Best Practices

Veeam Backup & Replication

Version 9

August, 2016

VEEAM IT JUST WORKS!™

Microsoft Partner
Gold Application Development
Gold Datacenter

vmware PARTNER TECHNOLOGY ALLIANCE

Table of Contents

Veeam Backup & Replication Best Practices for VMware

Version 9 Update 2

Build 9.0.0.1715.

Introduction

Welcome to the Best Practices guide for Veeam Backup & Replication.

About This Guide

This guide is developed by Veeam architects, and its content is also validated by support, developers and QA departments to ensure highest possible quality. If you have any questions or comments, please reach out the authors directly, or via your local Veeam Software representative.

If you possess a printed version of this e-book, you will notice many references to external resources for additional information. The e-book is optimized for digital consumption, and the most recent copy is always available at:

bp.veeam.expert

Intended Audience

This guide is intended for backup administrators or consultants managing Veeam Backup & Replication on a daily basis.

Most sections of this guide assume you already have hands on experience with Backup & Replication, and will serve as an "advanced user guide", meaning that more basic usage information, system requirements and the like must be found in User Guide in Veeam Helpcenter.

Service providers delivering BaaS and DRaaS with Veeam Cloud Connect should refer to the corresponding Veeam Cloud Connect Reference Architecture.

Authors

- Preben Berg (@poulpreben)
- Andreas Neufert (@AndyandtheVMs)
- Tom Sightler
- Pascal di Marco
- Stanislav Simakov (@ssimakov)

- Paul Szelesi (@PSzelesi)
- Luca Dell'Oca (@dellock6)

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.

Online Support

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

- Veeam Helpcenter at helpcenter.veeam.com
- Veeam Community Forums at forums.veeam.com

Customer Support

Should you have any technical concerns, suggestions or questions, please visit the Veeam Customer Portal at cp.veeam.com to open a case, search our knowledge base, reference documentation, manage your licenses or obtain the latest product release.

Company Contacts

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

Resource Planning

DNS Resolution

Domain Name System (DNS) resolution is critical for Veeam Backup & Replication deployment (VBR) and configuration. VMware components must be accessible via both forward (A) and reverse (PTR) lookups. If not some Veeam components may not work correctly.

If DNS resolution is not available you may add VMware vCenter, ESXi and managed Veeam servers to the local `hosts` file on *all* managed Veeam servers. When using this workaround it is recommended to add both short name and fully qualified domain name in the `hosts` file. DNS should be the preferred option.

When ESXi hosts are added to vCenter it is recommended to use FQDN. When using Network Backup mode the FQDN is returned via VMware API for Data Protection (VADP) so the backup proxy server must be able to resolve the FQDN via DNS. Using the `hosts` file the data transport path can be altered for Network Backup mode transfers.

Please see the below example.

Example `hosts` file

```
10.0.4.10    vcenter    vcenter.example.com

# 10.0.4.21    esx1    esx1.example.com # commented out management interface
# 10.0.4.22    esx2    esx2.example.com # commented out management interface

10.255.4.21    esx1    esx1.example.com # dedicated 10 GbE backup network
10.255.4.22    esx2    esx2.example.com # dedicated 10 GbE backup network
```

To explicitly alter the data transport path, the `hosts` file must be deployed on all backup proxy servers. For easier management, please see the Carbon module and `Set-HostsEntry` by Aaron Jensen.

Backup Server

Before installing the Veeam Backup & Replication server it is important to understand the different data streams generated by the Veeam Backup Server (VBR) Service's. Additionally there are considerations in regards to system requirements that may affect the desired deployment method.

Features & component requirements will affect your decision as to how you install the backup server ie: physical datacenter or different ESXi cluster. It could also mean that you choose to install additional backup servers or services in remote locations to optimize restore performance.

Deployment Method

You may deploy the Veeam Backup & Replication server as either a physical or virtual server. It will run on any server with Windows Server 2008 R2 or higher installed. Depending on the environment here are some guidelines that may help in deciding which deployment type is the best fit.

Virtual deployment

If installed in a virtual machine the VM can be replicated to a secondary location such as a DR site. If the virtual machine itself should fail or in the event of a datacenter/infrastructure failure, the replicated VM can be powered on via the VMware vSphere Client. Best practice in a two site environment is to install the Backup server in the DR site, in the event of a disaster it is already available to start recovery processes.

For most cases virtual is the recommended deployment as it provides high availability for the backup server component via vSphere High Availability or Fault Tolerance. Additionally it provides great flexibility in sizing and scaling as the environment grows.

Physical deployment

In small-medium environments (up to 500 VMs) it is common to see an all-in-one physical server running the Backup & Replication server, backup proxy and backup repository components. This is also referred to as "Appliance Model" deployment.

In large environments (over 2,500 VMs) installing Backup & Replication services on separate servers either virtual or physical may provide better performance. When running a large number of jobs simultaneously, consuming large amounts of CPU and RAM, scaling up the virtual Backup & Replication server to satisfy the system requirements may become impractical.

If installed on a physical machine, the Veeam backup server runs independently from the virtual platform. This may also be an ideal solution in case of disasters in the virtual environment.

Should the physical server itself fail, there are additional steps to take before reestablishing operations:

1. Install and update the operating system on a new server

2. Install Veeam Backup & Replication
3. Restore the configuration backup

In an enterprise environment you may choose to install an additional backup server to speed up the recovery process during a disaster. You may re-use existing infrastructure such as a proxy or repository server for the standby Backup & Replication server. During a disaster the configuration backup can easily be restored to this server. It is recommended to store the configuration backup using a file copy job in a location that is always available to this standby Backup & Replication server.

Backup Server Placement

The Backup server runs a number of processes, e.g. the Backup Service, Backup Manager services and in some scenarios a Mount Server as well. In this chapter we will evaluate how each of those components are affected by placement of the Backup & Replication server.

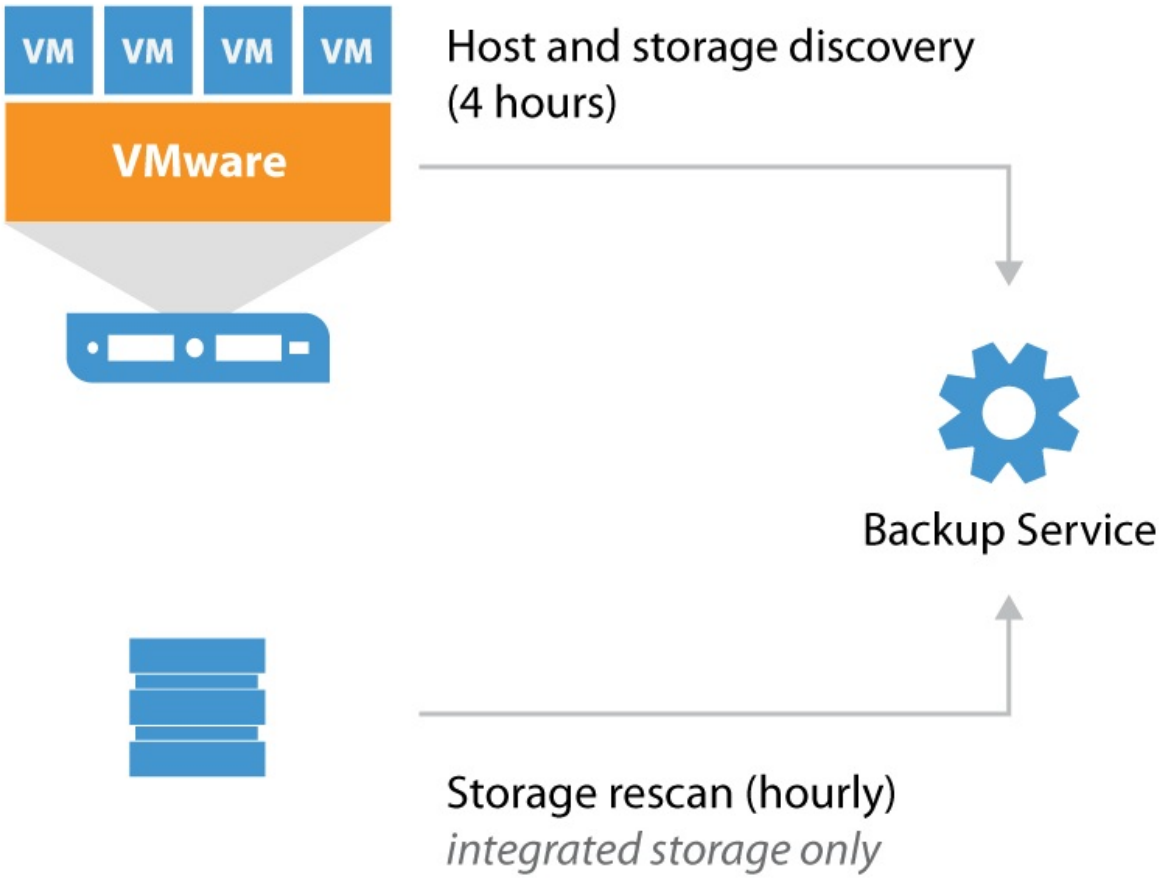
By evaluating the roles and understanding the data flow between the services it is possible to optimize overall backup performance and restore throughput significantly.

Host and Storage Discovery

To collect information about the virtual infrastructure all managed vCenters and their connected hosts and datastores are periodically rescanned. This rescan process is visible in the **History** tab > **System** section in the Veeam Backup & Replication console. As seen here, the Host discovery process runs every four hours. All the collected information is stored within the configuration database.

The amount of collected information is typically very small however the Host discovery process may take longer or even exceed the default schedule in highly distributed environments[1]. If hosts or clusters are connected to vCenter over a high-latency link you may consider deploying a Backup server locally on the ROBO, then you can create a vCenter service account with a limited scope to that particular location in order to reduce the window of the Host discovery process. If the ROBO uses a stand-alone host it is possible to add the host as a managed server directly instead of through vCenter.

Note: Avoid adding individual hosts to the backup infrastructure if using shared storage in a vSphere cluster.



If storage with advanced integration (HPE, NetApp, EMC) are added to the **Storage Integration** tab there will additionally be a Storage discovery process periodically rescanning storage hourly. This process checks all snapshots for virtual machine restore points for usage within Veeam Explorer for Storage Snapshots. The Veeam Backup & Replication server itself will not perform the actual scanning of volumes but it will use the management API's of the storage controller to read information about present snapshots. Only proxy servers with required storage paths available will be used for the actual storage rescanning process[2].

The following table shows the three different scanning workflows:

Adding new storage controller	Creating new snapshot	Automatic scanning
1. Collect specific storage information	1. Creating new Snapshot	1. Storage Monitor runs in background
2. List of volumes, snapshots, LUNs and NFS exports	2. Lists initiators	2. Detecting new volumes
3. Checking licenses, FC and iSCSI server	3. Testing iSCSI, NFS and FC from proxies	3. Scanning volumes for snapshots every 10 minutes
4. Lists initiators	4. Searching storage exports in VMware	4. Lists initiators
5. Searching storage exports in VMware	5. Mapping discovered VMs from datastores to snapshots	5. Testing iSCSI, NFS and FC from proxies
6. Mapping discovered VMs from datastores to snapshots	6. Export and scan the snapshots with proxies	6. Searching storage exports in VMware
7. Export and scan the snapshots with proxies	7. Update configuration database	7. Mapping discovered VMs from datastores to snapshots
8. Update configuration database		8. Export and scan the discovered objects with proxies
		9. Update configuration database

The scan of a storage controller performs, depending on the protocol, several tasks on the storage operating system. Therefore it is recommended to have some performance headroom on the controller. If your controller is already running on >90% CPU utilization, keep in mind that the scan might take significant time to complete.

The scanning interval of 10 minutes and 7 days can be changed with the following registry keys.

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`
- Key: `SanMonitorTimeout`
- Type: REG_DWORD
- Default value: 600
- Defines in seconds how frequent we should monitor SAN infrastructure and run incremental rescan in case of new new instances

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`

- Key: `SanRescan_Periodically_Days`
- Type: REG_DWORD
- Default value: 7
- Defines in days how frequent we should initiate periodic full rescan after Veeam Backup service rescan

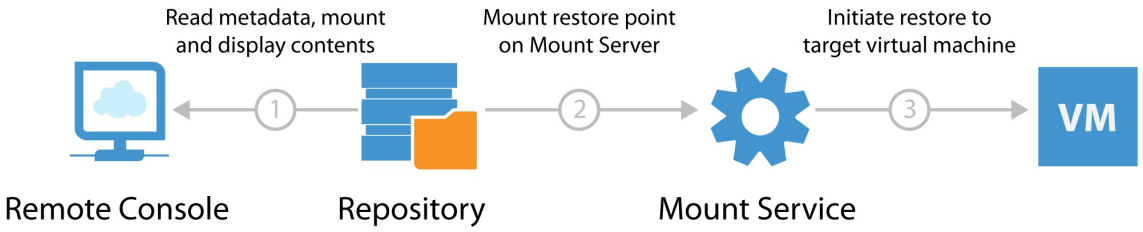
Per default Veeam will scan all volumes and LUNs on the storage subsystem. During rescan, each present snapshot produces a snapshot clone, mounts to a proxy server, scans the filesystem, lookup for discovered VMs and unmounts. This is repeated for every present snapshot.

Example: A storage system with 50 volumes or LUNs with 10 snapshots for each. Scanning the entire system means 500 (50x10) mounts and clones are performed. Depending on the performance of the storage system and the proxy server, this can take significant time.

To minimize the scan time it is recommended to select the volumes used by VMware within the setup wizard to avoid the overhead of scanning unused data volumes.

File-level Recovery Data Flow

To perform file-level restores for a Windows-based or other OS VM Veeam mounts all VM disk files from the backup files (stored on the repository server) to a Mount Service.



When file-level recovery is performed from the Veeam backup console, two mounts are initiated:

1. The remote console - for displaying restore point contents
2. The mount server - for performing actual restore traffic to the target VM

Note: For VMs not running a Windows operating system, a Linux based FLR helper appliance mounts the backup file for reading the file system.

Between 50-400 MB of data is transferred between the console and backup repository. If the first file mount is performed over a slow connection it may take considerable time to load the file-level recovery wizard. If there is significant latency between the backup repository and console, it is recommended to deploy an instance of the console on or closer to the repository server.

Veeam Enterprise Manager

Veeam Enterprise Manager is a self-service portal where administrators or service desk representatives can initiate restores for VMs, files, e-mail items and SQL databases.

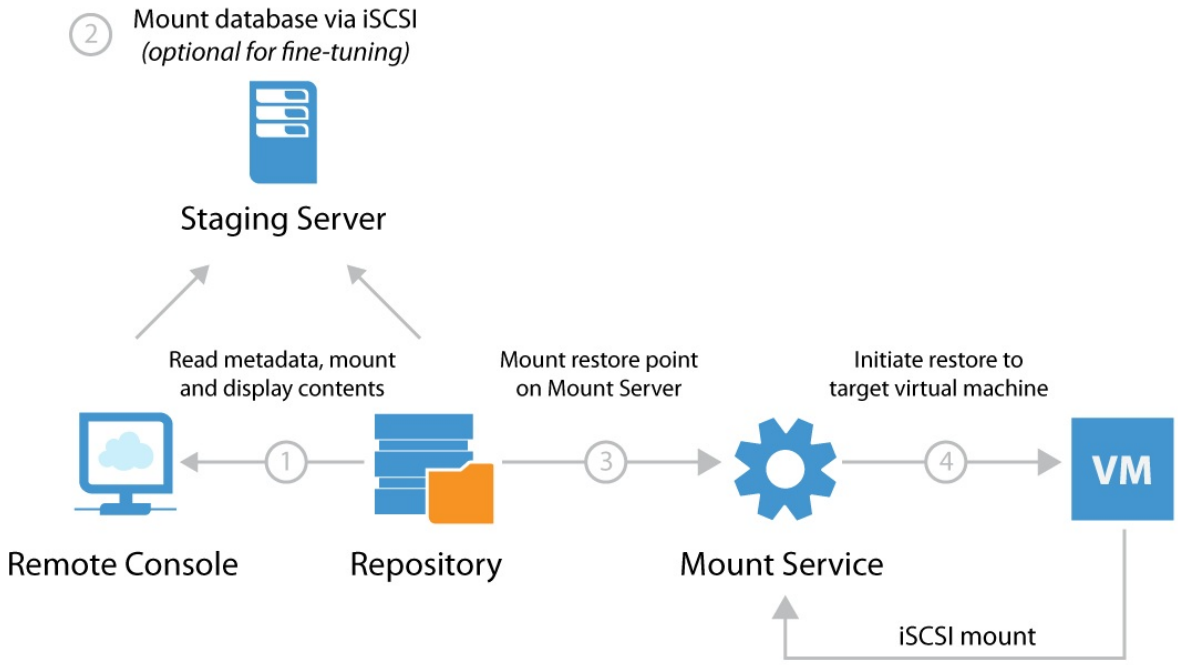
It is possible to avoid the first mount entirely by using "guest file system indexing"[3]. When guest file system indexing is enabled, the content of the guest VM is stored in the Veeam Catalog and presented through Veeam Enterprise Manager. Veeam Enterprise Manager will initiate the file-level restore with the mount server without requiring the first mount.

Note: If guest file system indexing is disabled restores may still be initiated through Enterprise Manager however they will still require the first mount to be performed with similar performance implications as previously described.

Veeam Explorers

Veeam Explorers are installed as part of the backup server and backup console when installed remotely. When performing item-level recoveries the file-level recovery engine is leveraged. Please see the previous section for deployment considerations.

The Veeam Explorer for SQL Server, SharePoint and Oracle all use a staging server to allow selecting a specific point in time for point-in-time restore. This introduces an additional connection as illustrated below.



Disaster Recovery Optimization

When using Veeam for replicating VMs to a disaster recovery (DR) site, it is recommended to keep the Backup & Replication server in the DR site alongside the replicas. When the backup server is located in the DR site it enables true "1-Click Failover" by being able to start Failover Plans immediately and thus eliminate manual reconfiguration before the failover process can be initiated.

Proper planning dictates that to get 1-Click Failover working it requires that the vSphere clusters in each location are connected to separate vCenter servers. In the event of an outage in the primary datacenter it is only possible for the Backup & Replication server in the DR site to initiate failover if the vCenter server itself is available.

