8 GB of RAM. (Since this test is focused on the user interface experience, no special preparation is needed from the performance perspective.)

Customers often drive this POC themselves. To assist customers with this task, Veeam has published a good Evaluator's Guide that includes configuration screenshots with minimal required background information. See Veeam Helpcenter for Evaluator's Guide: http://helpcenter.veeam.com/evaluation/backup/vsphere/en/ (VMware vSphere environments) and http://helpcenter.veeam.com/evaluation/backup/hyperv/en/ (Microsoft Hyper-V environments).

# Architecture Meeting and Enhanced POC Preparation

Based on the information gathered during the assessment phase and customer requirements, you may design a solution on paper and possibly implement it. Most likely such designs are going to change over multiple revisions during the implementation phase after communicating with other departments e.g. security, networking and storage teams. It may also happen that the customer comes up with the new demands based on new findings. This may delay in the implementation and ultimately lead to increased cost.

This chapter about The Architecture Meeting should help you avoiding such situations. We will explain how the approach used by Veeam architects can help you simplify and streamline the design phase and steer all project participants towards the same goals. This will optimize the implementation phase and ultimately cut cost due to less time spent revising the design and realigning stakeholders.

# The Architecture Meeting

This section describes how to hold an infrastructure discovery and assessment session with a customer. Below is an example of how Veeam Architects hold such meetings this with customers. The example below is just one example of many possible ways of the meeting implementation; please have a look at other chapters of this guide to prepare for such meeting.

## Infrastructure Discovery

1. Start with the first main customer datacenter. Figure out the following:
    a. Virtualization platform and version
    b. Main storage system, type, connection

c. Is storage virtualization used (between the storage arrays and hypervisor)?

2. Depict the second main customer datacenter (if available)

    a. Are there any storage replication/mirroring involved?

    b. Is Active/Active cluster used? For proper backup proxy implementation and backup mode selection, it is important to know where the data that you want to back up is located, and whether you can access all data from a single site.

3. Obtain information about network connections:

    a. Is there 10 GbE LAN?

    b. Is there a WAN connection between the 2 datacenters? This is necessary to know if you plan to use the Virtual Appliance or Network backup mode.

    c. What is the VMKernel Interface physical link speed? 10GbE gives you faster processing for the Network mode. To learn more, see the "Backup Proxy" chapter.

    d. Is vCenter Server physical or virtual? Where is it located?

4. Define the amount of production data:

    a. Number of VMs (this can help to design jobs)

    b. Used data (this can help to define the backup target and configure jobs settings)

    c. Number of ESXi hosts and number of used sockets (this regards Veeam licensing).

    d. Number of clusters

    e. Other information

5. Create the first Veeam implementation draft/sample scenario:

    a. Start with the repository, discussing customer demands. In the example, customer wanted to have the backup data in both datacenters. If so, you could decide to implement repositories on both sides (half of the data on each side) and use the backup copy job for replicating data to the second site.

b. Discuss proxy implementation. The customer agreed to implement physical proxy servers connected to their Fibre Channel network. As the customer used thick-provisioned VMware VM disks, this ensured a fast and reliable backup and restore. Check out the "Backup Proxy" section of this guide to determine the best proxy implementation and select a transport mode for the environment.

c. Plan for the backup server. In thie example, it was placed on a VM and replicated to the second datacenter. (The underlying datastore of the VM was not replicated /mirrored to the second site.)

d. Add other required components. The customer was already useing two IBM TS3500 libraries for long-term retention with the existing backup software (agents). They prepared a partition on each library with 4x LTO6 drives for use with Veeam. You would proceed and connect them to the 2 physical servers (having the proxy and repository roles assigned), and additionally assign the tape server role to these servers.

6. Define OS/applications:

   a. Create a list of used operating systems.

   b. Create a list of all applications starting with the most critical. Find out whether Microsoft SQL and Microsoft SharePoint are used, as it can influence the version and type of the Microsoft SQL Server on which the Veeam configuration database must be deployed (Express Edition may be not sufficient).

7. Define business-critical applications/VMs to plan for availability. Planning for backup is very important for them, as this mainly influence the RPO and stability of existing applications. It is even more important to plan for disaster recovery scenarios.

   a. Define the number of VMs that are business critical.

   b. Find out whether slower performance is OK at disaster recovery (consider using Instant VM Recovery). In this example, the customer used a third small datacenter with a single storage system (Quorum) for the storage virtualization. During the discussion the customer identified 50 VMs that were business-critical and needed full performance even at disaster recovery.

Thus, in the next step, you would add 2 ESXi hosts to that Quorum datacenter and replicate these 50 VMs every hour to that datacenter. The connection speed is to be 10 GbE. So, in case of disaster recovery the customer could just boot up all VMs with full speed.

**Important!**

It is very important to use all available Veeam possibilities to implement the best RTO and RPO times in customer's environment. For the VM recovery scenario, you can mix classic VM restore (best for small VMs), Instant VM Recovery (best for huge data servers) and VM replica failover (best for database systems with extreme I/O requirements). Together with the customer, check the "possible failure areas" (single storage system/ whole datacenter/ 1 datastore) and decide if the designed Veeam implementation fits into these needs and is in line with the budget.

## Network and Firewall

Veeam Availability Suite is very flexible and lets you implement different backup infrastructure schemes. Firewalls can be used between all backup infrastructure components. The only exception is RPC inspection functionality: it can cause delays in connections, and Veeam Backup & Replication can run into timeouts.

However, the best practice is to place backup infrastructure components in the same network segment as the corresponding VMware components to allow for efficient and fast usage of the network bandwidth.

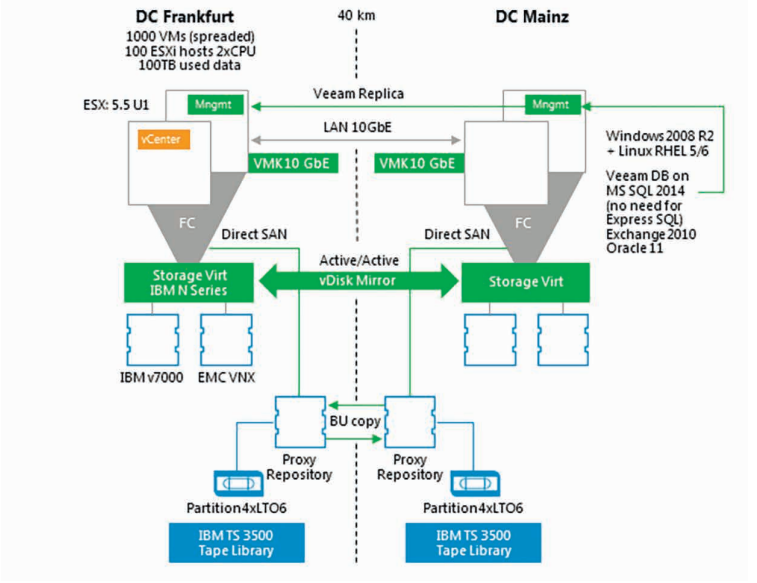
## Proxy/Repository Systems

Proxy and repository servers should be placed in the VMKernel networks. Veeam Backup & Replication uses the VMKernel interfaces to read out configuration data and disk data (in case of NBD), and to map Veeam vPower NFS datastores for fast recovery (Instant VM Recovery).

# Backup & Replication Server (Management)

As the backup server communicates mainly with the vCenter Server and other backup infrastructure components, it should be placed next to the vCenter Server in most cases.

The backup infrastructure for this sample scenario would look as follows:




# Veeam ONE

Veeam ONE components should be placed next to the vCenter Server and should be able to read from the backup server and ESXi hosts (over the CIM protocol) as well. See Veeam ONE documentation for more information: http://helpcenter.veeam.com/one/80/deployment/index.html?introduction.html.

# Veeam Backup Enterprise Manager

When Veeam Backup Enterprise Manager is used with Self-Restore Services, it should be placed in the internal DMZ in most cases.

## Restore Points

In the sample case, the customer needed daily backup with 14 restore points; the points were to be held on 2 sites (copied with backup copy job). The customer also wanted to offload the weekly full backups on tape and hold them for a period slightly longer than one year in both tape libraries.

The customer also needed to replicate the most critical VMs to the Quorum datacenter hourly, between 7:00 and 19:00. The number of replication restore points to be maintained was the maximum possible (here 28 restore points).

In many architecture meetings, planning for the retention policies is the most time-consuming part as you are likely to engage different administrators and management team members from different departments in this process. These team members have to translate their file-based existing restore point policies into a new way (image-level backup). It is a matter of concern because a longer retention chain will result in expensive storage space costs.

**Important!**

Remember to agree on backing up Microsoft SQL Server transaction logs with Veeam Backup & Replication.

If speaking about the storage sizing, the tool at http://vee.am/rps can help to illustrate the retention chains on disk and estimate the required capacity.

# Enhanced POC

After you have agreed and discussed the points above, proceed with the enhanced POC to demonstrate that Veeam Availability Suite can work in customer's environment with excellent speed.

Typically, the enhanced POC is carried out under the following conditions:

- The environment is close to the production environment, with all firewalls in place.
- Involved storage systems are similar to the production storage systems.

- Veeam storage integration is used whenever possible.
- To demonstrate the good working load balancing and scalability, 100-200 VMs are backed up/replicated.
- All major applications are backed up to test all restore scenarios.

## Preparation

1. Prepare for the POC planning with the Veeam User Guide and this document.
2. Check out the neecessary firewall ports and help the customer with the internal firewall change requests. Refer to the corresponding sections in the User Guide and this document.

**Tip**

Do the firewall planning very carefully: if something is not set up correctly, this will block the whole POC. In most cases, it is not easy to detect problems and do firewall changes when the POC is already running. However, it is a good idea to ask the customer to have the firewall administrator at hand in case you need an urgent change.

3. Create a separate vCenter Server account for Veeam ONE (*read-only* + *datastore browsing* + *CIM*) and Veeam Backup & Replication http://www.veeam.com/veeam_backup_8_permissions_pg.pdf so that you are able to track what users do.
4. If you want to use the storage integration feature, check out the corresponding chapter in this guide, set up the storage and the SAN network together with the storage administrators. Limit the scope of storage system rescan to the volumes used in the POC (requires Veeam Backup & Replication v8 Update 2 to be installed).
5. If you want to use SureBackup, make sure that a virtualized Domain Controller is present if needed (e.g., for Microsoft Exchange).
6. Let the customer prepare all used antivirus systems upfront so that you do not run into trouble. Check the “Antivirus” section of this guide.
7. Ask the customer to prepare a fair performing storage system for the POC. Avoid cheap NAS boxes for such big POC.
8. Let the customer prepare all operating systems and database installations. Set up Veeam Backup & Replication and backup infrastructure components together with the customer and place the folders correctly.

# Automation

The bigger is the environment, the more automation is needed to reduce the administration effort. For example, if operating 40 branch offices with independent Veeam installations, you may want to roll out and configure backup servers with scripts and automatically create jobs there as well. Another example is automatic job creation for 2.000-3.000 VMs with exactly the same configurations, which can limit user-caused misconfiguration.

## PowerShell

Operations in Veeam Backup & Replication can be automated with Veeam's own PowerShell plugin in the following areas:

- Installation
  See http://www.veeam.com/kb1833 for more details about the full rollout. Be aware that with automated rollouts, all software prerequisites need to be properly set up beforehand. So you are likely to write some automation rollout scripts for these packages as well.
- Updates
  http://helpcenter.veeam.com/backup/80/vsphere/index.html?update2_unattended.html
- Configuration
- Job creation/job editing
- Working with external schedulers (UC4/TWS and other) to start Veeam jobs
- Restores
- Reporting
- Datacenter migration (quick migration or replication)

The PowerShell plugin is available with all commercial versions of the product.

**Note**

Starting with v8 Update 2, PowerShell plugin is also available with Veeam Backup Free, although limited: http://www.veeam.com/blog/veeam-backup-free-edition-now-with-powershell.html

Our customers and partners use this functionality to scale out backup infrastructure environments to nearly 100 000 VMs under a single Veeam Backup Enterprise Manager instance with multiple backup servers located in different datacenters.

The best starting point to get in touch with the Veeam PowerShell plugin is to read the Veeam PowerShell User Guide:
http://www.veeam.com/veeam_backup_8_powershell_pg.pdf .

You can find help for the scripts in the Veeam Community Forum:
http://forums.veeam.com/powershell-f26/

If you need some examples, refer to the following section:
http://forums.veeam.com/powershell-f26/getting-started-and-code-examples-t13372.html

## RESTful API

In the Veeam Enterprise Manager, there is as well RESTful API that allows you to create workflows in orchestration tools or to integrate Veeam Backup Enterprise Manager (self-services) in your own "cloud" portal. Specifically, this is an option that comes with Enterprise Plus Editions and is focused on the hosting business.

User Guide: http://www.veeam.com/veeam_backup_8_web_api_pg.pdf

Community Forum: http://forums.veeam.com/restful-api-f30/

Examples:
http://helpcenter.veeam.com/backup/80/rest/beginner_example.html

# APPENDIX 1. VEEAM BACKUP & REPLICATION ANATOMY

You might have a basic understanding of how Veeam Backup & Replication components interact, but do you know what happens in detail with each component when you backup a VM, do a standard VM restore, an Instant VM Restore, a Windows File-Level restore, or replicate a VM? The next sections are dedicated to explaining in detail what actually happens during these processes.

## Backup

This section provides a step-by-step description of a VMware virtual machine backup process implemented in Veeam Backup & Replication.

### 1. Initialization Phase

A backup job can be started automatically or manually in the Veeam Backup & Replication console, Veeam Backup Enterprise Manager web console, by means of PowerShell, RESTful API and other.

In the initialization phase, Veeam Backup & Replication prepares resources necessary for a backup job. To help you better understand firewall settings and connection initiation flow, the process is illustrated by the diagram (see below):

1. When a backup job is initialized, the Veeam Backup Manager process is started on the Veeam backup server.

2. Veeam Backup Manager reads job settings from the Veeam Backup configuration database and creates a list of VM tasks to process (one task stands for one VM disk).

3. Veeam Backup Manager connects to the Veeam Backup Service. The Veeam Backup Service includes a resource scheduling component for managing all tasks and resources in the backup infrastructure. The resource scheduler checks what resources are available, and assigns backup proxies and repositories to process that job tasks using Veeam's load balancing.