In cases when it is impossible to have multiple vCenter instances across sites (e.g. Metro Cluster or similar active-active configurations), the recommended solution is to use vCenter Server and following these steps in event of a disaster:

1. Replicate vCenter from primary site to secondary site with low RPO
2. Configure VMware DRS affinity rules[4] for pinning replica vCenter VM to a specific host
3. Connect to specified host and manually power on replicated vCenter VM
4. Verify vCenter availability through Veeam Backup & Replication
5. Initiate Failover Plans

Examples

In this section we will outline two examples based on two enterprises with 50 remote/branch offices (ROBO). They have the following common characteristics:

- One vCenter Server in HQ managing all ROBO sites
- Local backup jobs for fast backup and restore performance
- Offsite copies from the ROBO consolidated at HQ for D/R protection

Example 1: Centralized Job Configuration

IT requires *one* central management console for the entire backup infrastructure, administration and job scheduling. The backup administrator can follow these guidelines:

1. Install and configure Veeam Backup & Replication in HQ
2. Add the vCenter Server via the Veeam Backup & Replication console
3. Add the ROBO backup server as Managed Server in the **Backup Infrastructure** tab
4. Configure the HQ backup server with the roles Backup Repository and optionally WAN accelerator
5. Configure the ROBO backup server with the roles Backup Proxy, Backup Repository and optionally as WAN accelerator[5]
6. Configure one or more Backup Jobs for each ROBO pointing to its local backup repository
7. At HQ configure one or more Backup Copy Jobs for each ROBO pointing to the backup repository
8. Install Veeam Backup Console on the ROBO backup server for faster restore via the local Mount Server

Note: The remote console installation files are on the same installation media as Veeam Backup & Replication (`\Backup\Shell.x64.msi`)

Constraints

Please consider the following constraint:

- If a WAN link between HQ and a ROBOs fails, no backup jobs will run, as the backup server will not be able to communicate with the remote ESXi hosts via the centralized vCenter Server

- When performing file-level restore for non-indexed virtual machines at the ROBO via Veeam Enterprise Manager the restore point will be mounted over the WAN link to HQ for displaying the contents of the restore point. Thus it is recommended to use indexing for such virtual machines

Example 2: Distributed Job Configuration

IT requires local backup jobs and backup copy jobs (with optional WAN acceleration) are created at the ROBO. For security considerations, each ROBO is provided with delegated access to VMware vCenter. Restore capabilities from backup copy jobs should be configured and managed at HQ as well as delegated restore and license management for all sites via Veeam Enterprise Manager. The backup administrator may follow these guidelines:

1. Install Enterprise Manager at HQ
2. Install and configure Veeam Backup & Replication on each ROBO
3. On vCenter Server, create separate service accounts per ROBO with a limited scope for displaying only relevant hosts or clusters
4. At the ROBO, add vCenter Server via the **Backup Infrastructure** tab using the scoped service account
5. *Optional:* At the ROBO, configure a local WAN accelerator and create or re-use an existing WAN accelerator at HQ (please note many-to-one configurations are supported)
6. At the ROBO, add and configure the Repository Server at HQ (please note many-to-one configurations are supported)
7. Configure one or more Backup Jobs at each ROBO pointing to its local backup repository
8. Configure one or more Backup Copy Jobs at each ROBO pointing to the centralized backup repository at HQ (use WAN acceleration as needed)
9. Install Veeam Backup & Replication Console at HQ. When using the remote console for connecting to remote instances, it is possible to leverage faster file-level or item-level restores at HQ via the console's built-in Mount Server

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.

> [1]. In very large or extremely distributed environments, it is possible to extend the schedule frequency by altering registry key `VolumesDiscover_Periodically_Hours` (REG_DWORD, default: 4) ↩
> [2]

> [2]. Storage rescan procedure > Re-Scanning Storage Systems ↩

> [3]. More information about guest file system indexing in Veeam Helpcenter > Guest file system indexing ↩

> [4]. VMware Distributed Resource Scheduler > VM-Host Affinity Rules ↩

> [5]. Remember to add sufficient resources if all three roles can run on the remote backup server. ↩

Sizing and System Requirements

In this section, we will describe how to configure and size the Veeam backup server.

Sizing with Veeam is cumulative in respect to configurations, if you want to create an all-in-one appliance (Appliance Model) add all the resource requirements together (CPU + Memory) to understand what in total you will need, the same goes if you only wish to have proxy and repository in one host.

Compute requirements

Recommended Veeam backup server configuration is **1 CPU core (physical or virtual) and 4 GB RAM per 10 concurrently running jobs**. Concurrent jobs include any running backup or replication jobs as well as any job with a continuous schedule such as backup copy jobs and tape jobs.

The minimum recommendation is 2 CPU cores and 8 GB RAM.

It is recommended to group multiple virtual machines into a single job for better efficiency and resource usage. With default configuration it is recommended to configure around 30 VMs per job. The recommendation can be increased by over 10x (300+ VMs) by leveraging additional features such as per VM backup files. Please refer to the Job Configuration section of this guide to learn more about job design.

All configuration and session information is stored in the configuration database. In larger environments the load on the SQL Server hosting the configuration database may be significant and is highly dependent on the amount of concurrently running jobs. For more information please see the Backup Server Database section of this guide.

Operating system

The Veeam backup server requires Microsoft Windows 2008 R2 or later. The latest supported version of Windows OS is always recommended (currently Microsoft Windows 2012 R2) as it will also support restoring from virtual machines with ReFS file systems or Windows Server Deduplication enabled.

For the full list of supported operating systems, please refer to the corresponding System Requirements section of the Veeam User Guide.

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:

Installation folder

Default location is `C:\Program Files\Veeam\Backup and Replication`

Plan for 40 GB. If installing in a virtual machine, thin disks may be used. By default the installer will choose the biggest drive space for the built in backup repository.

Log files

Default location is `C:\ProgramData\Veeam\Backup`

Log file growth will depend on the number and frequency of jobs and the VM count. 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.

It is recommended to not configure the logging level below 4, as it may complicate troubleshooting. Logging level 6 is very intrusive, and should only be configured for short periods of time when requested by Veeam Support.

Plan for 3 GB log files generated per 100 virtual machines, with a 24 hour RPO. For environments with more than 500 VMs it is recommended to change the default location to a different fast access disk. Many concurrently running jobs may produce a lot of write streams to log files, than can slow down operations for the Veeam Backup Service and Backup Manager processes.

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 Search Server and Indexing section of this guide. To change the default location, refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1453

vPower NFS folder

Default location is `C:\ProgramData\Veeam\Backup\NfsDatastore`

When booting VMs with Instant VM Recovery or SureBackup, this folder is used by default to store all configuration files and redo logs of the running VM. To offload the changes to a specific production datastore refer to the corresponding page of the Instant VM Recovery wizard.

We recommend installing vPower NFS Services on each Windows-based backup repository. For SMB/CIFS based repositories or deduplication appliances it is recommended to configure vPower NFS on the gateway server. For Linux-based repositories it is recommended to configure vPower NFS on a managed Windows machine as close as possible to the Linux repository (similar to selecting a Gateway Server for SMB/CIFS or deduplication storages).

The vPower NFS server is bound to backup repositories and the folder location is defined per server. To achieve best performance for VMs running off of vPower NFS please configure the fastest possible storage on the backup server or backup repository. To change the folder location please see the following steps.

1. in the **Backup Infrastructure**, select the **repository** you wish to change.
2. Right click the **repository** and go to **properties**
3. When the wizard opens navigate to the **Mount server** settings
4. using the browser buttons locate the new location for your vPower NFS storage
5. finish the wizard

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 longer period increase the space accordingly to fit the produced/estimated amount of changes generated by the running VMs (conservative average change rate can be defined as 100 GB per 1 TB VM per 24 hours - or 10%). Additional disk space is consumed when using Quick Migration. See more information here > Veeam Help Center > Performing Instant VM Recovery > Before You Begin.

Important! Make sure vPower NFS is configured correctly on the Veeam backup server itself as it will be used when deploying Virtual Lab for SureBackup or when performing file-level recovery for Linux-based VMs.

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.

Other software

It is strongly recommended that no highly-transactional and business-critical software is deployed on the same machine as the Veeam backup server. This could be (but not limited to) software such as Active Directory, Exchange Server or other intensive production databases on the SQL server instance.

It is recommended to follow antivirus exclusion guidelines as explained in Veeam KB 1999.

If it is not possible to connect to a remote SQL staging server for Veeam Explorers you can install Standard or Enterprise versions of SQL (depending on your licensing) locally for staging databases for item-level restores on the backup server. This installation can also be used to store the Veeam backup database if required as long as sufficient resources are assigned to the host machine, however do not run any instances in production from this installation that may affect the operation of the backups or restore processes. SQL express is included in the distribution but is limited to a 10GB database.

Note: Remote SQL Server for staging is supported from v9.0

Other software such as Microsoft Outlook (64-bit) for mail export to PST files via Veeam Explorer for Exchange, or a PDF viewer for reading Veeam documentation are considered non-disruptive.

Installing Veeam Backup & Replication updates

New Veeam releases and updates are installed on the Veeam Enterprise Manager and Veeam backup servers by the setup wizard or by using the unattended installation method (also referred to as “silent installation”). For detailed instructions check the latest release notes.

Note: Veeam Backup Enterprise Manager must be updated before updating Veeam backup servers.

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

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

Veeam Backup & Replication Database

Veeam Availability Suite which includes Veeam Backup & Replication and Enterprise Manager stores all information about backup infrastructure, jobs settings, job history, sessions and other configuration data in an instance of SQL Server.

When planning the Veeam Backup & Replication deployment you must choose the placement of the configuration database. It may be either a local or remote SQL Server with several licensing options available. Please see the following recommendations to ensure your Backup & Replication setup will scale to the size of your infrastructure.

SQL Server Edition

Microsoft SQL Server 2012 Express Edition is included in the Veeam Backup & Replication setup which is a convenient option for most smaller deployments. It does however have several limitations[1] which may affect performance:

- Each instance uses only up to 1 GB of RAM
- Each instance uses only up to 4 cores of the first CPU
- Database size cannot exceed 10 GB

It is recommended to install Standard or Enterprise Edition if any of the following apply:

- **When protecting more than 500 VMs**. It is recommended to use Standard or Enterprise versions of Microsoft SQL Server. The max database size is usually sufficient. Veeam Backup & Replication console and job processing may slow down as a result of CPU and RAM constraints on the SQL Server instance.
- **When using Files to Tape jobs extensively**, the database may grow significantly, and the 10 GB limitation may be exceeded quickly.
- **When unable to configure an external staging server**. For Veeam Explorer for Microsoft SQL Server or Veeam Explorer for Microsoft SharePoint. When working with databases larger than 10 GB, SQL Server Express cannot mount the databases.
- **When databases are using advanced features of Microsoft SQL Server**. Such as encryption or table partitioning, the licensing level of the staging server (local or remote) must match the level of the original instance.

If none of the above apply it is recommended to use Microsoft SQL Server Express Edition for the sake of simplicity.

Tip: Veeam Backup & Replication leverages Microsoft SQL Server 2014 enhancements (cardinality estimator has proved to show significant improvements for large queries), it is recommended where possible to update the database server to Microsoft SQL Server 2014 or Microsoft SQL Server Express 2014.

Database Placement

It is possible to leverage a remote SQL Server as staging server during restores in Veeam Explorer products. There are no specific edition requirements for neither SQL Express, Standard or Enterprise instance of SQL Server installed locally on the backup server. It is still recommended to run the SQL Server locally (when resource and planning allow) on the backup server for lowest latency and highest performance.

There may still be scenarios where a remote SQL Server is the better choice:

- SQL Clustering and AlwaysOn Availability Group on external SQL Servers can be used for configuration database high availability
- Failover to a standby backup server can be simplified by connecting to the configuration database directly without the need for restoring from a configuration backup
- Some enterprises have dedicated virtual clusters for SQL Server due to licensing constraints. In such cases, you may place the Veeam configuration database on existing instances to lower the overall TCO

Sizing

Veeam Backup & Replication may consume high amounts of CPU and RAM while processing backup or replication jobs. To achieve better performance and load balancing it is necessary to provide sufficient RAM and CPU resources to Veeam components. Remember to add additional resources, if the backup server is responsible for multiple roles, such as repository or backup proxy.

If possible follow these guidelines:

Number of concurrently running jobs	CPU	RAM
Up to 25	2	4 GB
Up to 50	4	8 GB
Up to 100	8	16 GB

Note: Concurrently running jobs include any job type with a continuous schedule such as Backup Copy Jobs.

When running more than 100 jobs concurrently increase compute resources in line with the table above to meet the resource need of the workload.

It is recommended to place the configuration database on fast, resilient storage subsystem. Performant storage for backing the configuration database will result in overall increased processing performance. Jobs with a lot of metadata such as very large SharePoint farms with thousands of sites, SQL Server instances with many databases or Files to Tape jobs may increase the I/O requirements for the configuration database.

SQL Server Configuration Tips

Veeam Backup & Replication does not require any specific settings[2] on the SQL Server in order to utilize the capabilities of Veeam Explorer for SharePoint or SQL. Both local and remote SQL Servers can be used for staging purposes, the corresponding requirements are detailed on Veeam Helpcenter should be met:

- Veeam Explorer for Microsoft SharePoint
- Veeam Explorer for Microsoft SQL Server

Tip:

- Enable and configure all features used by production databases.
- When possible use the highest license level and latest version and cumulative update level installed in any VM.
- Using an older version of SQL Server for the configuration database than running in a protected VM may result in warnings in job session logs when such VMs are processed.

If you plan to restore encrypted databases with Veeam Explorer for Microsoft SQL Server or SharePoint you will need a valid encryption certificate on the staging Microsoft SQL Server[3].

Follow Microsoft general recommendations for optimal SQL performance, for example, place the SQL `tempdb` on the fastest disks for best performance[5].

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 https://helpcenter.veeam.com/backup/vsphere/dbconfig_utility.html.

If using SQL authentication consider that all Veeam UI and Veeam PowerShell changes are communicated using this authentication.

Migrating Veeam Database

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

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

> [1]. Features Supported by the Editions of SQL Server 2012 https://msdn.microsoft.com/en-us/library/cc645993(v=SQL.110).aspx#CrossBoxScale ↩

> [2]. Generic requirements for SQL Server can be found here: https://helpcenter.veeam.com/backup/vsphere/system_requirements.html#backup_server ↩

> [3]. For restoring encrypted databases, please see: http://www.veeam.com/kb2006 ↩

> [5]. SQL Server `tempdb` Best Practices: http://blogs.msdn.com/b/cindygross/archive/2009/11/20/compilation-of-sql-server-tempdb-io-best-practices.aspx ↩

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/free/vbr_config.html

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

Planning for Disaster Recovery of Veeam Backup Server

If you have the backup server in the primary site it is recommended to replicate the Veeam backup server VM to the secondary site (verify network and IP mapping settings before you begin; refer to http://helpcenter.veeam.com/backup/vsphere/index.html?replica_job.html for details). If the configuration database is external, ensure this server is also replicated to the secondary site.

If the server is replicated successfully in the event of a disaster 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 Veeams File Copy Job to place a copy of the configuration backup at the DR site.

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.

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.

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
- *c:\programdata\veeam*

Folder on VM Guest OS (if VSS is used)

C:\Windows\VeeamVssSupport

Files

- *VeeamAgent.exe*

- *VeeamAgent64.exe*
- .vmdk .vbk .vlb .vib .vrb .vbm

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.

Consider that other services 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 please refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1055.

Veeam Backup Enterprise Manager

Whether to Deploy?

Enterprise Manager is intended for centralized reporting and management of multiple backup servers. It provides delegated restore and self-service capabilities as well as the ability for users to request Virtual Labs from backup administrators. It provides a central management point for multiple backup servers from a single interface. Enterprise Manager is also a part of the data encryption and decryption processes implemented in the Veeam solution and best practice 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 (http://helpcenter.veeam.com/backup/em/index.html?em_manage_keys.html) for the connected 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 Public Key Infrastructure (PKI).
- If an organization has a Remote Office/Branch Office (ROBO) deployment then leverage Enterprise Manager to provide site administrators with granular restore access via web UI (rather than providing access to 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 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.

If the environment includes a single instance of 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 files with a single click. To support this capability the VM restore point must be created with application-aware image processing enabled. Additionally, if guest file system indexing is enabled, it is possible to search for files across VM backups.

Note: It is possible to restore VM guest files even when application-aware image processing or file indexing is disabled. If both are disabled, the restore operator must type in guest OS credentials during a file-level restore.

The backup catalog on the Enterprise Manager server will be used to store indexing data replicated from the 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 "Search Server and Indexing" section of this document.

1-Click Application Item-level Restore

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 line 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 the 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 Backup & Replication 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. To do this an administrator of the VM can access the self-service web portal using the default URL: "https://ENTERPRISE_MANAGER:9443/selfrestore".

Tip: This feature is available only for the Windows-based VMs and requires Veeam Backup & Replication Enterprise *Plus* license. The VM needs to 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 guest processing 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

The 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 Indexing

Indexing and Search Overview

Veeam Backup & Replication performs backups at the image-level using APIs available from the underlying hypervisor. It has no direct visibility of the file structure after backup is finished. It is possible to Use File Level Recovery (FLR) wizard or Enterprise Manager 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 an 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 offers the capability to index files on VMs being backed up. Indexing is available for both Windows & Linux VMs allowing users of 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.

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 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 possible to exclude specific VMs from indexing as described for example in this section of the Veeam Backup Enterprise Manager User Guide

How Veeam Indexing Works

Veeam indexing creates a separate index file in the catalog for each restore point. These index files are used by Veeam Enterprise Manager to support file browsing or searching without a need to mount the restore point to the mount server. Users can quickly search for files across multiple restore points viewing the required file 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.

Enterprise Manager 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 If configured:

1. 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. Veeam uses the following software packages for it: mlocate, gzip and tar

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 the index is extremely fast and has minimal impact on network and VMware environment.
2. Once the index is created and stored on Veeam backup servers, the indexing service on Veeam Backup Enterprise Manager performs index copy — it aggregates index data for all VM image backups from multiple backup servers to the Enterprise Manager database while the original index files in the backup servers are deleted to conserve space. The consolidated indexes are stored on the Enterprise Manager server in the `C:\VBRCatalog\Index\Catalog` and are used for search queries .

Important To Note!

- To search within the index catalog it is necessary to deploy Veeam Backup Enterprise Manager, 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 indexes in the *VBRCatalog* folder of the backup server will never be collected or deleted and will eventually fill up the disk drive.

Temporary VM Disk Usage

During the indexing process indexing information is temporarily stored on the local VM guest requiring additional free space on the system drive.

Windows VM

Temporary space required on the first drive in the VM (`C:\` drive):

100 MB per one million files

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

Linux VM

Temporary space required in `/tmp`:

50 MB per one million files

Linux indexes require around 50% less space because `mlocate` does not index metadata such as timestamps and ownership.

Sizing Enterprise Manager Catalog

The Veeam Catalog Service is responsible for maintaining index data. When running on the 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 move index data from all managed backup servers into the Enterprise Manager local catalog deleting the files from the originating backup server 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.

Estimated used space of the final index file after compression is approximately 2 MB per 1,000,000 files for a single VM restore point on the Enterprise Manager server. The indexes are also stored in the backup files and temporary folders on the backup server.