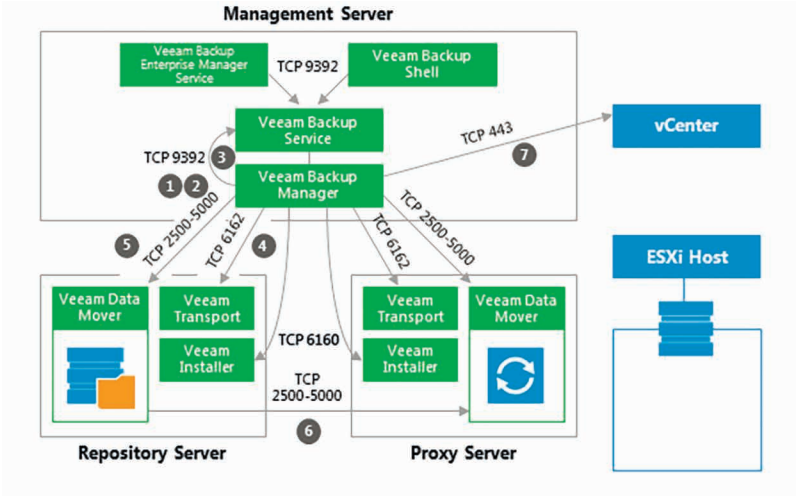
4. After the necessary backup infrastructure resources have been assigned, Veeam Backup Manager connects to the Transport Services on the target repository and on the backup proxy. The Transport Services, in their turn, start the Veeam Data Movers.

On the backup proxy, a new Veeam Data Mover is started for each task that the proxy is processing.

5. Veeam Backup Manager establishes a connection with Veeam Data Movers on the backup repository and backup proxy, and sets a number of rules for data transfer (such as network traffic throttling rules, and so on).

6. Veeam Data Movers on the backup proxy and repository establish a connection with each other for data transfer.

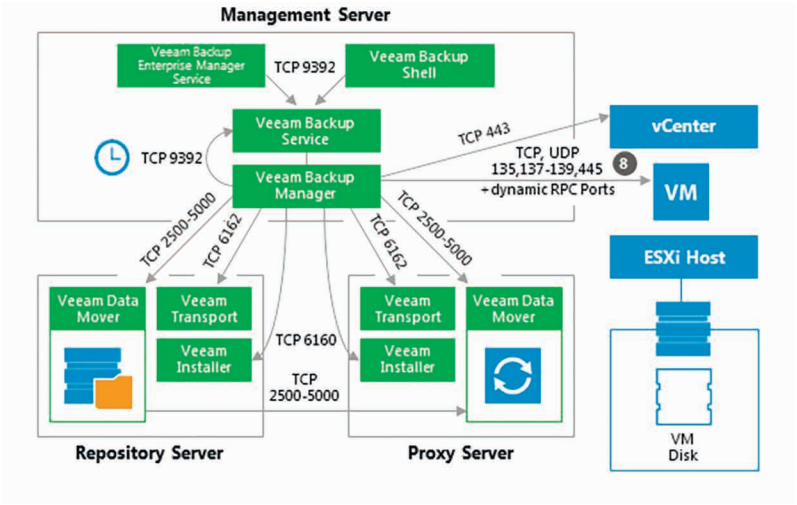
7. Veeam Backup Manager connects to the vCenter Server or ESXi host and gathers metadata about VMs and hosts engaged in the backup process. At this step, no connection between the Veeam backup server and VM guest networks is established.




## 2a. Guest Processing for Windows-Based VMs

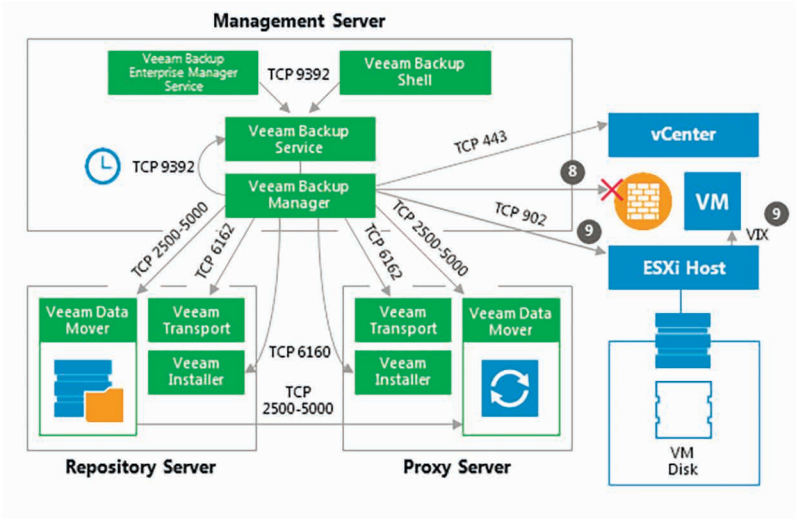
For VMs with Microsoft Windows guest OS, Veeam Backup & Replication obtains information about the guest's IP addresses from VMware Tools. Veeam uses these IP addresses to connect to the guest OS and perform in-guest processing tasks (if application-aware image processing is enabled).

If it is not possible to connect to the guest OS or the connection is blocked by a firewall, Veeam Backup & Replication tries to establish a connection using VIX, as described in section 2b.




# 2b. Guest Processing for Windows-Based VMs (VIX)

If there is no network connectivity to the VM guest OS, Veeam Backup & Replication uses the communication channel provided by VMware Tools (VIX) to interact with the guest OS and perform in-guest processing tasks.



# 2c. Guest Processing for Linux/Unix-Based VMs

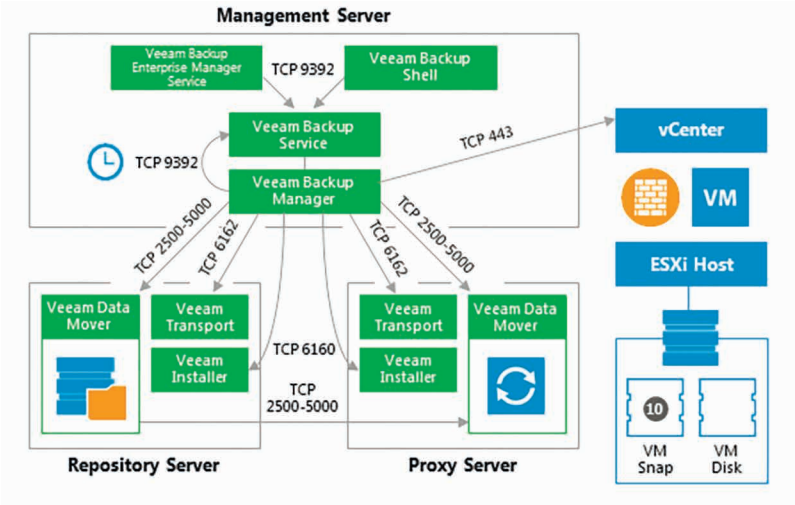
If pre-freeze and post-thaw scripts are enabled in the backup job properties, Veeam Backup & Replication obtains information about the guest's IP address from VMware Tools. Veeam uses this IP address to connect to the guest network over SSH and perform in-guest processing tasks. Scripts reside on the backup server and are injected in the guest OS at the time of backup.

If there is no network connectivity with a Linux-based VM, Veeam Backup & Replication will not fail over to the VIX communication channel. In such cases, as an alternative method, you can use VMware Tools quiescence and let VMware Tools run the necessary scripts that will need to be created inside the guest OS (see location details for Windows / Linux guest at: https://pubs.vmware.com/vsphere-50/topic/com.vmware.datarecovery.admin.doc_20/GUID-6F339449-8A9F-48C0-BE70-91A2654A79D2.html.

However, it is recommended to use Veeam's functionality to call pre-freeze and post-thaw scripts, as this method is more controllable by the Veeam code: all errors that occur during the backup process are written to Veeam logs (not VMware Tools).

# 3. Creating a VM Snapshot

Now, Veeam Backup & Replication requests the vCenter Server or ESXi host to initiate a VM snapshot creation. A VM snapshot is required to use VMware VADP backup methods and leverage features like VMware Changed Block Tracking (CBT).




# 4. Releasing the Guest OS Activities

Right after the VM snapshot is taken, all quiesced disk I/O activities in the guest OS are resumed.

# 5. VM Data Transport

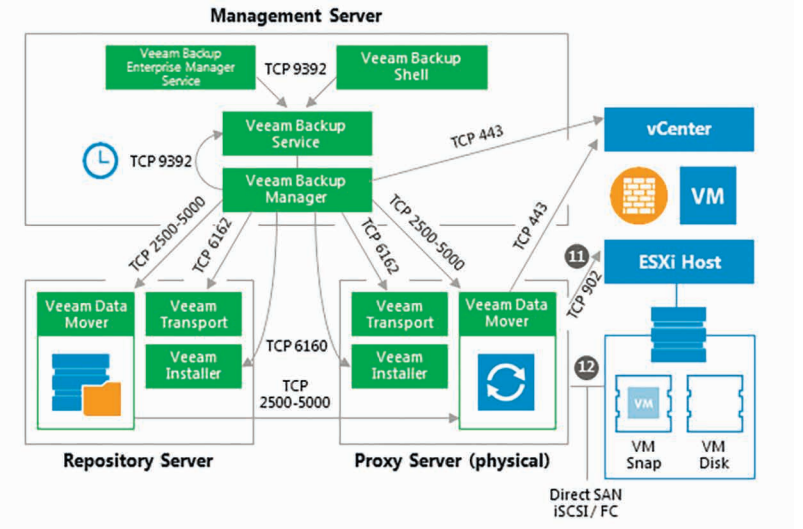
To read and transfer data from the VM snapshot, Veeam Backup & Replication can use one of the following transport modes:

- Direct SAN Access
- Virtual Appliance (HotAdd)
- Network (NBD)

For more information about each transport mode, see Veeam Backup & Replication User Guide or a corresponding section below.

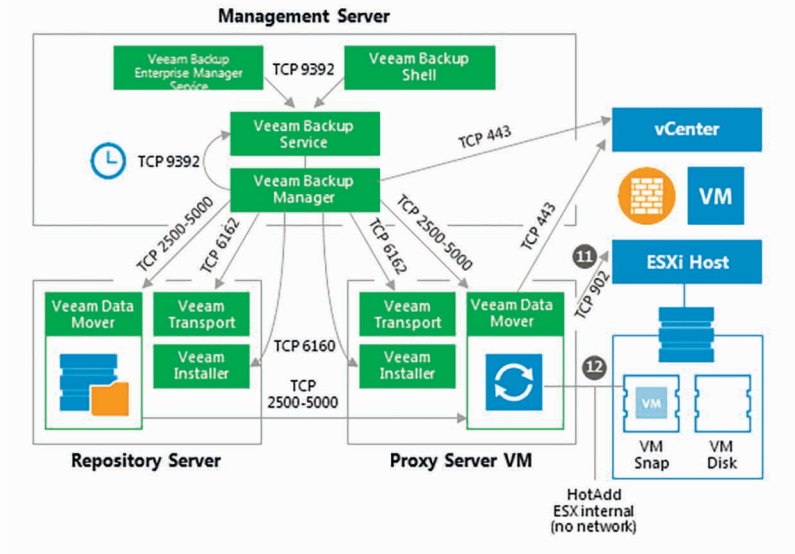
# 5a. Direct SAN Access Data Transport Mode

In the Direct SAN Access mode, Veeam backup proxy connects to the ESXi host where the VM resides, and reads the necessary VM configuration files (such as *.vmx). Backup proxies use VM configuration details to read VM data directly from the SAN.



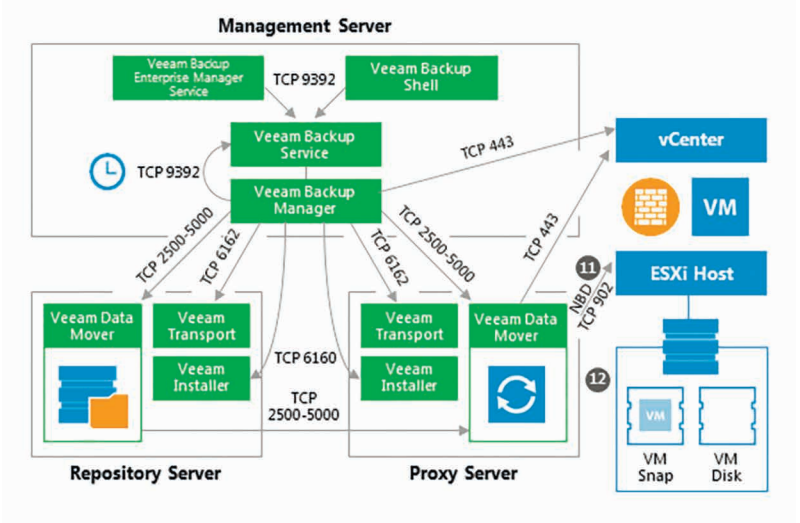
# 5b. Virtual Appliance Data Transport Mode

In the Virtual Appliance transport mode, Veeam backup proxy connects to the ESXi host where the VM resides, and reads the necessary VM configuration files (such as *.vmx). VM disks as of the snapshot state are hot-added to a virtualized Veeam backup proxy. The proxy reads VM data and unmaps the VM disks when finished.



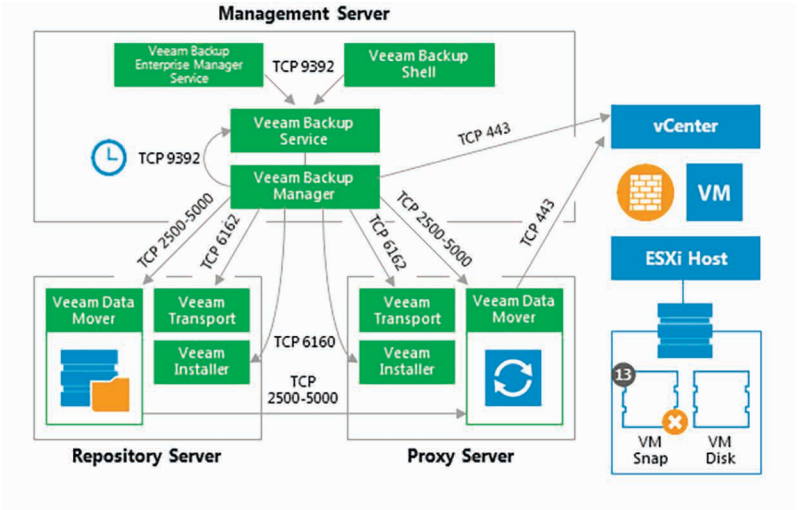
# 5c. Network Data Transport Mode

In the Network transport mode, Veeam backup proxy connects to the ESXi host where the VM resides, and reads the necessary VM configuration files (such as *.vmx). In this mode, the same data channel is used to read VM disk data, too.