Example

Below is an example that summarizes the information above. The example is given *per* indexed VM containing 10,000,000 files.

```
2 MB * 10 million files * 60 restore points per month * 3 months index retention = 3.5 GB
```

Recommended Settings

Follow these recommendations when setting up Veeam indexing:

- Place the catalog on a dedicated volume of high performance disk. To change the default Veeam Catalog folder location refer to this Veeam Knowledge Base article: http://www.veeam.com/kb1453.
- You can enable NTFS compression on the catalog folder. This can reduce the space requirements by well over 50%. 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.
- 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**.
- 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.

  Notes:

  To take advantage of indexing on SUSE Linux Enterprise Server (SLES) you must be running version 12 or above. In lower versions that do not contain by default the mlocate package you may try this OpenSUSE package http://software.opensuse.org/package/mlocate

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

Using Veeam Backup Search (Optional Component)

In its early versions Veeam did not have its own indexing engine, instead it used the Veeam Backup Search component to connect to the Microsoft Search Server 2010 that provided search capabilities. Now Veeam has its own built in indexing engine developed specifically for this purpose.

It is no longer a requirement to have a Veeam Backup Search configured as Veeam Integrated indexing engine can be more performant.

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

- 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. If you are using 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. 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

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 advanced deployment scenarios, the backup proxy role is usually assigned to between one or several Windows servers (recommend using 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 infrastructure. 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 Backup & Replication during backup, replication (at source and target), VM copy, VM migration jobs or VM restore. They are also used to detect and scan snapshots to enable Veeam Explorer for Storage Snapshot features if a compatible storage system was added to the Backup & Replication Server.

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).
- BitLooker: Applies to VMs running Windows OS and using NTFS. For more information, see the corresponding section of this guide > Deduplication and Compression - BitLooker
- Inline source side data deduplication to optimize information received by vSphere Change Block Tracking (CBT)
- Inline compression
- AES256 encryption if the corresponding option is selected in the data transportation or backup data 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 assigning the proxy role to it Backup & Replication installs the necessary components

starting the required services on that server. When a job is started the Veeam Backup & Replication server becomes the point of control for dispatching tasks to proxy servers using its built-in load balancing algorithm.

Like any backup vendor using VMware VADP, Backup & Replication integrates VMware VDDK in the Veeam Transport Service. This is necessary for management interaction with vCenter and ESXi hosts, while in some scenarios, VDDK is bypassed to optimize performance.

Backup data VDDK based transport modes underlay some limitations so Veeam developed it´s own more advanced communication protocols to address them. For example Veeam can backup multiple disks of the same VM at same time with HotAdd mode or can read data with Direct NFS mode directly out of NFS based storage systems.

Intelligent Load Balancing

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), 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.

The default proxy server is configured for 2 simultaneous tasks at installation, whereas subsequently added proxy servers analyze the CPU configuration. The proxy server automatically proposes configuring 1 task per CPU core. During deployment, it is determined which datastores the proxy can access. This information is stored in the configuration database, and is used at backup time to automatically select the best transport mode depending on the type of connection between the backup proxy and datastore.

First Backup & Replication checks if data processing can be assigned to a backup proxy with the Direct Storage mode (which includes Direct SAN Access and Veeam's own special Direct NFS Access), then it checks whether a Virtual Appliance or Hot-Add proxy can be used. Then it looks for a Network Mode(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:

1. 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.
2. All tasks are standing in a “VM to process” queue, when a proxy’s task slot becomes free Backup & Replication will automatically fill it up with the next VM disk backup task.

3. Priority goes to the disk that belongs to an already processed VM, after that VMs of already running jobs have next higher priority. Short-term scheduled jobs take priority over long-term scheduled jobs (like daily or weekly jobs).

Tip: At the repository which writes the backup data, only one write stream is started to the backup storage per job. Job priority together with jobs to be run with a large amount of VMs can lead to the situation that only 1-3 target write streams can become the bootleneck in the backup processing. Consider enabling "Per VM backup chain" at the repository to address this.

Tip: Default recommended value is **1** task per Core/vCPU, with no less than 2 cores as minimum. To optimize the backup window, you can cautiously oversubscribe the **Max concurrent tasks** count, but monitor CPU and RAM usage carefully.

Veeam Backup & Replication supports parallel processing of VMs/VM disks:

- It can process multiple VMs within a Job simultaneously increasing data processing efficiency.
- If a VM was created with multiple disks Veeam will try to process these disks simultaneously to reduce VM backup time to minimize VMware snapshot lifetime.
- Priority goes to already running parallel processes for VM disks backups. After all these disks are backed up, VMs from existing running jobs take priority. As well as highly frequent scheduled jobs which are prioritized higher.

To achieve the best backup window it is recommended to slightly oversubscribe the tasks slots and start more jobs. This allow Veeam to leverage the maximum of the task slots and lead into an optimal backup window.

Note: Parallel processing is a global setting that is turned on by default. If you had upgraded from older versions please check and enable this setting.

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". It performs 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 on the proxy. These processes can operate in either read or write mode. When used on a proxy server for backup, they are only performing read operations, while "write" mode is used for writing data on a target backup proxy (replication). Veeam agents in write mode are also used on all repository types, but will not be discussed in this chapter.

Transport Modes

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.

- **Direct Storage Access:** in this mode backup proxy server has direct storage access to the location 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 with FibreChannel, FCoE, iSCSI or NFS
- On a VM with iSCSI and NFS
- Both options can be used for Storage Snapshot Integration which uses the Backup from Storage SnapShot (BfSS) feature with a suitable storage array.
- **Virtual Appliance mode:** To work in this mode the backup proxy must 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 replication 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.

- **Network mode:** another option is to use the VMKernel Interfaces of VMware ESXi hosts to read and write VM data. As it needs no special configuration working 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/write through the VMKernel Interfaces by NFC protocol (via TCP port 902). Data Transport by use of NFC is limited by VMware to protect management datastreams to around about 40% of the available throughput. The Network proxy can be deployed as a physical server or a VM.

  Note: This backup mode is also referred to as NBD.

The following sections explain transport modes in detail.

Direct Storage Access

Under "Direct Storage Access" mode selection Veeam summarize the VMware own "Direct Storage Access" and the Veeam own "Direct NFS Access" modes.

The Direct Storage Access Mode uses a direct data path (a Fibre Channel or iSCSI) between the VMFS datastore and the backup proxy for data transfer. The Proxy need at least read access to the datastores so Fibre Channel Zoning, Networking and LUN masking on the Storage need to reflect this.

To use Direct NFS backup mode the Proxies need access to the NFS network and need to be member of the NFS Storage System "export policy" for read- write access.

Pros

- Direct Storage Access mode provides very fast and the most reliable predictable backup performance (typically, using 8 Gb Fibre Channel or 10 GbE for iSCSI and NFS).
- Produces zero impact on vSphere hosts and VM production networks for backup data transport.
- It is also possible to perform full VM restore using Direct Storage Access. This mode will be used automatically if eligible backup proxies are available in the backup infrastructure, and the VM disks are thick provisioned.
- Direct Storage Access is the fastest backup and restore mode for NFS datastores. It uses multiple concurrent read and write streams with an increased queue depth.
- Direct Storage Access for NFS datastores will mitigate the "VM stun" issues that may be caused by Virtual Appliance Mode (hot-add).
- Direct Storage Access for FC and iSCSI can be used for replication at the target for the initial replication (with thick provisioned disks) only. For NFS datastores, Direct Storage Access can be used for initial and incremental replication passes. There are no differences on the source replication proxy.

Cons

- Typically Direct Storage Access requires a physical server for a Fibre Channel, iSCSI or NFS connection. For virtual only deployments, Direct Storage Access for iSCSI and NFS is possible, but would transport the data through networks of the ESXi hosts.
- Restore via Direct Storage Access using Fibre Channel or iSCSI is possible only for thick-provisioned VM disks. At restore the data stream needs to be coordinated in the background with vCenter or an ESXi host which can slow down the restore speed. Consider adding additional hot-add proxy servers for restore (FC/iSCSI only).
- Direct SAN mode (FC/iSCSI only) is the most difficult backup mode to configure as it involves reconfiguring not only the storage but also the SAN, (Fibre Channel 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, NIC drivers and firmwares are up-to-date and that multi path driver software (e.g. MPIO) is properly configured.

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

Example

If datastores or virtual raw device mapping (vRDM) LUNs are connected via shared storage using Fibre Channel, FCoE or iSCSI, you may 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.

Ensure that a connection between the storage and backup proxy can be established. Verify FC HBAs, zoning, multipath, driver software and iSCSI configurations including any network changes. To test the connection it is best practices to count the volumes at windows disk manager adding one disk per storage system at a time. Later add all the others. This will streamline the search for errors.

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

- 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 this, the Veeam logging of the particular process can become the bottleneck you can disable this with the registry setting

  - Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`
  - Key: `VDDKLogLevel`
  - Type: REG_DWORD
  - Value: 0
  - Default: 1

  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 Storage Access-related challenges).

- To achieve performance versus cost use fewer proxies with more CPU cores available. This will help to fully utilize the HBA or NIC capacity of each proxy server. A 2 CPU System with 2x 12 cores is seen as a good balanced configuration between throughput and costs.

Security Considerations for Direct SAN mode with FibreChannel or iSCSI

During deployment of the proxy role to a Windows VM, 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. However if you want to be able to restore thick VM disks by this mode you have to disable `automount` and enable the volumes. If you set the disks read only Veeam will choose another restore mode and it will failover to NBD mode through the same proxy.

- Veeam deploys VMware's VDDK Kit into the backup proxy server, in most cases this VDDK Kit coordinates read and write processes (Direct SAN restore) with VMware vSphere allowing VMware's Software to control the read and write streams in a reliable manner.

If necessary you can take additional measures as follows:

- Disable automount. Open a command box and type "diskpart". Then disable automaount with "automount disable".

- 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 storage. When possible implement read-only LUN masking on the storage system or read-only zoning on the Fibre Channel switches (possible on most Brocade variants).

If a VMFS datastore is manually brought online in Windows Disk Management by mistake, and disk resignaturing is initiated, the datastore will become unavailable, and VMs will stop. Please contact VMware Support for assistance with recreating the VMFS disk signature. 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

Summary

Use Direct Storage Access whenever you have the chance to do so for fast backup speed and reduced load on the ESXi host. For fibre channel and iSCSI based datastore it processes backup data reliably and very fast. Consider using HotAdd proxies as these restores are much faster than Direct SAN restores. Direct SAN uses VDDK which can cause excessive metadata updates while hot-add restore bypasses VDDK.

For NFS datastores Veeam's Direct NFS mode is the mode you should choose for backup and restore as it delivers top speed without any negative side effects. You can use it for virtual and physical proxy deployments.

Virtual Appliance Mode

Veeam's Virtual Appliance (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 having 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.

One example of where Virtual Appliance mode can be recommended for proxies is in highly dynamic environment where it can be difficult to maintain access to newly created datastores for Direct Storage 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 not possible to leverage Direct Storage modes, for example in the case of local datastores, then using VMware Hot-Add will likely provide the optimal throughput.

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) it can add up to 1-2 minutes per VM. 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.

Veeam developed Direct Storage Access for NFS based datastores to overcome the problems with disk hot-add and release which causes significant stuns for NFS based VMs). Direct Storage Access should be used for all virtual and physical proxy deployment to backup and restore NFS datastore based VMs.

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 satisfactory performance in environments running 1 GbE configurations.

Cons

- If working in this mode the 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.
- In situations with a high number of VMware clusters with individual datastores a minimum of one proxy per cluster is needed, this can increase management overhead.

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 situation is most likely to happen on NFS based storage). To mitigate the issue, Veeam implemented the following functionality:

- Veeam Snapshot Hunter. This feature automatically initiates disk consolidation for VMs in the "Virtual machine disks consolidation is needed" state. For more information please see Snapshot Hunter section
- Bypassing Virtual Disk Development Kit (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 "hot-add I/O bursts" during hot-add operations are mitigated by bypassing VMware VDDK during restores and replication.
  - When performing writes via hot-add and VDDK, excessive metadata updates on the VMFS datastore will occur. This significantly impacts performance for other

workloads on the datastore, and slows down restore throughput. Bypassing VDDK helps overcoming this limitation

- 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 . To avoid this completely we highly recommend you to use the Direct NFS backup mode for backup and restore of NFS datastore based VMs.

Note: For additional tips refer to the “Impact of Snapshot Operations” section of this guide.

Recommendations

- You will need at least one type of (virtual) SCSI controller added to Proxy Server VM that is used somewhere at the VMs in your infrastructure to allow VMware to HotAdd the VM disks at backup.
- Add an extra SCSI controller to allow for more VM disks processing in parallel (check the corresponding Veeam proxy settings, default value is 4). The limit for a single controller is the maximum number of devices per SCSI controller (15). Max SCSI controllers per VM is 4 = 60 disks max. Adding one additional SCSI controller is usually sufficient.
- When deploying hot-add backup proxies avoid cloning existing VMs as this may lead to identical UUIDs 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.

  Specific client OS limitations for Hot-Add processing are documented in Veeam Backup & Replication Release Notes at https://www.veeam.com/veeam_backup_9_0_release_notes_en_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.

Network Mode

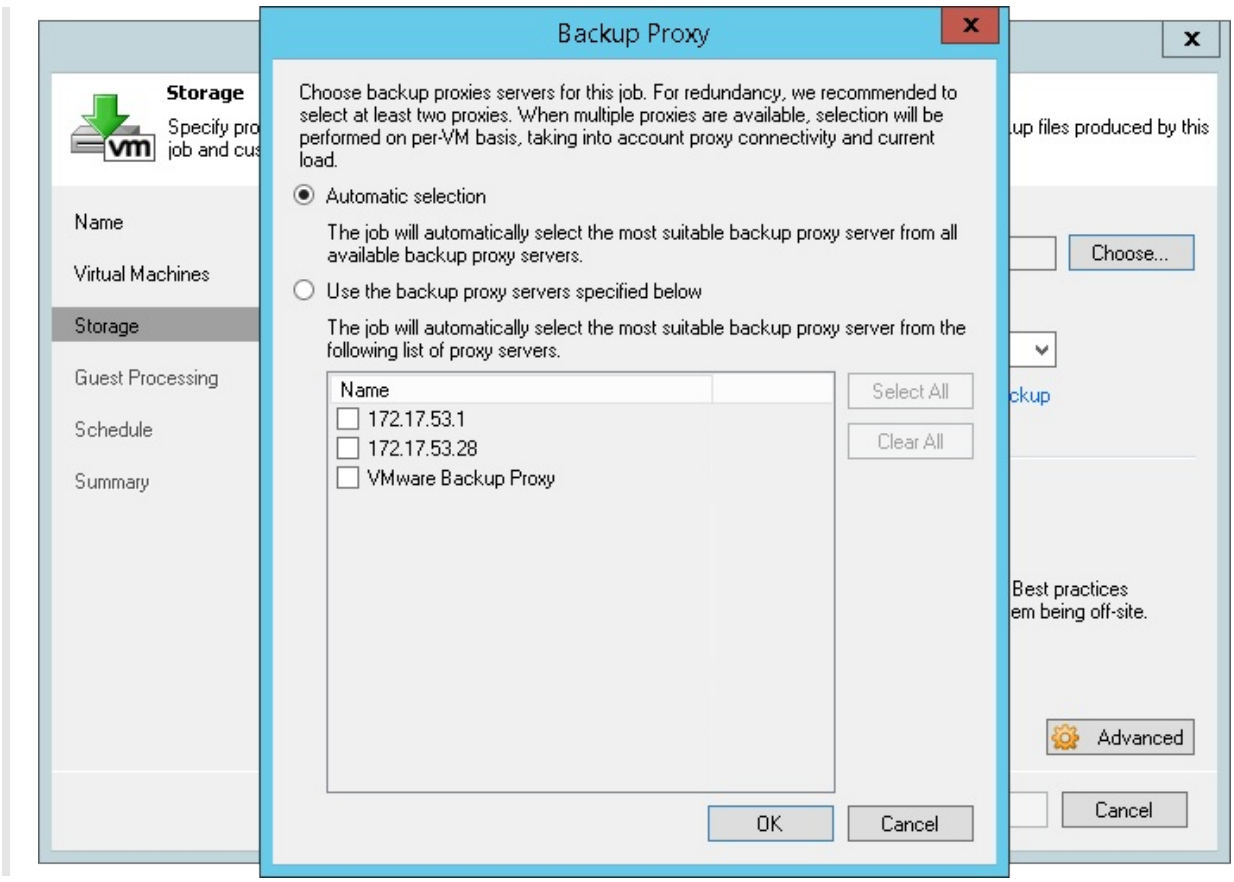
Network mode is by far the easiest backup mode to implement as it requires no additional configuration. Veeam uses the same interface to backup and restore VMware configuration files and to read Change Block Tracking (CBT) information.

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: https://helpcenter.veeam.com/backup/vsphere/used_ports.html.

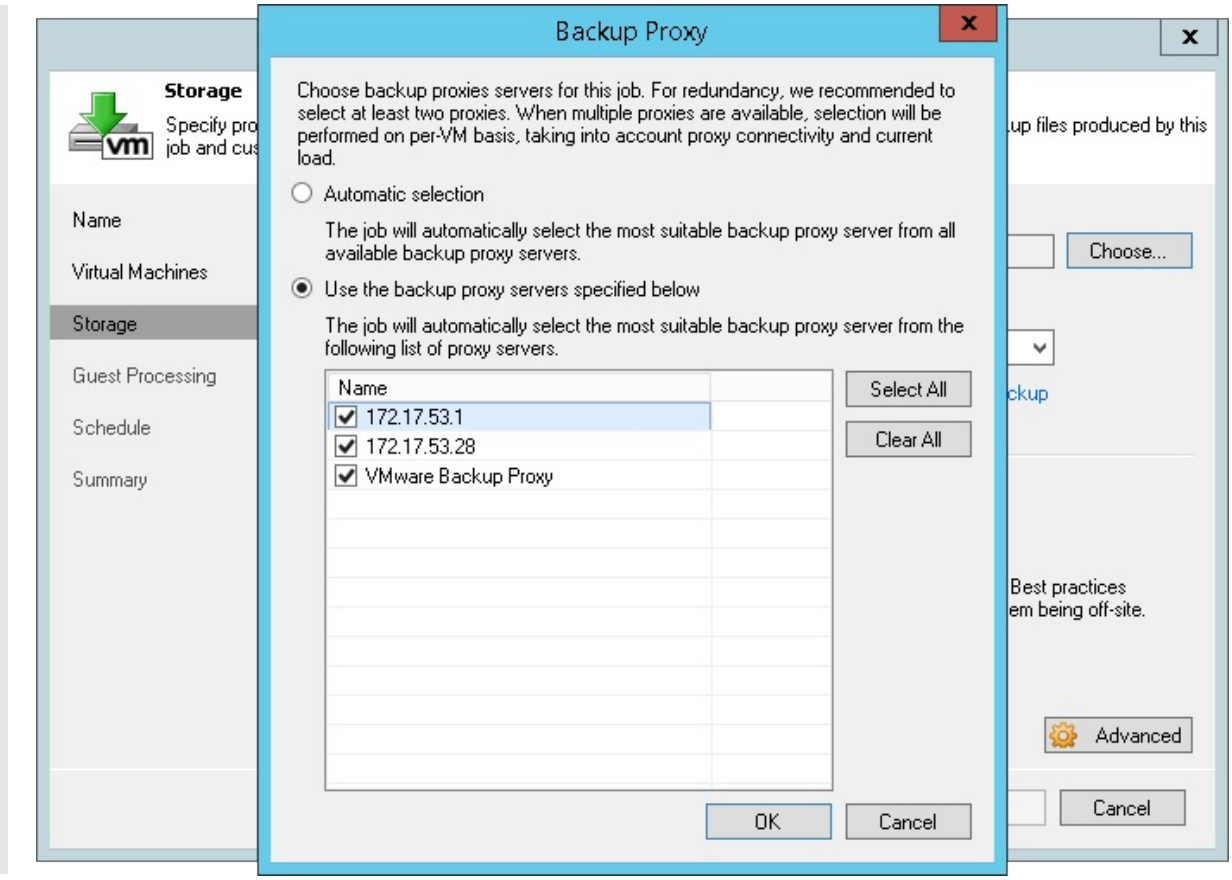
Note: It is highly recommended to maintain a good network connection between the VMware ESXi VMKernel port and Veeam Backup & Replication as it will be used by many other features like Instant VM Recovery, Virtual Lab and SureBackup, Linux FLR appliance, config files backups etc.