# 6. Committing VM Snapshot

After Veeam backup proxy finishes reading VM data, Veeam backup server requests the vCenter Server or ESXi host to initiate a VM snapshot commit.



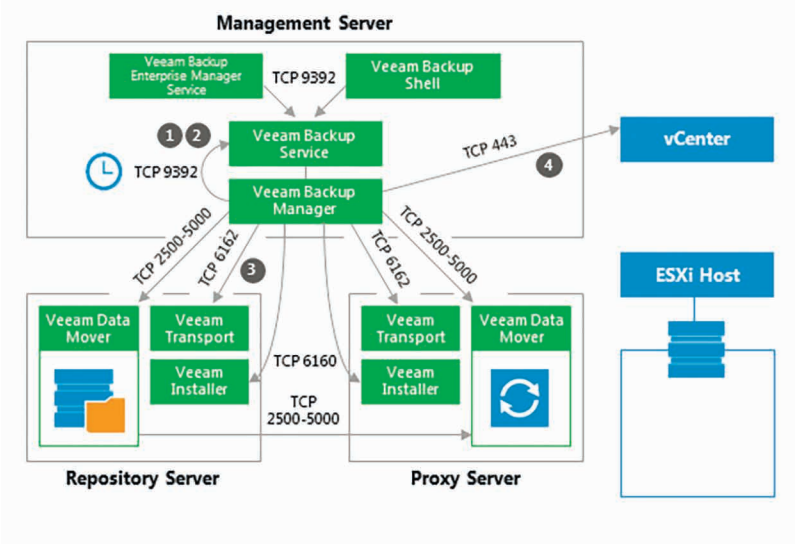
# VM Restore

This section provides a step-by-step description of a full virtual machine restore process implemented in Veeam Backup & Replication.

## 1. Initialization Phase

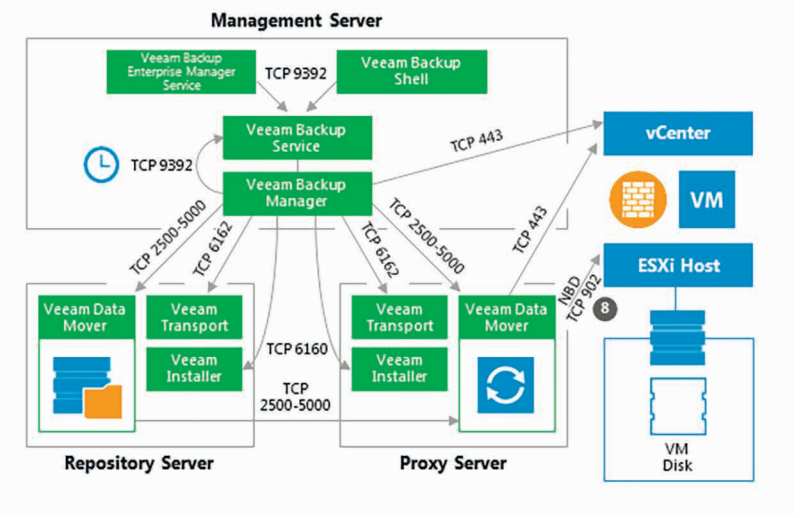
In the initialization phase, Veeam Backup & Replication prepares the resources necessary for full VM recovery. It performs the following steps:

1. Starts the necessary processes on the Veeam backup server.
2. Checks available backup infrastructure resources and assigns a proxy server for transferring restored VM data to the target host/datastore.
3. Communicates with Transport Services on the backup proxy and backup repository where the backup files reside. Transport Services, in their turn, start Veeam Data Movers. Veeam Data Movers on the backup proxy and repository establish a connection with each other for data transfer.
4. Connects to the vCenter Server or ESXi host where the restored VM will be registered.



# 2. Restoring VM Configuration

Veeam Backup & Replication retrieves VM configuration data from the backup and restores it on the chosen ESXi host/datastore. Next, it instructs VMware vSphere to register the restored VM on the host. If a user selects to change VM configuration (for example, disk format or network settings) during restore, Veeam makes the necessary amendments.

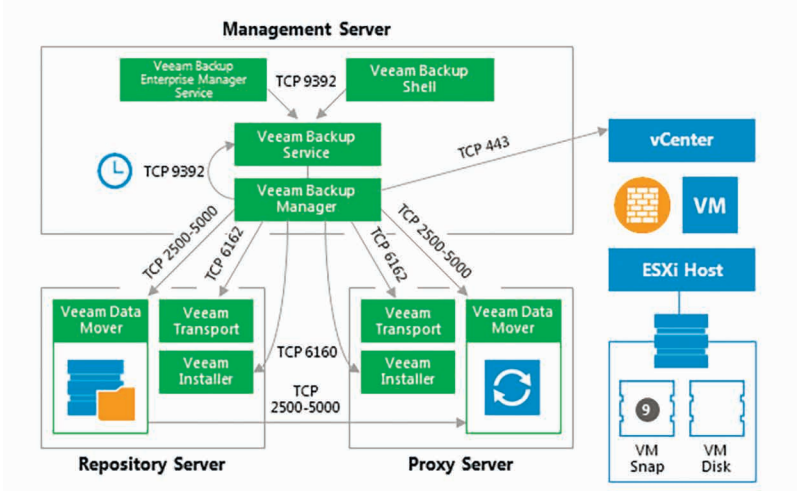


# 3. Creating VM Snapshot

Veeam Backup & Replication requests the vCenter Server or ESXi host to initiate a VM snapshot creation on the restored VM.

**Important!**

A snapshot is not taken if a VM is restored to a VVOL datastore due to vSphere VDDK limitations (see https://www.vmware.com/support/developer/vddk/vddk-600-releasenotes.html#compatibility).




# 4. VM Data Transport

Veeam Backup Manager instructs VMware vSphere to create virtual disks for the VM.

To write VM disk data to the target datastore, Veeam Backup & Replication can use one of the 3 transport modes:

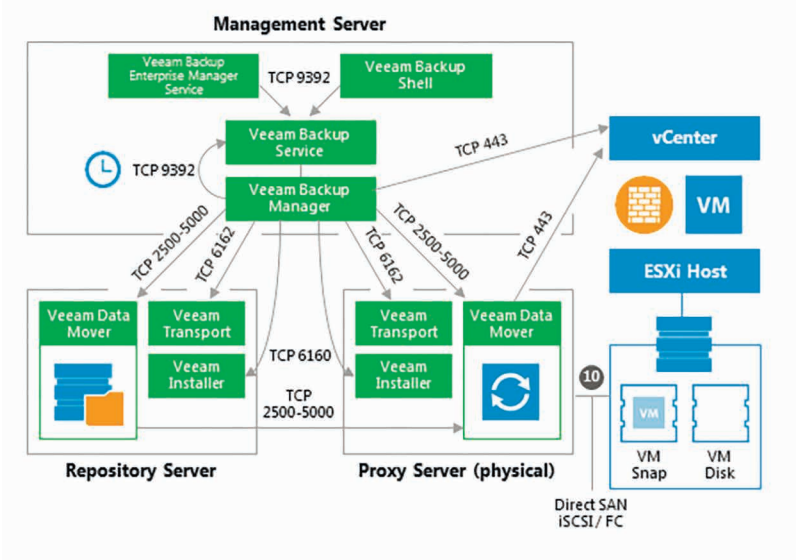
- Direct SAN Access
- Virtual Applicance (HotAdd)
- Network (NBD)

For more information about each transport mode, see Veeam Backup & Replication User Guide and the corresponding sections of this document.

# 4a. Direct SAN Access Data Transport Mode

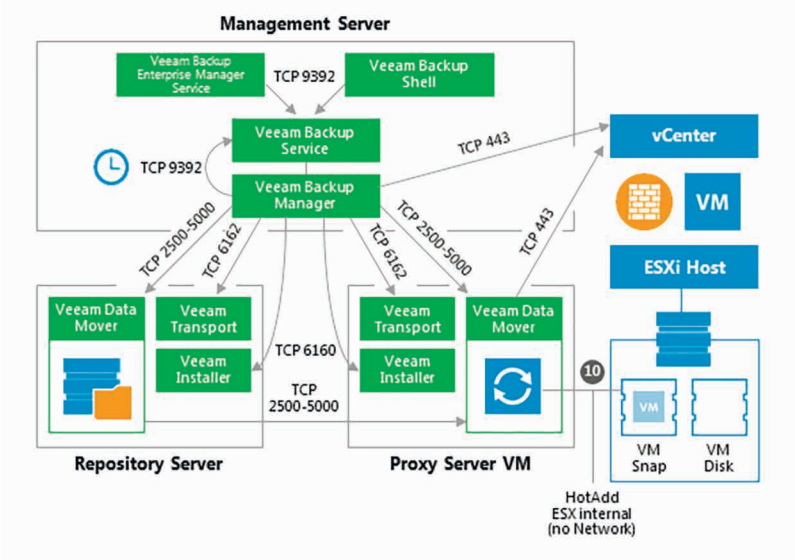
This mode is available only for VMs that have all disks in thick provisioning.

In the Direct SAN Access mode, Veeam Backup & Replication connects to the ESXi host where the restored VM is registered. The ESXi host locates the VM disks, retrieves metadata about the disk layout on the storage, and sends this metadata to the backup proxy. The backup proxy uses this metadata to copy VM data blocks to the datastore via SAN.



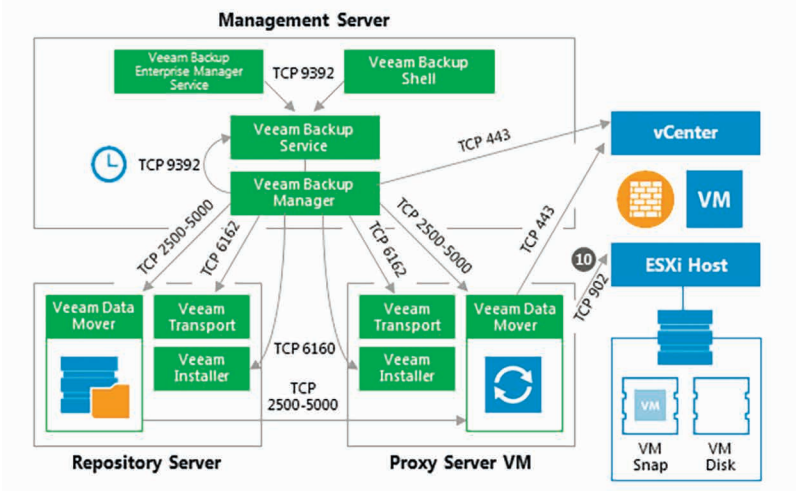
# 4b. Virtual Appliance Data Transport Mode

In the Virtual Appliance transport mode, VM disks from the backup are hot-added to a virtualized Veeam backup proxy. The proxy connects to the ESXi host where the restored VM resides and transfers disk data to the target datastore through the ESX(i) I/O stack. When the data transfer process is finished, disks are unmapped from the backup proxy.



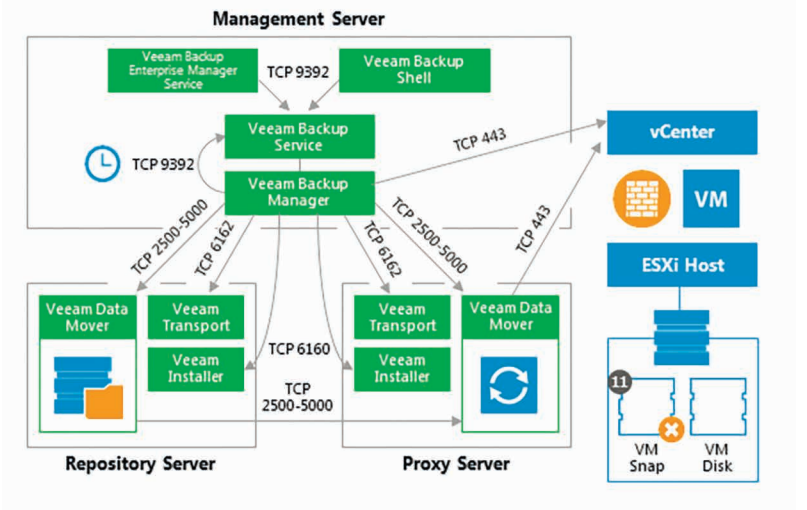
# 4c. Network Data Transport Mode

In the Network transport mode, Veeam backup proxy connects to the ESXi host where the restored VM resides, and writes VM disk data to the target datastore through the LAN channel.




# 5. Committing VM Snapshot

After the proxy finishes writing VM disk data, Veeam Backup & Replication requests the vCenter Server or ESXi host to initiate a snapshot commit for the restored VM.



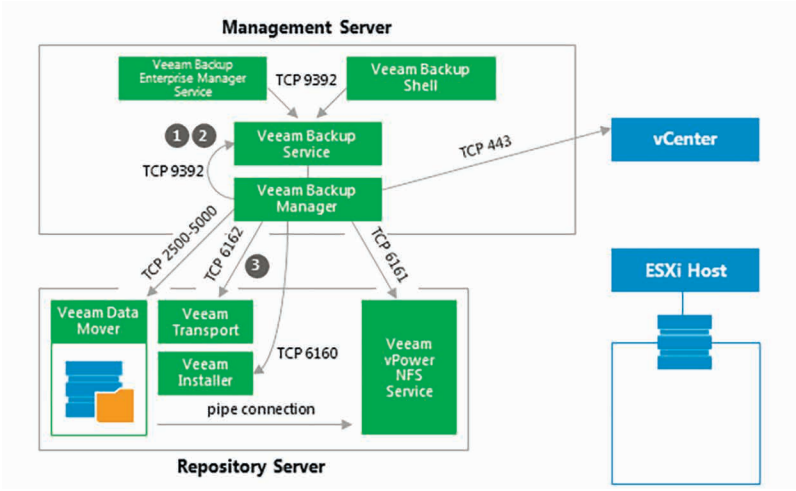
# Instant VM Recovery

This section provides a step-by-step description of the Instant VM Recovery process implemented in Veeam Backup & Replication.