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 (the Interface IP that is bound to the FQDN 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 with same speed.
- 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 in cases where Direct NFS is unavailable 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 specifically as the load is spread across all ESXi hosts.
- 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).

- It 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-6.x. Sometime less.
- It can be rather slow on 1 Gb Ethernet (about 10-20 MB/s) due to throttling mechanisms so restores via the Network mode can take quite a long time with 1GbE.

Tip: You can influence the usage of the specific VMKernel interface by modifying the DNS name resolution for the ESXi hosts on the Backup & Replication Sever and on all Proxy Servers, for example, by adding entries in the *hosts* file of the Windows OS or by using special DNS configuration. See section on DNS at the start of this documented.

Recommendations

When you choose the network mode not having to deal with hot-add vCenter and ESXi overhead or physical SAN configurations. The network mode (NBD) is a very reliable way to do backups. In emergency situations when you need fast restore the following tips can be helpful:

- Consider setting up at least one virtual backup proxy for hot-add based 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. Thin disk restores have to transport less data.
- Another way to overcome this limitation is to use Instant VM Recovery with Storage vMotion (if licensed on the VMware vSphere side) as it is not affected by any throughput limitations of VMware.

General recommendations:

- 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).

- To protect VMware, Veeam reduces the amount of Network Mode data transport connections to 7. You can add a registry key to Veeam to increase that number.

  ViHostConcurrentNfcConnections = 7 (or higher) (DWORD)

  in the

  HKLM\SOFTWARE\Veeam\Veeam Backup and Replication registry key.

More data connections uses more RAM buffers and the default reserved RAM on ESXi hosts can produce failed jobs. Consider increasing NFC RAM buffer sizes on the ESXi hosts if you increase the above registry setting for example from 16384 to 32768 MB. http://kb.vmware.com/kb/2052302 After increasing NFC buffer setting, you can increase the following Veeam Registry setting to add addition Veeam NBD connections:
HKLM\SOFTWARE\VeeaM\Veeam Backup and Replication
ViHostConcurrentNfcConnections

DWORD (decimal) Default: 7

Be careful with this setting. If the buffer vs NFC Connection ratio is too aggressive, Jobs may fail.

Backup from Storage Snapshots

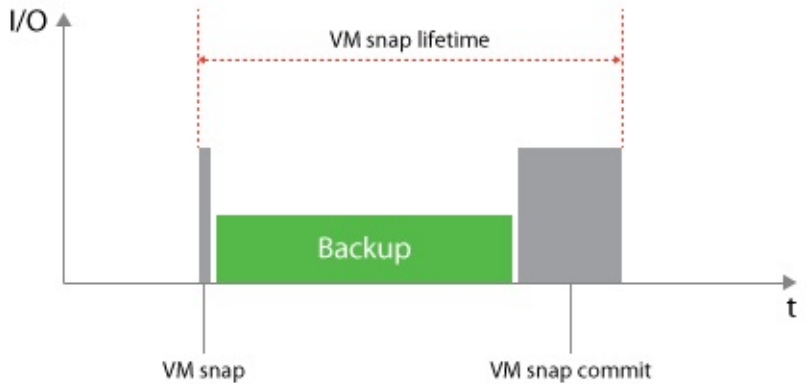
Veeam Backup & Replication offers integration with certain storage arrays for VM snapshot offloading. The following storage vendors and arrays are currently supported:

- HPE StoreVirtual (LeftHand)
- HPE StoreServ (3PAR)
- NetApp Data ONTAP (FAS, V-Series and IBM N series)
- EMC VNX, VNXe and Unity

Licensing and system requirements are described in the Veeam User Guide: Backup from Storage Snapshots.

The storage integration covered in this section is VMware only and does not apply for Hyper-V.

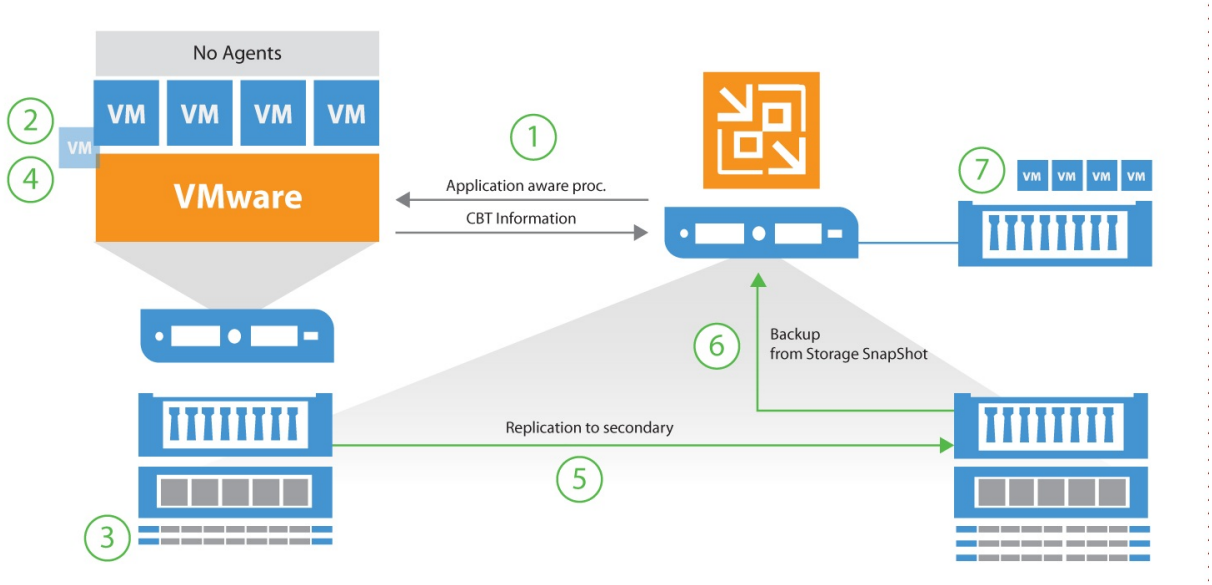
Backup from Storage Snapshots (BfSS) is a feature included in the deep storage array integrations and a way to optimize and enhance VM backups in a very easy way. The main objective for implementing BfSS is to minimize the lifetime of a VM snapshot, which reduces the time for VM snapshot commit and I/O the vSphere environment.



For regular VADP based backups, the VM snapshot is created and remains open (VM snap lifetime) until the VM backup is completed. Especially with large or highly transactional VMs, that can lead to large snapshot delta files being created during the backup followed by hours of snapshot commit tasks within vSphere producing high I/O on the production storage. Ultimately, these long snapshot commits may lead to unresponsive VMs. For more information about the impact of VM snapshots please see the "Interaction with vSphere" section of this book.

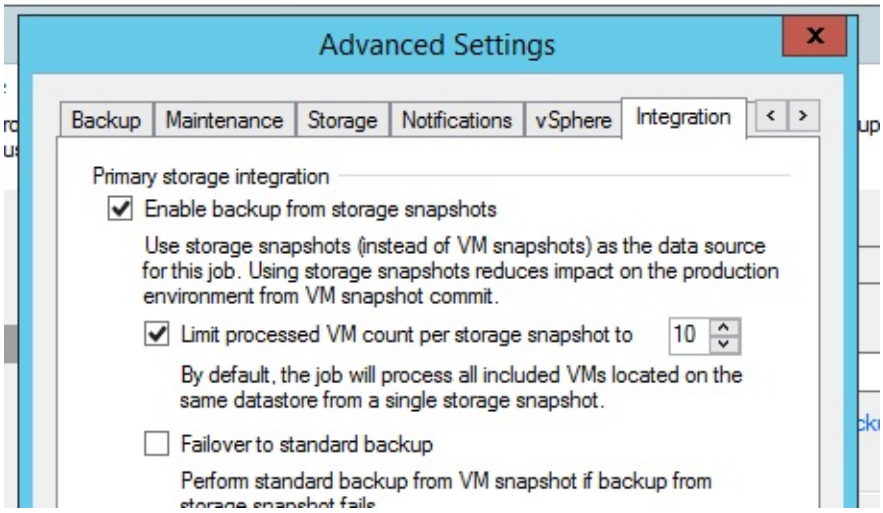
How it works

By using BfSS, the VM snapshot lifetime will be significantly reduced. In this section, we will go through the steps performed.



1. Application-aware processing ensures transactional consistency within the VM
2. Veeam requests a VM snapshot via VMware APIs
3. Immediately after creating the VM snapshot, a storage snapshot request is issued for saving the VM *including* the application consistent VM snapshot within the storage snapshot.
4. When the storage snapshot has been created, the VM snapshot is deleted
5. *(NetApp only - optional)* Trigger a replication update to secondary storage via SnapMirror or SnapVault
6. Mount storage snapshot to the Veeam backup proxy server
7. Read data from the storage snapshot and write to a Veeam backup repository

VM processing limit

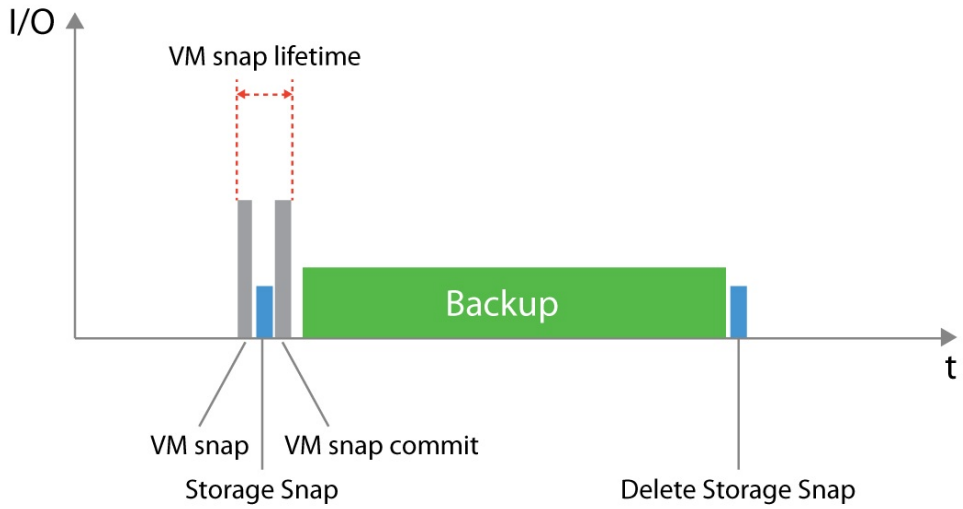


When adding large amounts of virtual machines to a job, by default steps 1 and 2 (above) are repeated until all virtual machines within the job have successfully completed. Only then will BfSS proceed to step 3 and issue the storage snapshot. If adding 100s of jobs to a backup or replication job, this could cause a very high VM snapshot lifetime for the first VMs in the job list.

When configuring such large jobs, it is advised to configure the maximum number of VMs within one storage snapshot. The setting is available in the advanced job settings under the **Integration** tab.

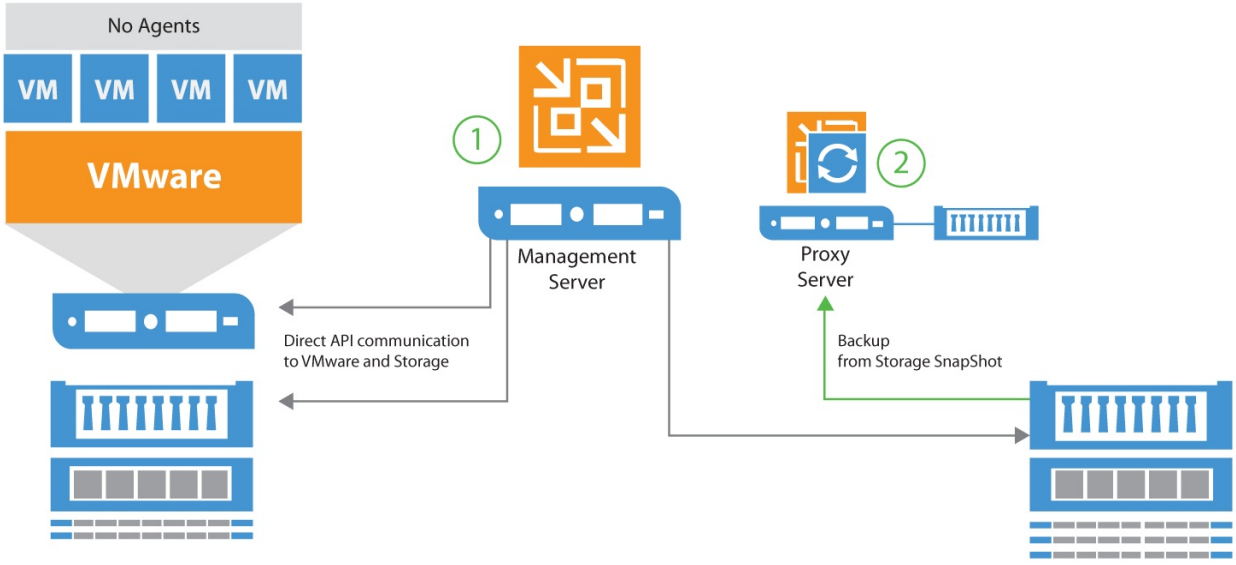
Example: When creating a job with 100 VMs, and setting the limit to 10, BfSS will instruct the job manager to process the first 10 VMs (step 1 and 2), issue the storage snapshot and proceed with the backup (step 3-7). When step 7 has successfully completed for the first 10 VMs, the job will repeat the above for the following 10 VMs in the job.

As seen below, when ensuring proper configuration of BfSS, minimal VM snapshot lifetime is achieved, and reduces overall I/O penalty on the production storage for highly transactional VMs.



Configuration

Enabling BfSS requires minimal configuration, but understanding the tasks and responsibilities of involved components are key when troubleshooting and optimizing for high performance and low RTPO.



The backup server is responsible for all API requests towards vSphere and storage arrays for determining present volumes, snapshots and all necessary details such as initiator groups, LUN mappings and which protocols are available.

The proxy server(s) are used for reading data from the storage snapshot and sending it to the backup repository. To leverage Backup from Storage Snapshots, the following configuration requirements must be met:

Backup server must have access to the management interfaces of the storage array. All additional prerequisites such as LUN mappings, creation of initiator groups for iSCSI, altering NFS exports and snapshot management is subsequently handled via this connection.

Backup proxy servers must be able to directly access the storage array via the same protocol used for connecting the production datastore (FibreChannel, iSCSI or NFS). As opposed to using Direct Storage Access, it is not a requirement for the proxy server to have access to the production datastore itself, as it reads data blocks directly from the cloned storage snapshot.

As described in previous sections, the backup server and proxy server can be deployed on one single server or scaled out on different servers. In most environments, where BfSS is applicable, the components are usually separated for scalability considerations.

When to use

When using Backup from Storage Snapshots, overall jobs processing may take longer, as additional steps are performed such as mapping vSphere Changed Block Tracking (CBT) to offsets of the storage snapshot, and the snapshot must be cloned and mounted on the backup proxy server. The mount overhead can take several seconds on block protocols as HBAs or initiators must be rescanned. It mostly affect FC deployments.

With this in mind, using BfSS on small VMs or VMs with a very low change rate is not advised. As the VM snapshot lifetime on such VMs is very short, the benefits of using BfSS are minimal.

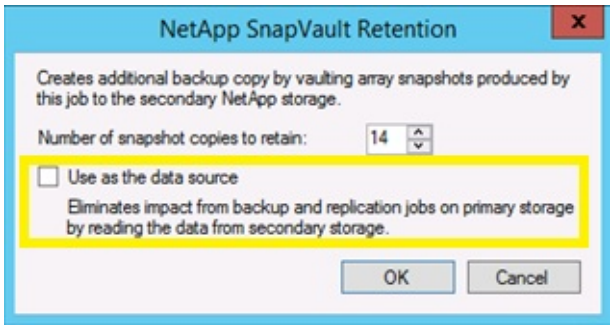
In most environments, large VMs or highly transactional VMs producing large amounts of changed data benefit most from using BfSS. Using the VM Change Rate Estimation report in Veeam Availability Suite, you may quickly identify such VMs.

VMs with either virtual or physical Raw Device Mapping (RDM) are not supported with BfSS. Such VMs will failover to backing up via standard methods if allowed in the job settings.

NetApp Data ONTAP

Specifically for NetApp Data ONTAP, Veeam offers some specific additional capabilities.

Backup from secondary snapshots

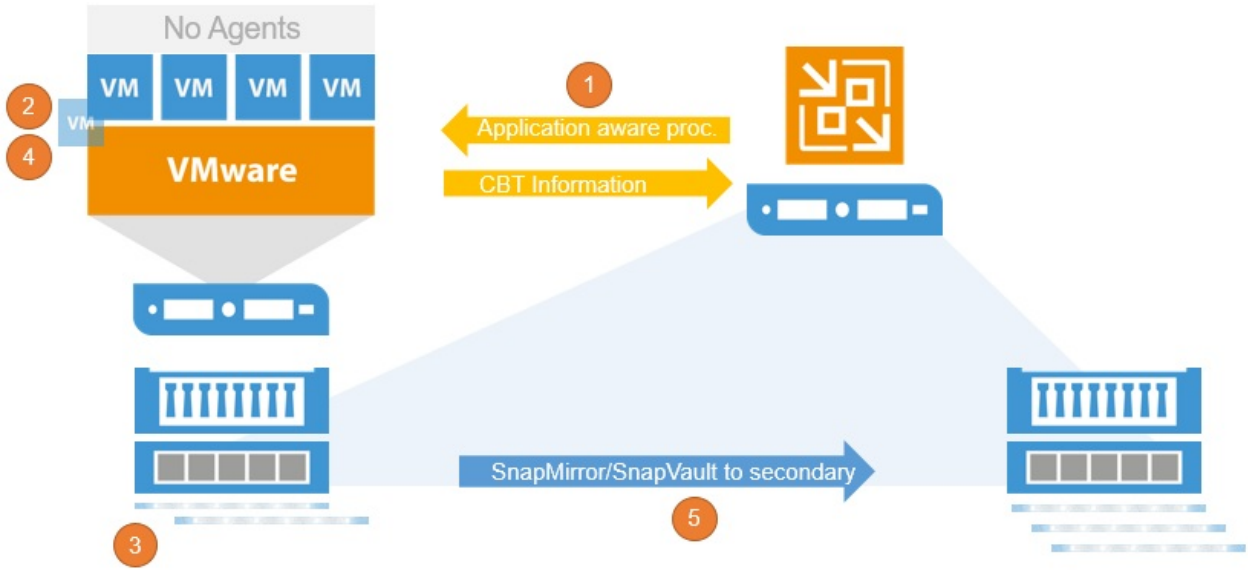


Backup from Secondary Snapshots. In case you use NetApp SnapVault or SnapMirror, Veeam can create a primary snapshot, update the secondary (SV/SM) Snapshot and backup the CBT changes to the backup file. It is configured with a job setting in the "Advanced" section if Veeam should allow fallback to the primary snapshot for backup. You can find the setting within the secondary destination window of your backup job and enable “Use as the data source”.

Snapshot Orchestration

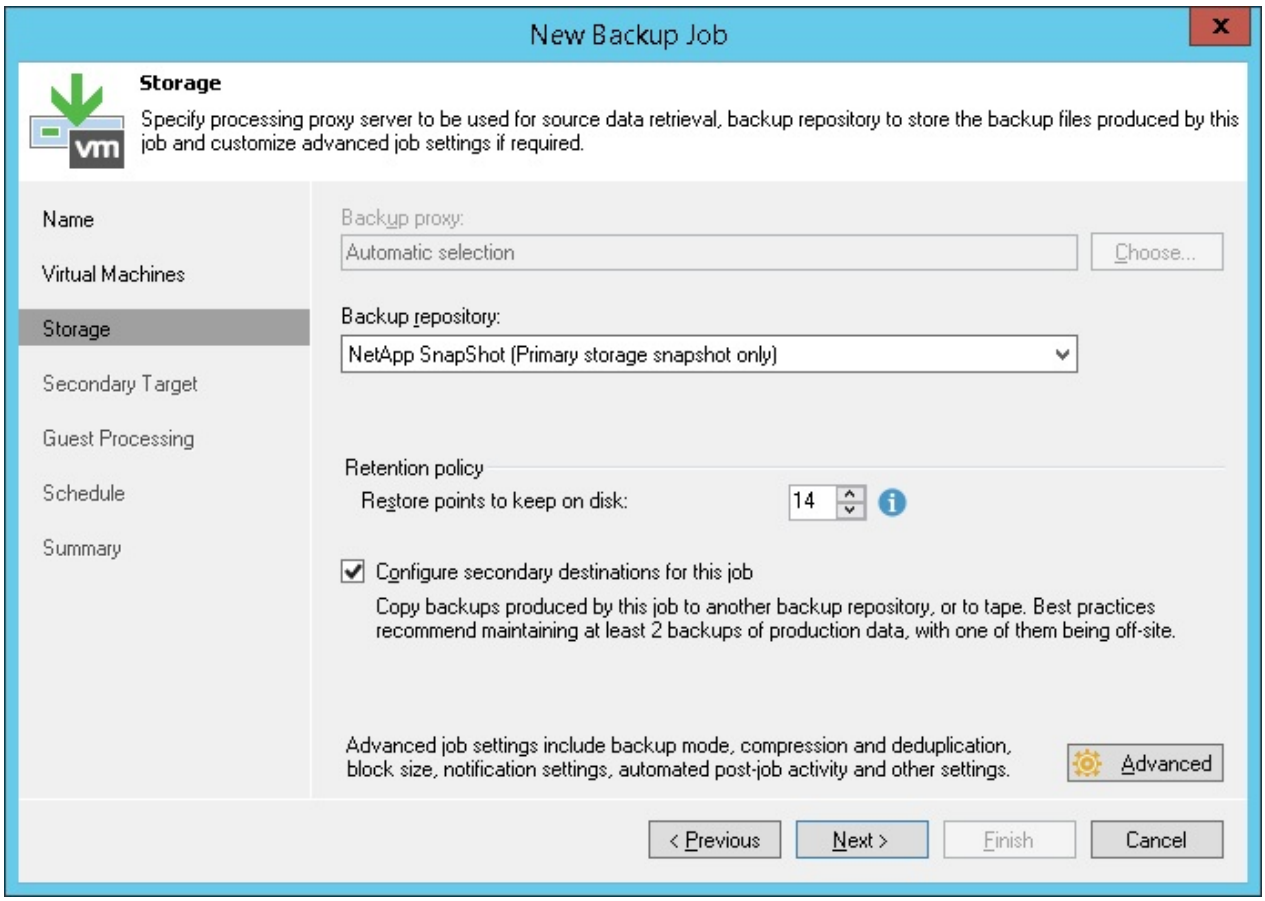
For NetApp ONTAP storage systems Veeam offers a SnapShot Orchestration only feature. SnapShot orchestration means to use storage SnapShots as backup target. The feature can be used without any need to run a real backup to an external repository. Veeam is taking care of all required storage related tasks like data retention, SnapShot management and SnapMirror/SnapVault updates to secondary sides.

The workflow for Storage Orchestration is:



1. *(Optional)* Application-aware processing ensures transactional consistency within the VM
2. Veeam requests a VM snapshot via VADP
3. Immediately after creating the VM snapshot, a storage snapshot request is issued for saving the VM *including* the application consistent VM snapshot within the storage snapshot.
4. When the storage snapshot has been created, the VM snapshot is deleted
5. Trigger a replication update to secondary storage via SnapMirror or SnapVault

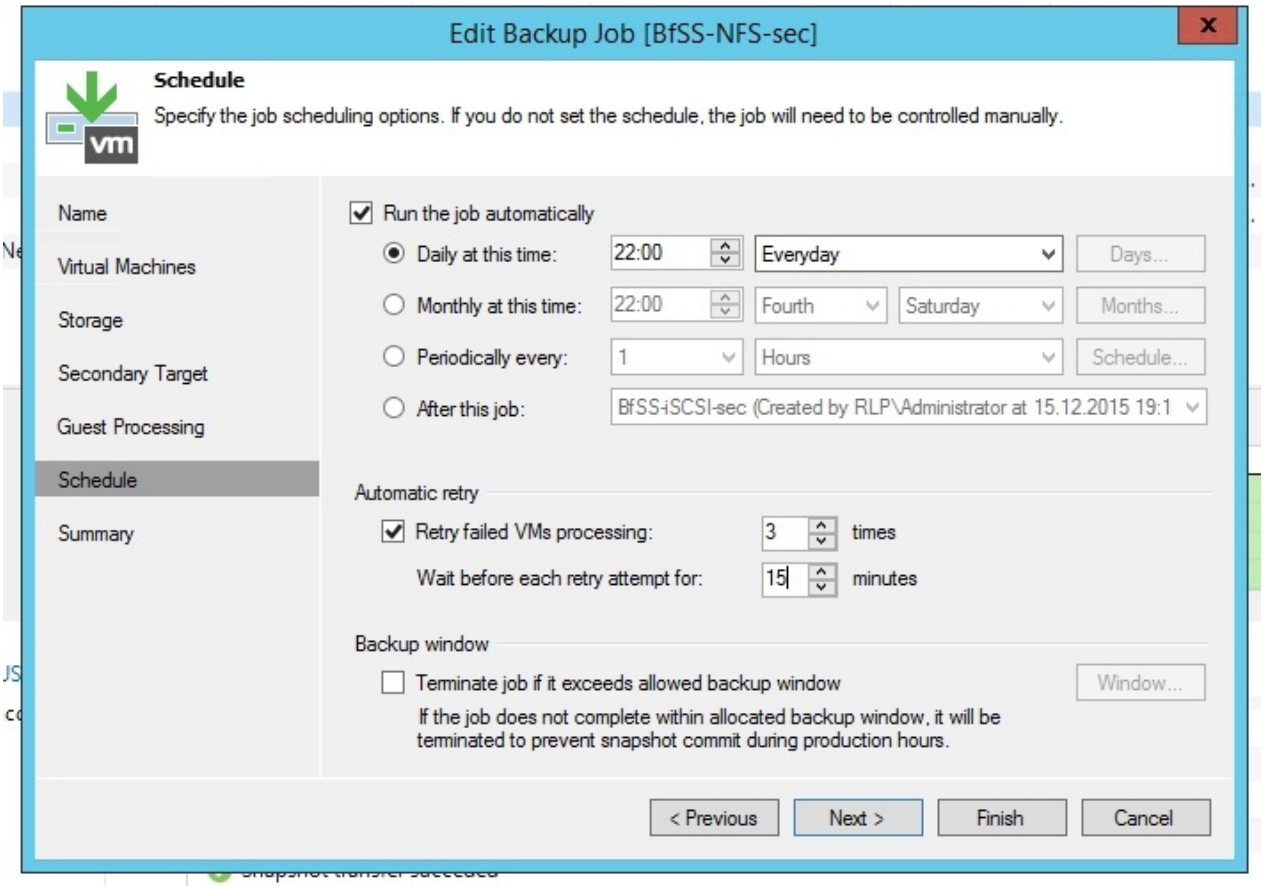
To configure a “SnapShot only” job set the Repository to "NetApp SnapShot only"



The retention policy defines the number of storage snapshots to keep. To store 5 snapshots a day for 1 week, configure the retention to 35 restore points with a daily schedule. If the job is configured with a high or lower schedule frequency, adjust the number of restore points accordingly.

If you use a secondary NetApp ONTAP system with SnapMirror and/or SnapVault you can set the box for a secondary destination and set the retention.

When using Snapshot Orchestration please take care of the retry scheduler setting.



If you have for example 100 VMs in one job and 10 of these VMs are failing in the first run Veeam will rerun the job based on the retry settings. If the setting is set to 3 (default) Veeam will try 3 more time to process the failed VMs. For every successful retry Veeam will create a new Snapshot. If all retries are needed to proceed the failed VMs that ends in 3 Snapshots for one run. It is recommended to not set the value higher than 3 or disable the automatic retry to avoid a high number of Snapshots being created during every run.

One of the big benefits is that you are still able to use all Veeam restore capabilities from storage snapshots. For more please refer to the Veeam Explorer for Storage Snapshots section.

Selecting a Transport Mode

Depending on the size of the environment, there are different recommendations for selecting a transport mode. For simplicity, a couple of definitions will be used in this section:

Name	Definition
Very small	Single host with local disks as primary datastores. Typical ROBO configuration.
Small	2-4 hosts with shared storage. Typical ROBO configuration or small datacenter
Medium	4-20 hosts with shared storage
Large	20-100 hosts with shared storage
Enterprise	Over 100 hosts

Keep in mind that within larger datacenters, multiple definitions may apply. As an example, it is possible that a separate management or DMZ cluster without shared storage could benefit from using the "Very small" or "Small" recommendations, while the main production environment is leveraging recommendations based on "Medium" to "Enterprise" datacenter size.

Very small

- 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 without any special configuration needed.
- NBD over 1GbE VMKernel interfaces can be used for failover.
- Direct Storage Access mode or Backup from Storage Snapshots modes are typically unavailable, as the disks of the host are local and thus cannot be mounted to an external proxy server.

Small and Medium

- If storage integration is available, use Backup from Storage Snapshots (BfSS)[1]

- For NFS based Storage, use Direct Storage Access
- For shared storage connected via FC or iSCSI, you can choose one of the following two modes:
    - **Physical proxy:** Direct Storage Access will provide the best backup performance. For example, you can configure a physical server with access to FC datastores on the local site and perform backups to a local repository. If you use thin-provisioned disks for the VMs, configuring a dedicated backup proxy for restoring via Virtual Appliance (hot-add) mode can help to increasing restore performance.
    - **Virtual proxy:** The Virtual Appliance (hot-add) mode is a good an fast backup mode. Avoid to backing up VMs on NFS datastores using hot-add. Use Direct Storage Access or NBD backup modes instead.
- 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.
- When using NBD, check the Network Mode chapter for tuning tips.

Large

In addition to the above considerations for Small and Medium, please see the following guidelines:

- When Direct Storage Access, or Backup from Storage Snapshots are unavailable, and when virtual proxy servers are disallowed, Network Mode (NBD) is the only choice. In such cases, 10GbE interfaces are a must.
- For virtual only deployments (virtual proxies only) in environments with many isolated clusters, using network mode (NBD) may be ideal. As hot-add requires at least one proxy within each cluster, it may require many more proxy servers compared to using network mode.
- A combination of hot-add mode for large clusters and NBD mode for smaller clusters may be ideal.

Enterprise

In addition to the above considerations for Large, please see the following guidelines:

- In large enterprise scale environments, the deployment of Veeam components, configuration and job creation is typically automated using the Veeam PowerShell SDK.
- To balance the management load, it is recommended to use multiple Veeam backup servers for at least every 5,000 VMs and federate them for central reporting and administration by using either Veeam Enterprise Manager, Veeam Managed Backup Portal, Veeam Management Pack for Microsoft System Center Operations Manager or Veeam ONE.
- When running a central backup server and with multiple branches connected to it, a dedicated backup server is recommended for at least every 200 branches. Consider using Veeam Enterprise Manager for federation.

> [1]. In case storage integration is used with Backup from Storage Snapshots (BfSS), the overhead of mapping blocks from VMware CBT and the storage snapshot can increase processing time and lead to longer backup windows. To mitigate, consider the majority if the VMs can be backed up with one of the other transport modes and use BfSS only for the largest VMs or high change rates (typically 10% of VMs). Veeam ONE Change Rate Estimation report can help to identify such VMs. ↩

Sizing a Backup Proxy

Getting the right amount of processing power is essential to achieving the RTPO defined by the business. In this section, we will outline the recommendations to follow for appropriate sizing.

Processing Resources

As described above, you may define the max concurrent tasks value in the backup proxy settings. It is best practices to plan for 1 physical core or 1 vCPU and 2 GB of RAM for each of the tasks. A task processes 1 VM disk at a time and CPU/RAM resources are used for inline data deduplication, compression, encryption and other features that are running on the proxy itself.

In the User Guide it is stated that proxy servers require 2 GB RAM + 500 MB per task. Please consider these values as minimum requirements. Using the above mentioned recommendations allow for growth and additional inline processing features or other special job settings that increase RAM consumption.

If the proxy is used for other roles like Gateway Server for SMB shares, EMC DataDomain DDBoost, HPE StoreOnce Catalyst or if you run the backup repository on the server, remember stacking system requirements for all the different components. Please see related chapters for each components for further details.

Tip: Doubling the proxy server task count will - in general - reduce the backup window by 2x.

Total needed Task Slot Numbers

A general sizing rule of thumb is, to use 1 physical CPU core or vCPU and 2 GB RAM for each 30 VMs within an 8 hour backup window. Depending on the infrastructure and mainly the storage performance, these number can turn out to be too conservative, we recommend to do a POC and find out the specific numbers for the environment.

Calculating Overall Task Count Examples

Sample infrastructure has the following configuration:

- 480 VMs

- 48 TB used data
- Backup window: 8 hours
- Change rate: 5%

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

Using the "30 VMs per CPU core" rule, we get following result:

- 480 VMs / 30 VMs per core = 16 CPU cores

Each CPU core must have 2 GB RAM:

- 16 CPU cores x 2 GB RAM = 32 GB RAM.

Result: **16 CPU cores and 32 GB RAM**.

- For a physical server, it is recommended to install dual CPUs with 8 cores each.
- For virtual proxy servers, it is recommended to configure multiple proxies with maximum 8 vCPUs to avoid co-stop scheduling issues.

If you need to achieve a 2x smaller backup window (4 hours), then you may double the resources for a total of **32 CPU cores and 64 GB RAM** - 2x the amount of compute power (possibly split across multiple servers).

The same rule applies if the change rate is 2x higher (10% change rate). To process a 2x increase in amount of changed data, it is also required to double the proxy resources.

Note: 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?

- For per job backup files: 30 VMs per job
- For per VM backup files: 300 VMs per job

Consider that some tasks within a job are still sequential processes. For example a merge process that write the oldest incremental file into the full file is started after the last VM finishes backup processing. If you split the VMs into multiple jobs these background processes are parallelized and overall backup window can be lower. Be as well carefull with

big jobs when you use Storage Snapshots at Backup from Storage Snapshots. Guest processing and Scheduling of jobs that contain multiple snapshots can lead into difficult scheduling situation and Jobs that spend time waiting for (free) ressources. A good size for Jobs that write to per VM chain enabled repositories is 50-200 VMs per Job.

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 10-16 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 2x10 Core CPU with 48GB RAM and 2x 16 Gbps FC cards for read + 1-2 10GbE Network cards.

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). However if you have a lot of free CPU ressources at the backup time window, you can consider to use *High* compression mode.
- **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** - see the corresponding section of this document.
- **3rd party applications** – it is not recommended to use an application server as a backup proxy.

Backup Repository

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

A 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 runs the 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.

You can balance the load across the backup infrastructure by setting up several repositories in the environment and limiting the number of concurrent jobs for each repository, or if you have a proper license you can leverage Scale-out Backup Repository as explained later on in this section.

The 3-2-1 rule

The 3-2-1 rule states that an environment, in order to be properly protected, has to have **3 copies of data, stored on 2 different media, with at least 1 copy in a different location**. Each of the parts of the rule involves the use of a storage device, that's why a Backup Repository is such a key component in each Veeam deployment.

The 3-2-1 rule however is a data protection strategy, whereas **availability** requires the different storage implemented in this strategy to support additional capabilities like:

- Instant VM recovery
- File transforms
- Distant copies
- Item restoration
- Sure Backup

This is the reason why v9.0 introduced two major new features for Veeam backup repositories : **Scale-out Backup Repository** and **Per-VM Backup chains**.

Repository Type

Being storage-agnostic, Veeam Backup & Replication supports a wide range of repository types, each offering its own set of specific capabilities. So when deciding on repository storage, you might consider the following:

- Capacity
- Write performance
- Read performance
- Data density
- Security
- Backup file utilization

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

Organization policies may require different storages for backups with different retention. In such scenario, you may configure two backup repositories:

- A high-performance repository hosting several recent retention points for instant restores and other quick operations
- A repository with more capacity, but using a cheaper and slower storage, storing long-term retention points

You can consume both layers by setting up a backup copy job from the first to the second repository, or leverage Scale-out Backup Repository, if licensed.

Server-Based Repository: DAS or SAN?