## 1. Initialization Phase

In the initialization phase, Veeam Backup & Replication prepares resources necessary for Instant VM Recovery. It performs the following steps:

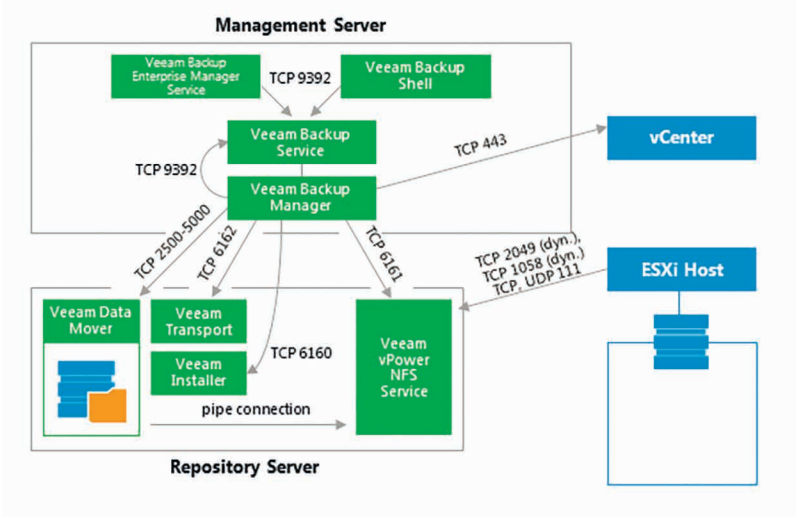
1. Starts the Veeam Backup Manager process on the Veeam backup server.
2. Checks with the Veeam Backup Service whether the necessary backup infrastructure resources are available for instant VM Recovery.
3. Communicates with the Transport Service on the backup repository to start Veeam Data Mover.




## 2. NFS Mapping

When backup infrastructure resources are prepared, Veeam Backup & Replication maps an empty NFS datastore to the selected ESXi host. It uses the Veeam vPower NFS Service for this purpose.

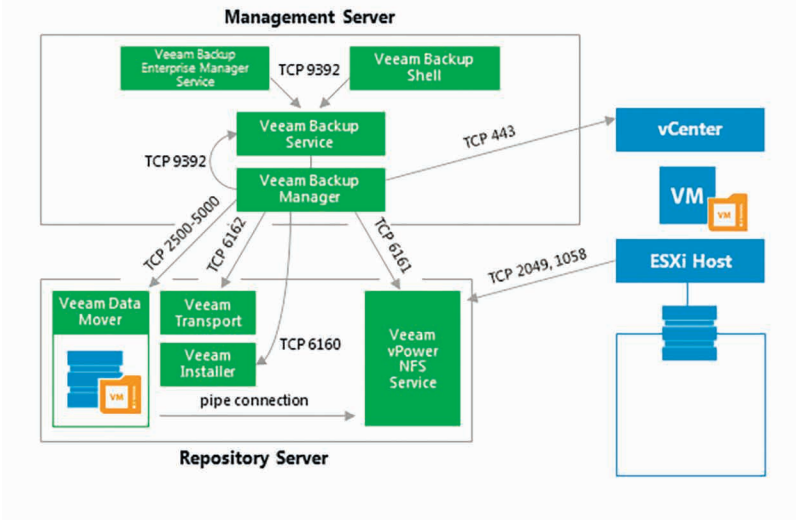
Next, Veeam Backup & Replication creates in the Veeam NFS datastore VM configuration files and links to virtual disk files. Virtual disk files remain in the backup on the repository, while all changes to these files are written to the cache file.




# 3. Registering and Starting VM

The VM runs from the Veeam NFS datastore. VMware vSphere treats the Veeam NFS datastore as any regular datastore. For this reason, with the recovered VM you can perform all actions that vCenter Server/ESXi supports for regular VMs.

To migrate VM disk data to a production datastore, use VMware Storage vMotion or Veeam Quick Migration. For details, see Veeam Backup & Replication User Guide.



# Windows File-Level Restore

This section provides a step-by-step description of Microsoft Windows file-level restore process for a VMware virtual machine implemented in Veeam Backup & Replication.

## 1. Initialization Phase

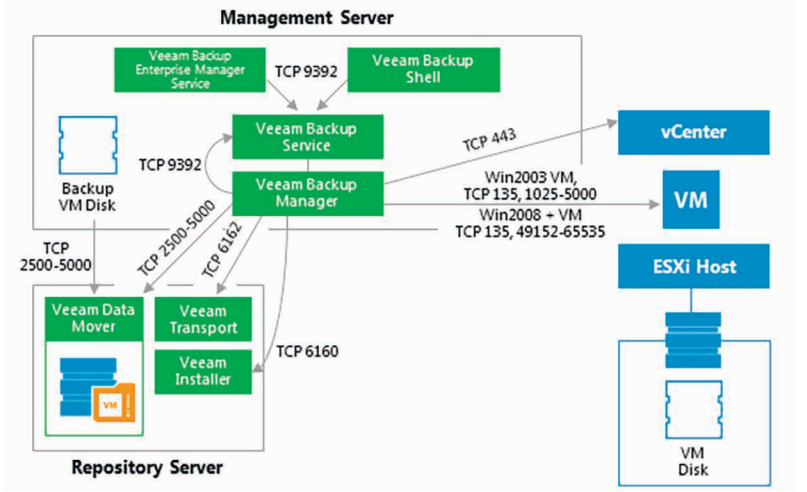
In the initialization phase, Veeam Backup & Replication prepares resources necessary for Microsoft Windows file-level restore. It performs the following steps:

1. Checks with the Veeam Backup Service whether the necessary backup infrastructure resources are available for Microsoft Windows file-level restore.
2. Starts Veeam Data Movers on the Veeam backup server and backup repository.
3. Mounts the content of backup files to the backup server with the help of Veeam's proprietary driver.

The backup files remain on the backup repository. Guest files inside the backup can be accessed in Veeam Backup browser or Microsoft Windows File explorer on the backup server, mapped by default in the *C:\Veeam-FLR* folder (can be changed via registry key).

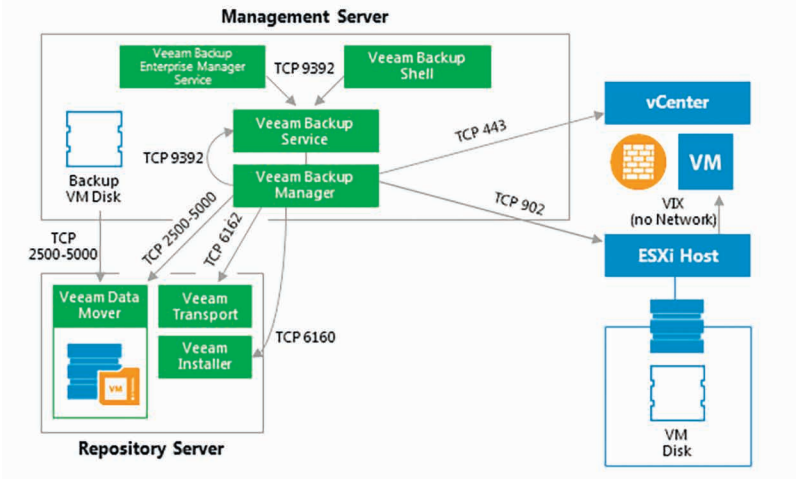
## 2a. Restoring Windows Guest OS Files (Network-Based)

To restore guest files back to the original VM, Veeam Backup & Replication establishes a connection with the VM Guest OS. It obtains information about the guest IP address from VMware Tools. Veeam uses this IP address to connect to the guest OS and perform in-guest file recovery.