Direct-Attached Storage

This is a cheap, easy-to-use solution that can be very efficient in terms of performance; however, if not used as part of a Scale-out Backup Repository, it is less manageable due to non-transportable volumes, capacity growth, and so on.

- Since a DAS storage can be fully dedicated to backup operations, this type of repository is considered to offer a good balance between “performance” and “cost” factors.
- A strong benefit of a DAS repository is that it supports the features offered by Veeam Backup & Replication in a very flexible way. In particular, it provides good read and write performance, sufficient for Veeam 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 an organization's requirements.

Tip: To achieve optimal performance, it is often required to install a battery module to the server's controller card in order to enable write-back mode for the internal cache. A DAS is a shelf with disks, and all the intelligence of the solution is delegated to the controller installed in the connected server.

Pros	Cons
Cost	Manageability
Performance	Single point of failure
Simplicity	Monolithic

SAN Storage

This is a more advanced and manageable solution that offers the same advantages as DAS, and adds more advantages like higher availability and resiliency.

The volume size and quantity are easily adjustable over time, thus offering a truly scalable capacity.

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 capabilities	

Windows or Linux?

The main difference between Windows and Linux in regards to Veeam repositories 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 or RDM mounted to the Repository VM).

Best practice is to avoid using the same storage technology that is used for the virtualized infrastructure, as the loss of this single system would lead to the loss of both copies of the data, the production ones and their backups.

In general we recommend whenever possible to use physical machines as repositories, in order to maximize performance and have a clear separation between the production environment that needs to be protected and the backup storage.

SMB Repository

While an SMB repository (also called CIFS, even if this term is technically wrong) 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.

Gateway Server

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

- Automatic selection of the server as the SMB gateway proxy (that is, the server that will host the target-side transport component and thus perform the role of "data writer" towards the SMB share itself).
- Specify a specific server (among the available managed Windows servers in Veeam Backup & Replication) as a SMB gateway proxy.

The second option is very helpful in situations where the SMB share is located on a remote location, since it avoids that the automatic selection uses a server that is not local to the SMB share, thus having all synthetic operations or backup copy jobs occuring over the WAN link (which is usually slower than the local link). It's always recommended to use an SMB gateway server as close as possible to the SMB storage. By specifying the SMB gateway you have a better chance of keeping the data flow under control and avoid data crossing the WAN links unnecessarily.

As single stream performance for SMB repositories may be suboptimal, you can potentially increase performance of your SMB storage by configuring several repositories pointing to the same folder using different gateway servers. With multiple proxies, the automatic SMB gateway may be a good option and can be configured by selecting **Automatic** from the drop-down list.

Tip: Gateway servers must be properly sized as regular Windows repositories. If you are using Automatic mode, remember that the same machine could be elected backup proxy and gateway server simultaneously. Apply sizing it accordingly.

Another option for increasing the number of streams is using per VM backup files. Please see the corresponding section of this guide for more information > Per VM backup files

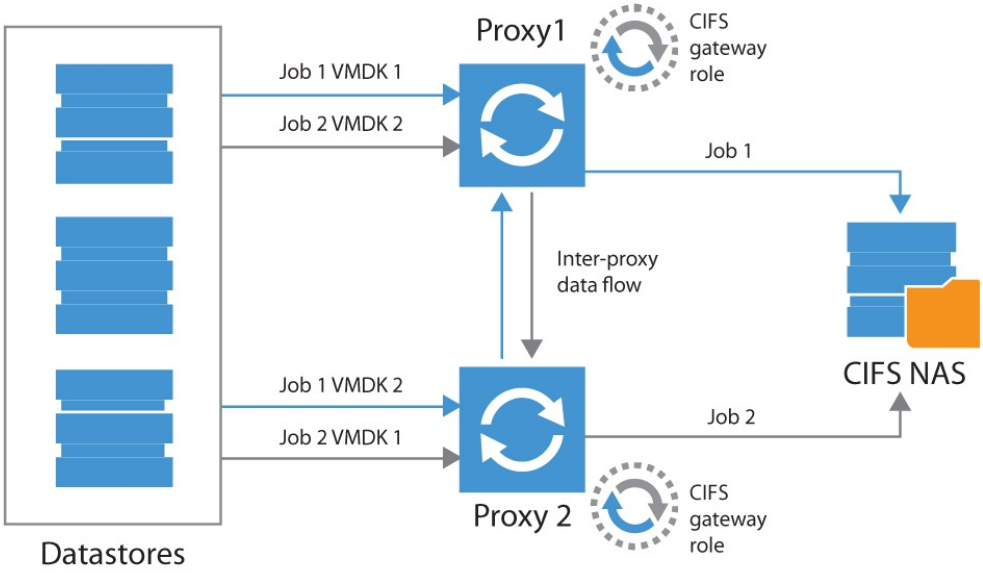
Load Balancing (with Automatic Selection)

Even when multiple proxies are used to process a given backup job, only **one*** Windows server (called “gateway server") per backup chain will be used to write data to the SMB share. In **Automatic** mode the first selected proxy in the running job will become the gateway server. If per-vm backup files are enabled, this applies to each per-vm chain, thus multiple gateway servers may be started concurrently.

Here are some recommendations for managing backup proxies used as gateway servers:

- The networking between the multiple proxies should be sized correctly to allow data to flow from each proxy to the gateway server.
- As the first backup proxy of a job is used as the gateway server, it may happen that all the gateway server instances of different jobs (or per-vm backup file chains) are started on the same proxy. This requires proper sizing of CPU and RAM; ensure resource monitoring is in place.

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



NOTE for Creative: rework the image with SMB instead of CIFS NAS

Scaling out using this approach will allow for processing larger amounts of data and optimize the throughput of the SMB 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 have non official implementations of the SMB protocol that may improve performance test results, but may also corrupt backup files. For these devices it is discouraged to use SMB.

Deduplication Appliances

Overview

Deduplication applied to storage is a technique aimed at reducing the storage space consumption.

Deduplicated storage systems are often optimized for write operations and can offer rather high ingest rates. However, any random read I/O (thus, restore operations) may suffer from re-hydration processes required during restores. For this reason we recommend to use these devices mainly as secondary targets, where parameters like price per GB are more important than restore performance.

Using a Deduplication Appliance

As a storage-agnostic product, Veeam Backup & Replication can use any deduplication appliance as a repository in different use cases: primary backup repository, backup copy repository, and Virtual Tape Library (VTL) container.

Deduplication Appliance as a Primary Backup Repository

Unless you are using DDBoost protocol on EMC DataDomain storage or Catalyst on HPE StoreOnce, you should configure primary jobs for forward incremental with active full backups - since jobs with transformation will require block "de-hydration" and then "re-hydration" on the storage. Such operations require significant time and I/O.

Note: "Re-hydration" is the act of restoring the original blocks in a non-deduplicated form. 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 deduplicated area for the most recent restore points (such as ExaGrid).

The downside of active fulls is the need to transport the entire amount of virtual machines on a weekly/monthly basis. This can lead to long snapshot commit, so this mode needs to be planned carefully. It is recommended to limit the use for primary backup jobs to the integrated deduplication appliances, where synthetic operations can be used.

Using Deduplication Appliance as a Backup Copy Repository

By default a backup copy job applies transformations to the backup chain. This could lead to the "de-hydration"/"re-hydration" overhead at the end of the backup copy job cycle, due to synthetic full or transformation). When using non integrated appliances, use the option of Active Fulls for Backup Copy jobs.

If one of the integrated appliance is used, synthetic operations will be performed on the appliance itself, so they will require minimal additional time and lower I/O.

Using Deduplication Appliance as a Virtual Tape Library

If a deduplication appliance is used in Virtual Tape Library (VTL) mode, it is required to store the backup files in a staging area, which is uncompressed. Sending compressed and/or deduplicated backup files to a VTL will compromise the efficiency of the deduplication appliance.

The repository used for staging should be configured with "Decompress before storing" advanced option enabled, which ensures previously applied compression at the job level is ignored.

Also, ensure that the appliance meets Veeam tape requirements described in the User Guide.

File-Level Recovery and Veeam Explorers

By design, Veeam Explorers perform a large amount of random read operations on the backup repository. To optimize for such operations on deduplication devices, following the job and repository configuration best practices (see below) is paramount. If the recommendations are not fully implemented, this may lead to significant waiting time when launching file-level recovery or Veeam Explorers.

To further reduce restore time, it is recommended to enable file-level indexing for backup jobs located on deduplication repositories. Indexing VMs will remove the waiting time for mounting a restore point when browsing contents via Enterprise Manager.

Best Practices

In this section, we will distinguish between integrated and non-integrated deduplication appliances. Integration is available for:

Integrated appliances are:

- HPE StoreOnce - via Catalyst API
- EMC DataDomain - via DDBoost API
- ExaGrid - via integrated Veeam datamover

If the mentioned integration API is unavailable due to licensing restrictions, or if any other deduplication appliance is used, the appliance should be considered *non-integrated*.

In order to optimize throughput for deduplication appliances, please use the following configuration guidelines:

Job configuration

The following settings are configured in the backup job "Edit" wizard under Storage > Advanced. Options not defined in this table are optional and not related to backup repositories using deduplication storage.

Configuration tab	Setting	Value
Backup	Backup mode	Incremental
Backup	Create synthetic full backups periodically	Enabled - if integrated
Backup	Transform previous backup chains into rollbacks	Disabled
Backup	Create active full backups periodically	Enabled - if non-integrated
Maintenance	Perform backup file health check	Disabled
Maintenance	Defragment and compact full backup file	Disabled
Storage	Enable inline data deduplication	Disabled
Storage	Exclude swap file blocks	Enabled
Storage	Exclude deleted file blocks	Enabled
Storage	Compression level	Optimal
Storage	Storage optimization	Local target (16TB+ backup files)
Storage	Enable backup file encryption	Disabled

Hardware assisted encryption is available for EMC DataDomain via DDBoost, but must be configured in the integration specific repository configuration. If enabled on the job level data reduction efficiency will be significantly degraded.

Repository configuration

The following settings are configured in the "Edit Repository" wizard under Repository > Advanced.

Setting	Value
Align backup file data blocks	Enabled - only if repository uses fixed block size deduplication (almost never true)
Decompress backup data blocks before storing	Enabled
This repository is backed by rotated hard drives	Disabled
Use per-VM backup files	Enabled

Deduplication integration specifics

EMC DataDomain

Selecting DataDomain as a repository will automatically recommend job and repository settings according to best practices. For more information, refer to vendor guidelines.

DDBoost allows for the following capabilities:

- Source side deduplication between the Veeam gateway server and DataDomain appliance. This will reduce the amount of data sent over the network to the appliance
- 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 Fibre Channel SAN, providing a totally LAN-free backup solution

For more details, refer to the DDBoost configuration guide by Rick Vanover: Configuring EMC Data Domain Boost with Veeam Availability Suite (still applicable for version 9).

Chain Length Limitation

Consider that DataDomain can support only up to 60 incremental restore points for a single full backup. For details, refer to the Veeam Backup & Replication User Guide: Limitations for EMC Data Domain

ExaGrid

ExaGrid appliances run an integrated Veeam data mover similar to a Linux based backup repository. With ExaGrid, there is no requirement for a Windows based gateway server.

See Using Veeam Backup and Replication Software with an ExaGrid System for more information.

ExaGrid recommends configuring 1 job per repository. Thus, if you want to achieve parallel processing, create several repositories and setup 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 have sufficient capacity for an uncompressed full backup so that each backup can fully be written there and processed. This ensures

SureBackup, Instant VM Recovery and item-level restores will be usable for the latest restore point without rehydration overhead.

HPE StoreOnce

Selecting StoreOnce appliance as a repository will automatically recommend job and repository settings according to best practices. For more information, refer to vendor guidelines.

When using HPE Catalyst, consider the following recommendations:

If the Catalyst Store is configured as **High Bandwidth** on the appliance, Low Bandwidth mode can be forced using the following registry value (ideally, work around the issue by configuring both Primary and Secondary modes to "Low"):

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup Transport`
- Key: `UseLowBandwithMode`
- Type: REG_DWORD
- Value: 1 (default: 0)

If the Catalyst Store is configured as **Low Bandwidth**, additional payload verification is introduced. Over high latency connections, disabling the verification may improve performance. However, the defaults should be left for local connections.

See the following registry keys:

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Wow6432Node\Veeam\Veeam Backup Transport`
- Key: `PayloadChecksumsDisabled`
- Type: REG_DWORD
- Value: 1 (default: 0)

and

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Wow6432Node\Veeam\Veeam Backup Transport`
- Key: `BodyPayloadCompressionDisabled`
- Type: REG_DWORD
- Value: 1 (default: 0)

Chain Length Limitation

Consider that HPE StoreOnce can support only up to 7 restore points. For details, refer to the Veeam Backup & Replication User Guide: Limitations for HPE StoreOnce

Windows Server 2012 Deduplication

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 (use parameters `/Q /L /A:64K`)
3. Follow compression and deduplication guidelines for non-integrated deduplication storage in previous chapter.
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.

Configuration Guidelines

Parallel Processing

A repository can be configured to limit the amount of parallel tasks it can process at a time; with parallel processing enabled (by default) a *task* is one VMDK handled by the proxy during a backup job, or by a repository during a backup copy job. 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 sizes

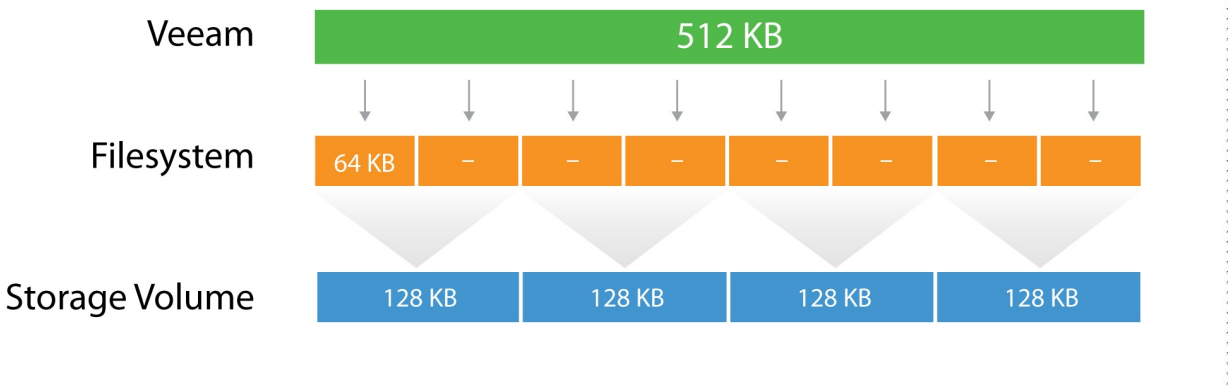
During the backup process data blocks are processed in chunks and stored inside backup files in the backup repository. You can customize the block size during the Job Configuration using the **Storage Optimization** setting of the backup job.

By default block size is set to **Local target**, which is 1 MB before compression. Since compression ratio is very often around 2x, with this block size Veeam will write around 512 KB or less to the repository per each block.

This value can be used to better configure storage arrays; especially low-end storage systems can greatly benefit from an optimized stripe size.

There are three layers where the block size can be configured: Veeam block size for the backup files, the Filesystem, and the Storage volumes.

Let's use a quick example:



The Veeam block size of 512KB is going to be written in the underlying filesytem, which has a block size of 64k. It means that one block will consume 8 blocks at the filesytem lavel, but no block will be wasted, as the two are aligned. If possible, set the block size at the filesytem layer as close as possible to the expected Veeam block size.

Then, below the filesytem there is the storage array. Even on some low-end storage systems, the block size (also called stripe size) can be configured. If possible, again, set the stripe size as close as possible to the expected Veeam block size. It's important that each layer is aligned with the others, either by using the same value (if possible) or a value that is a division of the bigger one. This limits to a minimum the the so called **write amplification**: with a 128KB block size at the storage layer, a Veeam block requires 4 I/O operations to be written. This is a 2x improvement compared for example with a 64KB stripe size.

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, as explained in the previous paragraph, the file system may also benefit from using a larger cluster size (or Allocation Unit Size). For example, during formatting of NTFS volumes, Allocation Unit Size is set to 4KB by default. To mitigate fragmentation issues, configure to 64 KB whenever possible.

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

Using "Large File" Switch for NTFS

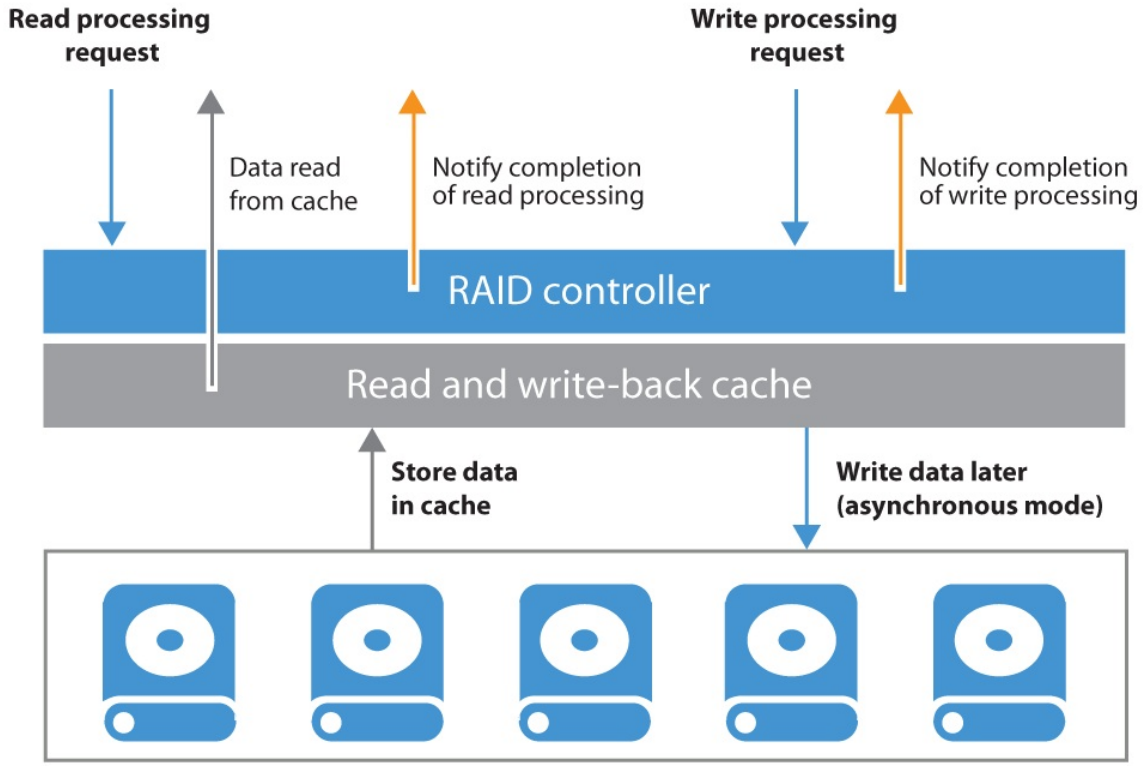
A file size limitation can be occasionally reached on NTFS, especially on Windows 2012 R2 with deduplication enabled. This happens due to a hard limit reached on the file records size because of the high level of file fragmentation. To mitigate the issue, we recommend to format Windows NTFS repositories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

Keeping File Size Under Control

Try to avoid 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 of the files themselves, it is recommended to keep backup chain size (the sum of a single full and linked 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.



Repository Sizing

In mid-sized or enterprise environments, the recommended amount of CPU for a repository is 1 core per concurrent job that processes data on a repository server. At least 2 cores allow for the Operating System to be more responsive.

It is recommended to configure 4 GB RAM per core. The same amount of resources are needed for 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 required disk space, you should know the following:

- Total size of VMs being backed up
- Frequency of backups
- Retention period for backups
- Will jobs use forward or reverse incremental

Also, when testing is not possible beforehand, you should make assumptions on compression and deduplication ratios, change rates, and other factors. The following figures are typical for most deployments; however, it is important to understand the specific environment to figure out possible exceptions:

- Data reduction thanks to Compression and Deduplication is usually 2:1 or more; it's common to see 3:1 or better, but you should always be conservative when estimating required space.
- Typical daily change rate is between 2 and 5% in a mid-size or enterprise environment; this can greatly vary among servers; some servers show much higher values. If possible, run monitoring tools like Veeam ONE to have a better understanding of the real change rate values.
- Include additional space for one-off full backups.
- Include additional space for backup chain transformation (forward forever incremental, reverse incremental) – at least the size of a full backup multiplied by 1.25x.

Note: When using deduplication appliances, please contact 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.

A repository sizing tool that can be used for estimation is available at http://vee.am/rps. Note that this tool is not officially supported by Veeam, and it should be used "as is", but it's nonetheless heavily used by Veeam Architects and regularly updated.

Tip: With Veeam Availability Suite, you can use Veeam ONE together with Veeam Backup & Replication. Among the many reports, Veeam ONE has the VM Change Rate Estimation report from the “Infrastructure Assessment” report pack; this can be used 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 growth and consequent space consumption. The report provides recommendations for adjusting 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, please refer to its Reviewer's Guide at https://www.veeam.com/documentation-guides-datasheets.html

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)

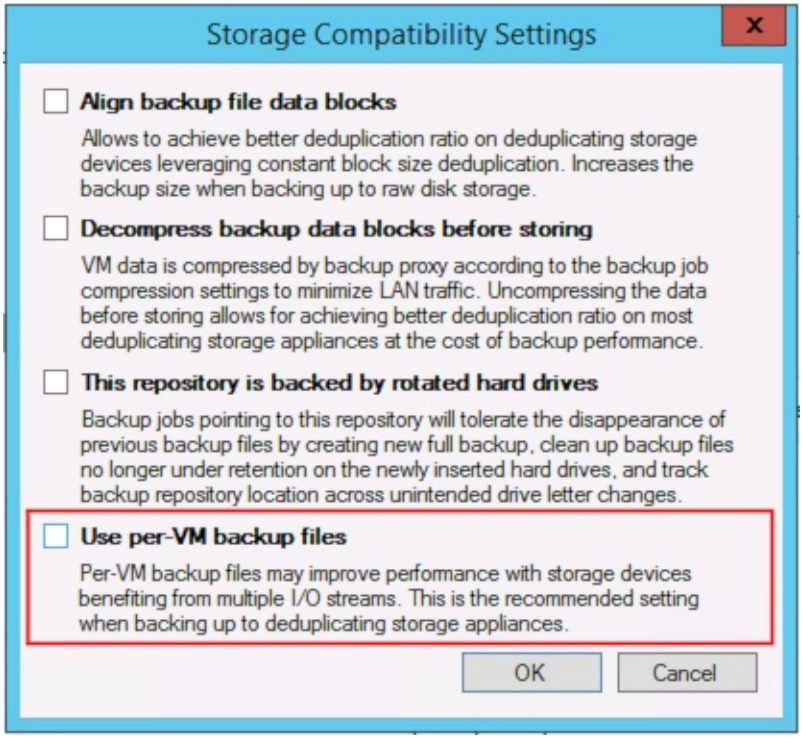
To summarize, when estimating the size of the repositories, use the following best practices:

- 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.

Per VM backup files

It is possible to write one backup file chain per each VM on a repository, compared to the regular chain holding data for all the VMs of a given job. This option greatly eases job management, allowing to create jobs containing much more VMs than in previous versions, and also enhances performance thanks to more simultaneous write streams towards a repository, even when running a single job.

In addition to optimizing write performance with additional streams to multiple files, there are other positive side effects as well. When using the forward incremental forever backup mode, you may experience improved merge performance. When backup file compacting is enabled, per VM backup files require less free space. Instead of requiring sufficient space to temporarily accommodate an additional entire full backup file, only free space equivalent to the largest VM in the job is required. Parallel processing to tape will also have increased performance, as multiple files can be written to separate tape devices simultaneously.



Per VM backup files is an advanced option available for backup repositories, and it disabled by default for new backup repositories. If enabled on an existing repository, an active full backup is required after the option has been enabled.

Maximum number of VMs per job

With per VM backup files the recommendation for number of VMs per job can be increased significantly. Even if technically jobs containing five thousands VMs have been successfully tested in a lab, feedback from the field shows the sweet spot at around 300 VMs per backup job, more for management reasons and unexpected side effects than pure performance matters. When designing your jobs, keep in mind that several operations such as synthetic operations, health checks and Backup Copy Jobs will be pending until all VMs in the job have completed successfully. For those reasons, extremely large jobs may be impractical.

Performance

To avoid counter productive effects, attention should be paid on not having too many write threads towards a storage used as a repository. For example, a low range NAS storage will probably not react very well to a high amount of parallel processes created by per VM backup files. To limit this effects, refer to Reposiory configuration options, especially the **Concurrent tasks** limit.

Scale Out Backup Repository

Veeam Scale-out Backup Repository is a logical entity made of multiple “simple” repositories, grouped together into a single abstracted object, that can be used as a target for any backup and backup copy job operation.

Scale-out Backup Repository is an extremely easy way for both medium and large customers to extend repositories when they run out of space. Instead of facing the long and cumbersome relocation of backup chains, users will be able to add a new extent (that is any of the “simple” backup repositories supported by Veeam Backup & Replication) to the existing Scale-out Repository — or group multiple repositories to create a new one.

The only requirement is the ownership of a proper license, and that at least two simple repositories have been added to Veeam Backup & Replication already.

NOTE: the default backup repository created during the installation cannot be used in a Scale-out Backup Repository as long as it’s the target of Configuration Backup, as this type of job is not supported by Scale-out Backup Repository. If the default repository needs to be added to a Scale-out Backup Repository, consider first to change the target of Configuration Backup.

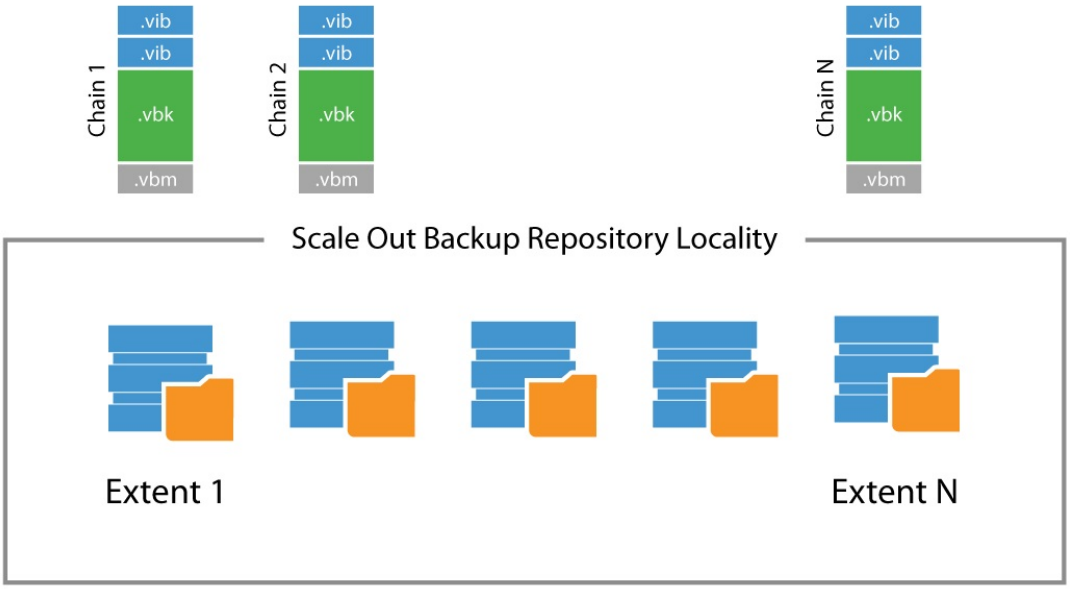
For additional technical information, the online documentation is available here : https://helpcenter.veeam.com/backup/vsphere/backup_repository_sobr.html.

File placement policies

Scale-out Backup Repository has two different options for file placement.

Data Locality

This is is the default policy, and it works by placing all the dependent files of a backup chain into the same extent. Every extent grouped with this policy has the same chances of receiving a backup chain as the algorithm treats them equally, and the major parameter for the initial placement is the free space value.



The failure domain is a single extent, as the loss of a given extent impacts only the backup chains stored into that extent. Policy can be violated by Veeam itself if, for example, one of the extents has no free space left, and the additional incremental is stored in a different extent. This because the priority is always to complete a backup or backup copy.

Performance

Performance policy places dependent incremental backup files on a different extent from the corresponding fulls. In order to choose which extent will hold the different files when using the performance policy, for each extent users are able to assign it a “role”.

Important: When using integrated deduplication devices, virtual synthetic operations may not work, if the full and incremental backup files are placed on separate extents. Please use Data Locality mode instead.

Users can configure each repository of the group to accept full backups, incremental backups or both. As soon as a new backup chain is stored into a performance Scale-out Backup Repository, the different files are place in accordance to the policy itself.

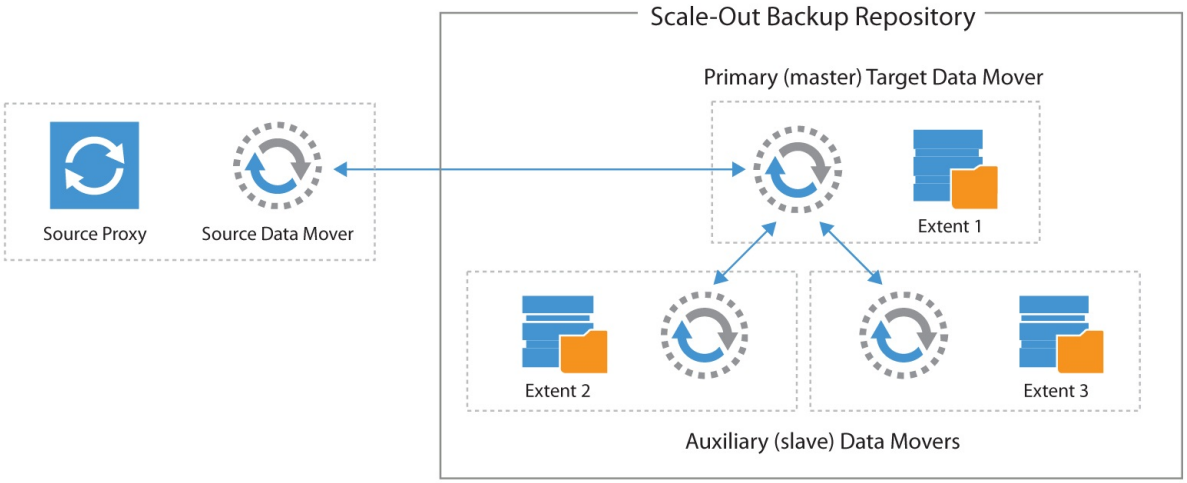
Note: in order to leverage the performance policy correctly users need to use at least two different repositories. Even if it's possible to assign both roles to the same repository, this configuration makes little sense and the best results can be obtained by splitting full backup files and incremental backup files over different physical extents.

Performance policy increases the failure domain — a backup chain is split over at least two repositories, thus the loss of one of the two corrupts the entire backup chain. This is a consideration that Veeam architects need to evaluate carefully. There is a trade-off between the increased performance guaranteed by the performance placement policy, and the increased failure domain.

Scale-out Backup repository and network considerations

Scale-out Backup Repository is, as the name implies, a scale out arhitecture, based on multiple agents, with a notion of master and slave repository datamovers.

During backups, the master agent is always started where the write is happening. During restore, the master is always started where the VBK is located, as most blocks are likely retrieved from this location.



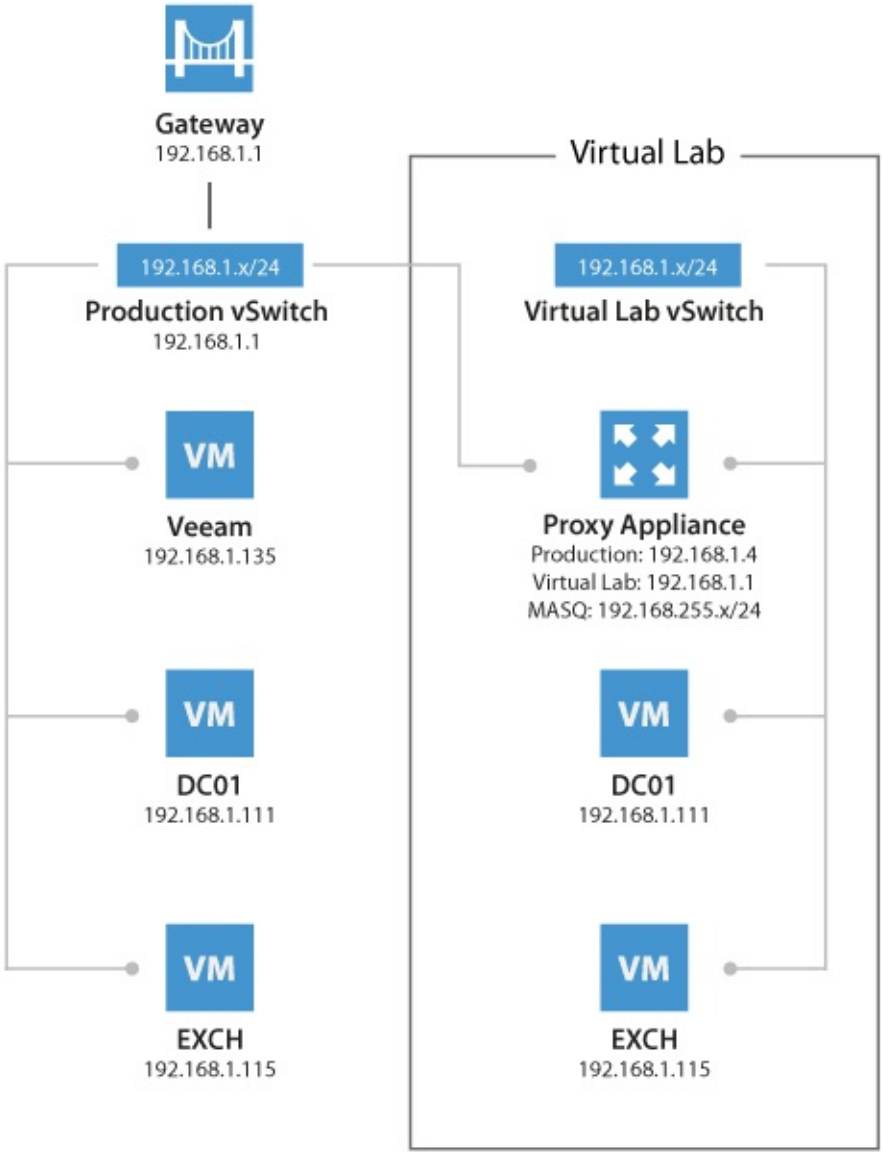
A master datamover is the only repository agent receiving data from a source datamover (a proxy in a backup job or a source repository in a backup copy job). A master datamover is able to communicate if needed with other slave datamovers to retrieve their data.

As in any scale-out solution, careful design should be applied to the network, as communications between the different data movers may increase network consumption, regardless the policy in use or the specific design of the scale-out architecture. When using Scale-out Backup Repository, 10 Gb networks are always recommended.

vPower NFS and Virtual Lab

Virtual Lab Appliance Overview

The Virtual Lab appliance operates as a gateway to offer network connectivity 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 coming from the production network using static mapping. If VMs running in the isolated network need Internet access, the Virtual Lab appliance can act as a proxy server.



When a SureBackup job is executed the static routes to reach the masquerated networks are temporarily added to the routing table on the Veeam backup server. To review the routing table, you can open a command prompt on the Veeam backup server and execute:

```
route print -4
```

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

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

To avoid network reconfiguration of physical components, place the backup server and the Virtual Lab appliance in the same network subnet.

Check Veeam Backup & Replication documentation for configuration details:

- https://www.veeam.com/veeam_backup_9_0_evaluators_guide_vpower_vsphere_en_pg.pdf
- https://helpcenter.veeam.com/backup/vsphere/verification_perform.html

How SureBackup Job Works

SureBackup leverages the capabilities of the Virtual Lab appliance to create an isolated environment where different tests can be executed against VMs. These VMs are powered on directly from the backup files using the vPower technology.

Booting the 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 masquerated networks is added dynamically to the routing table of the Veeam backup server. Those static routes define the IP address of the Virtual Lab appliance as the gateway towards the masquerated networks.

Booting Virtual Machines

1. If the Application Group is based on backups, Veeam publishes and registers VMs using Veeam vPower NFS from the repository containing the backup file. This step is skipped if the VMs are replicas.
2. Veeam reconfigures the VMs and connects them to the isolated port groups of the Virtual Lab. If a network connection is configured to be connected to a port group that is not available in the Virtual Lab, those network are disconnected automatically.
3. Veeam creates a snapshot for the VMs in order to redirect write operations to a production datastore selected during the Virtual Lab configuration.
4. If the domain controller role is selected, registry settings are injected in the VM 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 of VMs. 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 the VMs. This will slow down SureBackup jobs significantly. Therefore, it is always recommended to install VMware Tools on a verified VM.

Testing Virtual Machines

1. **VMware Tools heartbeat** is used for verifying that the VM OS is successfully started.
2. **PING** tests are initiated according to the masqueraded network configuration. The ping is sent from the Veeam backup server using the static routes added during the job execution. Since the masquerade network is not part of the Veeam backup server's own subnet, the packet is sent to the gateway matching the Virtual Lab network (usually the virtual lab appliance).
3. **Application-specific testing** uses scripts and 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. The built-in role for SQL Server provides additional testing (see next section), and custom scripts may be used for third party applications. Requests are sent from the Veeam backup server, and the routing to the virtual machine is handled by the Virtual Lab proxy appliance.