# 2b. Restoring Windows Guest OS Files (Networkless)

If there is no network connectivity with the VM guest OS, Veeam Backup & Replication uses the communication channel provided by VMware Tools (VIX) to interact with the guest OS and perform in-guest file recovery.




# 3. Dismounting Backup Content

After all restore activities are completed and the user closes the Veeam Backup browser (or the browser is closed by timeout), the content of the backup files is dismounted from the backup server.

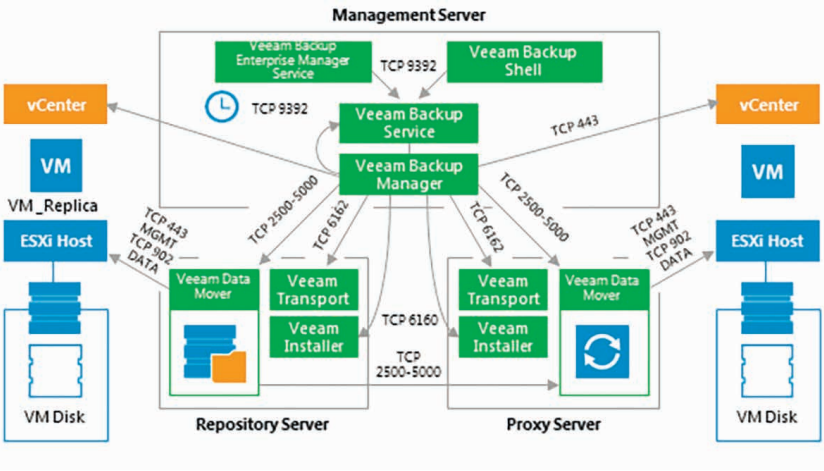
# Replication

This section provides a step-by-step description of a VMware virtual machine replication process implemented in Veeam Backup & Replication.

In many aspects, the replication initialization phase is similar to the initialization phase of the backup process. Veeam Backup & Replication starts the necessary processes, builds the list of VMs to replicate, assigns backup infrastructure resources for the job and starts Veeam Data Movers on two backup proxies (source and target) and the backup repository that is used for storing replica metadata.

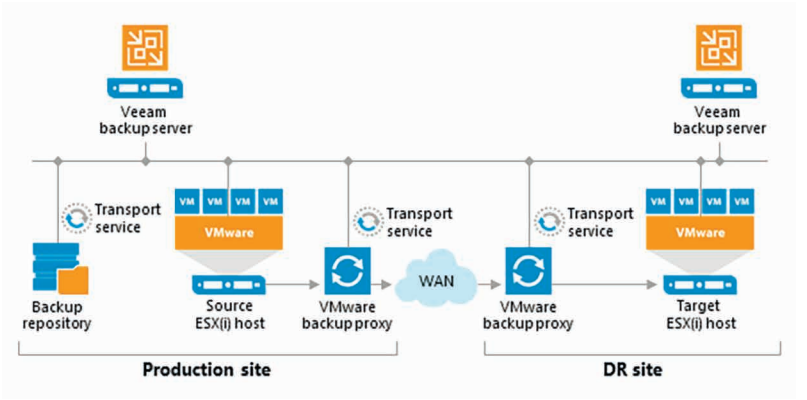
Next, Veeam Backup & Replication performs in-guest processing tasks, triggers VM snapshot creation, registers a replica VM on the target host and performs data transfer from the source host and datastore to the target host and datastore. The source and target proxies can use one of 3 available data transport modes for reading data from source and writing data to target.

This diagram illustrates the replication process with the NBD transport mode used for reading and writing VM data. For examples of the Direct SAN Access and HotAdd transport modes, see the "Backup Anatomy" section above in this Appendix.




Note that Veeam uses backup repository to store replica metadata.

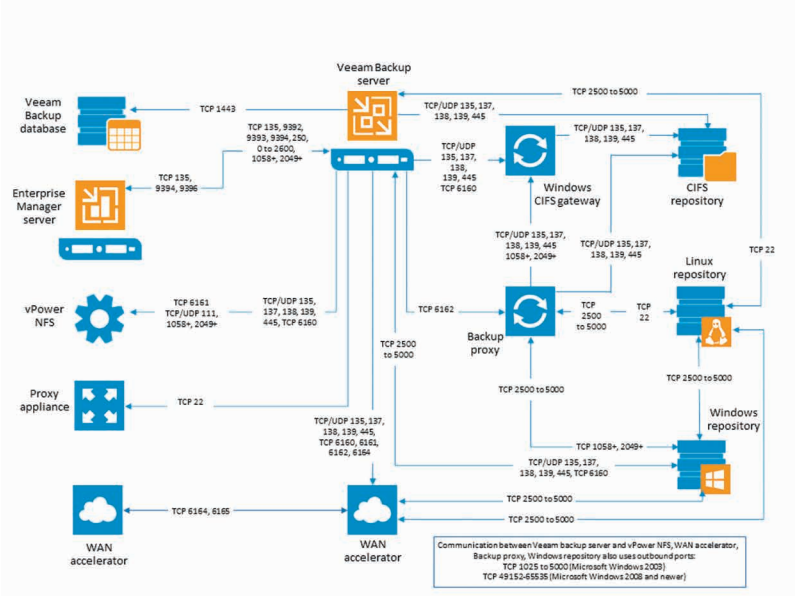
The following diagram illustrates a possible placement of the Veeam Backup & Replication components in a distributed environment, with a WAN link between the production and DR sites.



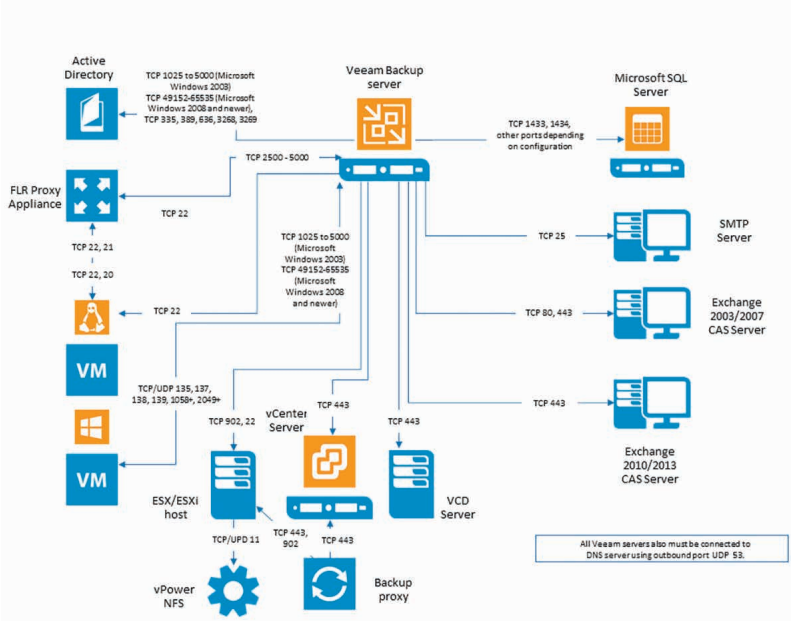
# APPENDIX 2. NETWORK CONNECTIVITY DIAGRAMS

There is a detailed list of ports used by Veeam Backup & Replication available at http://www.veeam.com/kb1518, but sometimes a more visual approach is helpful – you can use the diagrams below for that purpose.

## Veeam Backup Server



# Infrastructure Servers



# Veeam Backup Enterprise Manager