4. **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.

   Note: This feature reads the entire backup file, and requires significant time to complete.

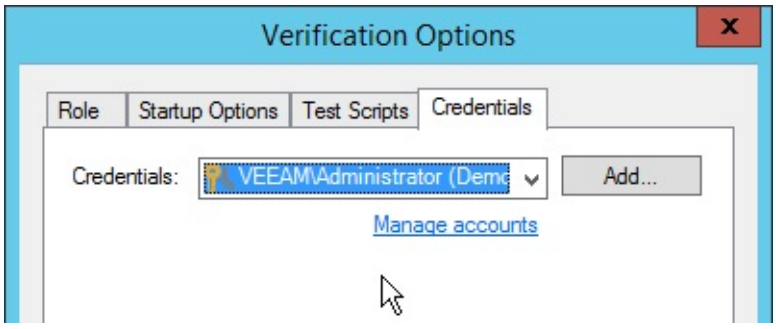
If Linked Jobs are configured for the SureBackup job, linked VMs will start booting once all virtual machines explicitly defined within the Application Group have been successfully booted and verified. Remember that by default 3 VMs are tested at the same time in a Linked Job. There may be more than 3 VMs linked, but the following ones will stay in the testing queue. The limit can be adjusted in the SureBackup job configuration wizard, and may be increased if the backup repository can handle the load accordingly.

Checking SQL Server Database Availability

A dedicated Visual Basic script is included 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 configure one or more specific instances to be tested. The script enumerates all databases and checks if these databases are available, using the `USE <db>` statement.

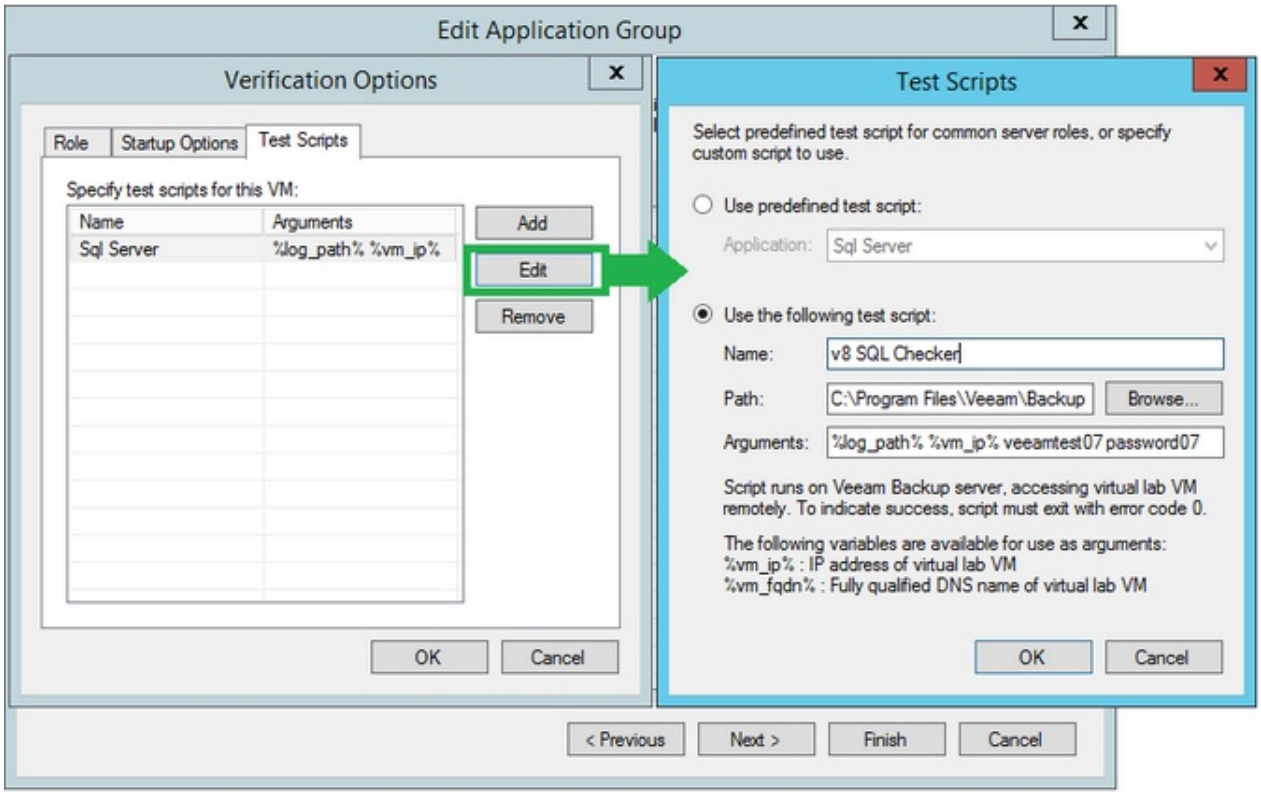
When running scripts that require authentication, when executed the script will impersonate the service account under which the Veeam Backup Service is running (default is SYSTEM). To specify different credentials configure them in the 'Credentials' tab in the Application Group settings.



Important! To ensure successful authentication it is required for the specified user to have *public* access to all databases.

The `SqlChecker.vbs` script also accepts two additional parameters to use SQL authentication instead of Windows based authentication. In order to use SQL authentication you need to add a custom test script instead of the built-in SQL Server role, and specify the following path and arguments:

- Name: SQL checker
- Path: *Browse for the* `Veeam.Backup.SqlChecker.vbs` *file*
- Arguments: `%log_path% %vm_ip% sa sa_account_password`



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. This is the procedure to do so:

1. Open the Veeam installation folder and look in 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 the generated code locally within the VM. To do so, use `\Putty\plink.exe` shipped with the product and located in 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` from the Veeam backup server in order 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 described 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. To do so:

1. Open up **Statistics** for a SureBackup job.
2. Right-click the VM you want to troubleshoot.
3. Select **Start**.

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

If the selected VM is in an application group, this VM and previous ones are started. If the VM is part of a linked job, the entire Application Group and the selected VM is started.

This mode 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: On the Virtual Lab appliance, ICMP traffic is blocked on all network interfaces connected to isolated networks, unless you check the "Allow proxy appliance to act as internet proxy for virtual machines in this lab". Unless checked, this may cause some versions of Windows Server to switch a network interface to a different network profile (i.e. from Domain to Public). This may lead to some tests failing.

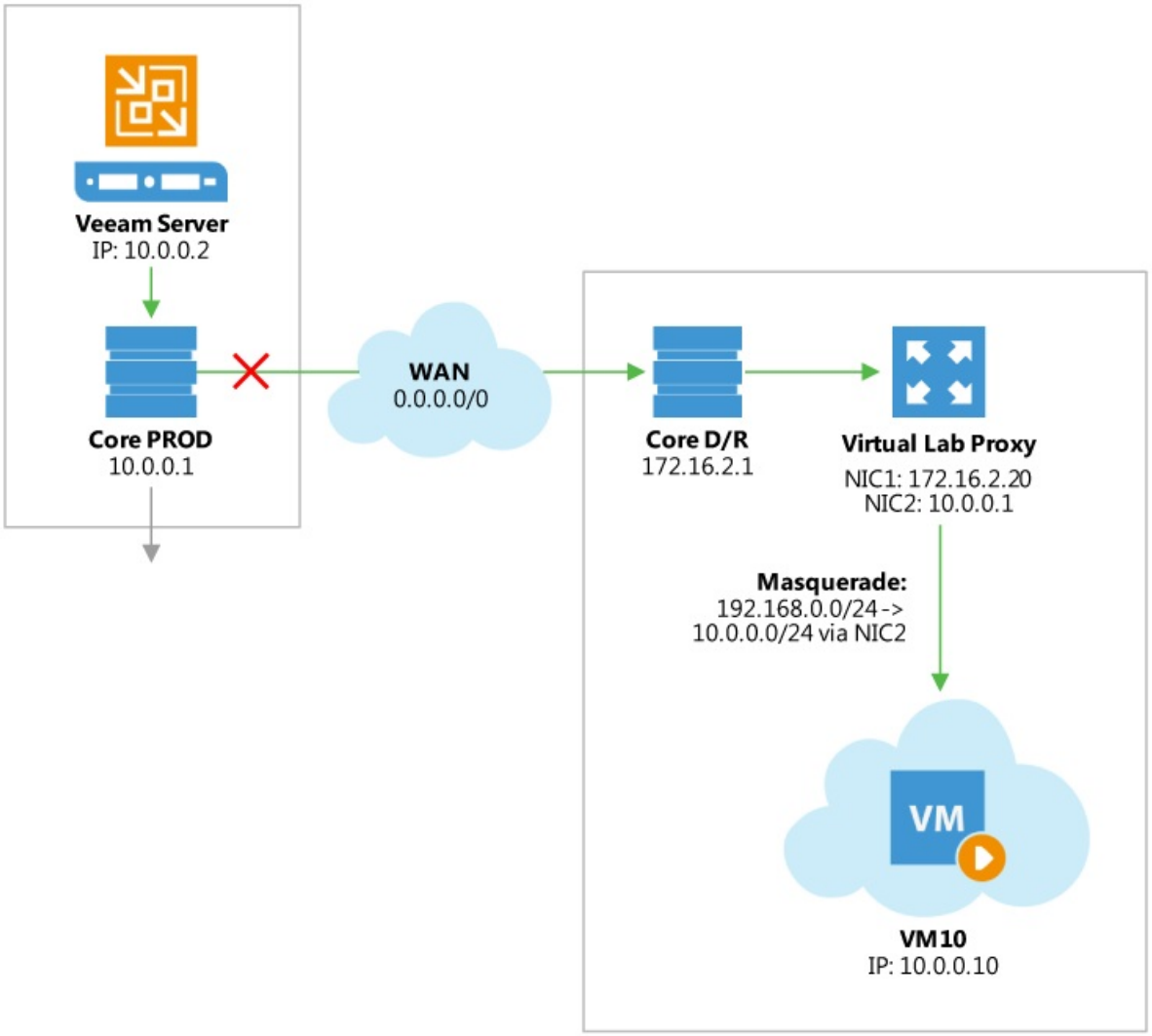
Virtual Lab in Complex Environments

When using standard vSwitches in a VMware vSphere infrastructure, the Virtual Lab proxy appliance and the isolated networks must 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 networks are not known at the core network in terms of VLAN tagging or routing.

When 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 once they are started.

Important! Please check the following help article and the links at the bottom of the webpage before you configure Virtual Labs for Distributed vSwitch: Advanced Multi-Host Virtual Labs.

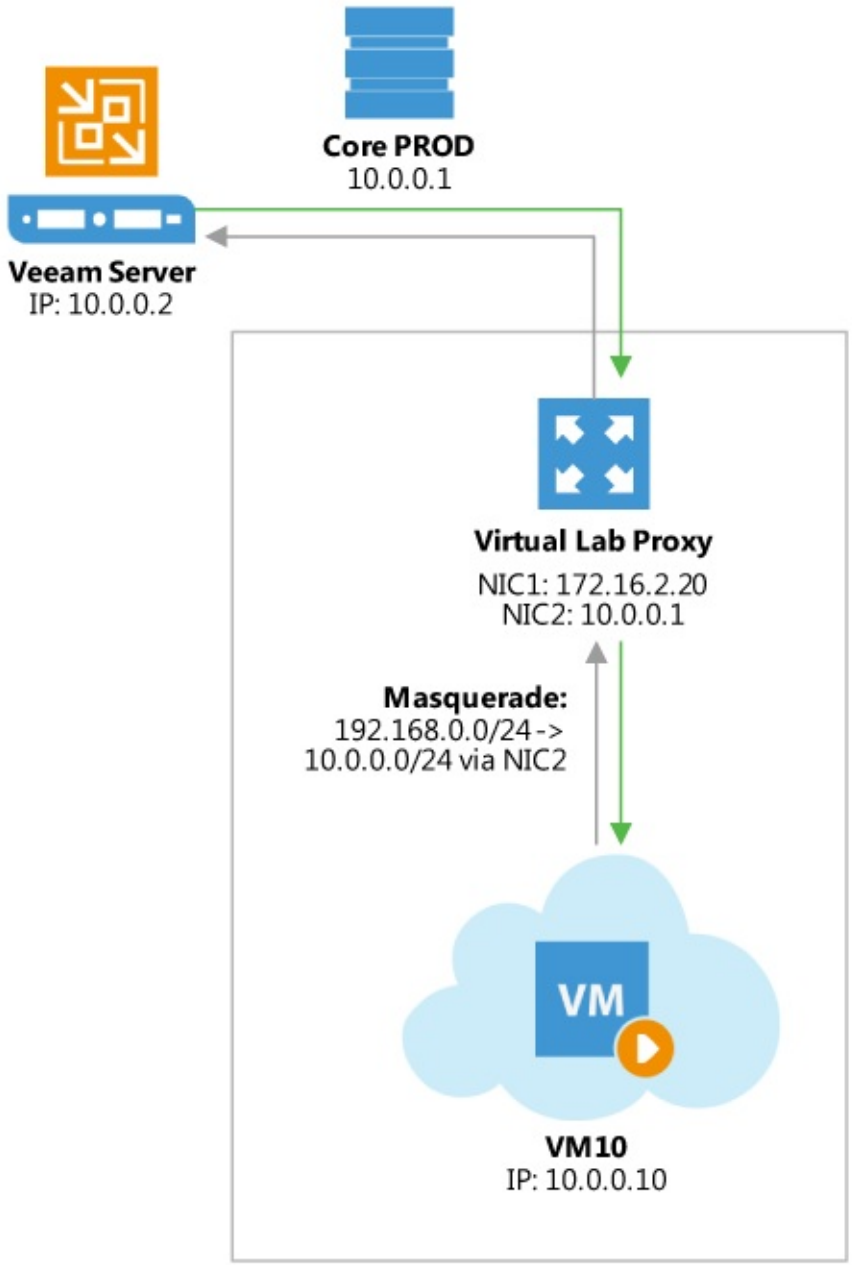
Even in environments where Distributed vSwitch is available, make sure that the Veeam backup server and the Virtual Lab proxy appliance are placed in the same subnet to prevent network packets (sent to the masquerading IP subnets) from being routed.



Most DR datacenters are configured with different IP networks from production to allow for “active-active” configurations. In such cases, layer 3 (L3) is used for networking configuration and routing is in place to establish communications between the production site and the DR site.

In these scenarios, it is recommended to deploy a Veeam backup server at the DR site. This will help getting the Virtual Lab working and ensure correct "1-Click Failover" and failback if the production site becomes unavailable.

For more information, please see the Backup Server Placement section of this guide.



WAN Acceleration

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. This technology is specifically designed to accelerate Veeam job. Any other WAN acceleration technology should be disabled for Veeam traffic.

To determine whether WAN acceleration is necessary in an environment, it is important to understand what particular savings can be achieved.

Determining Required Bandwidth

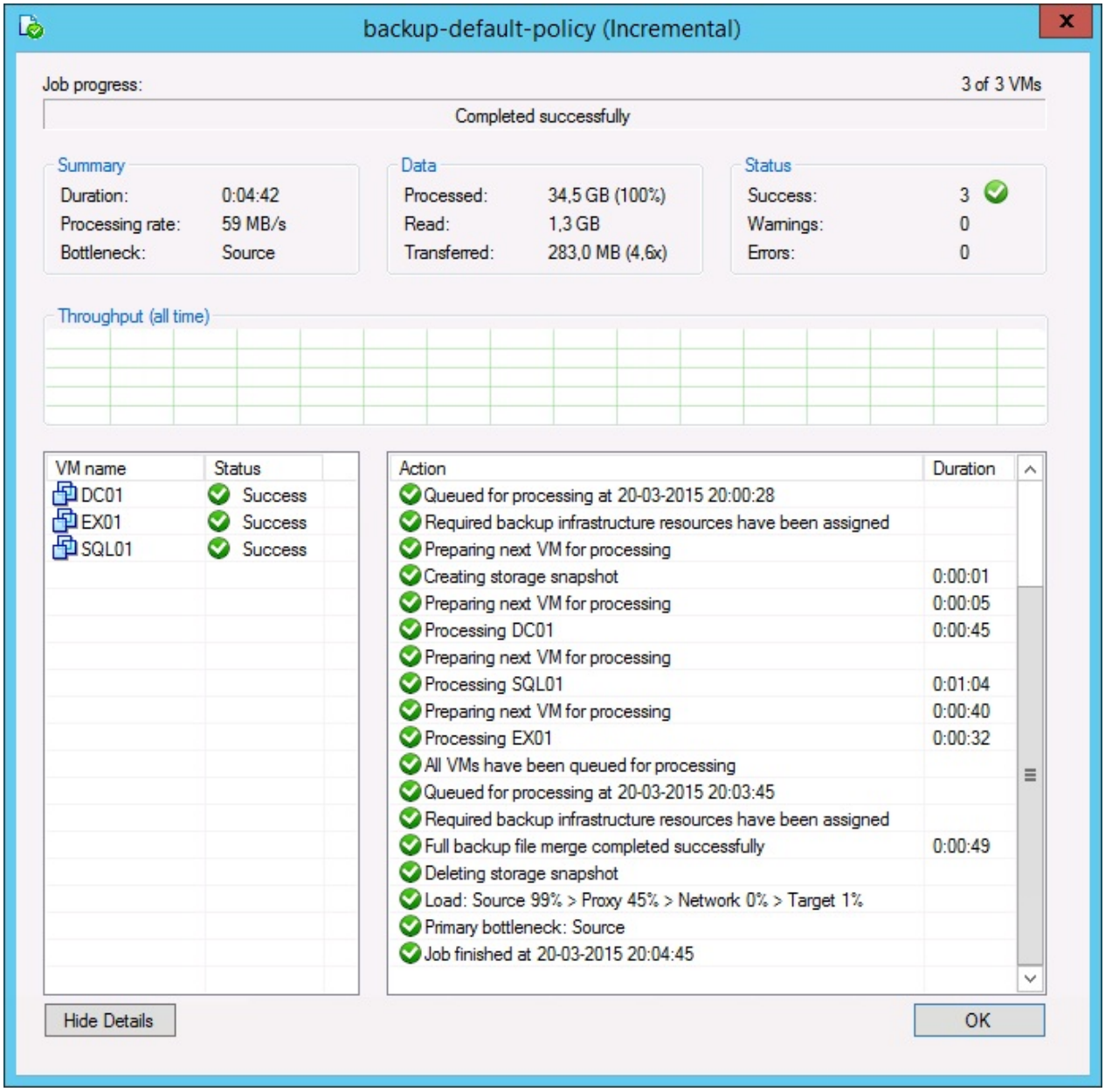
When using WAN acceleration on links with very poor bandwidth, you may have to manually seed the initial copy to the target. For more information, refer to the WAN Acceleration section of the Veeam Backup & Replication User Guide.

The WAN accelerator uses its own digests based on the hashes of the blocks inside a VM disk, which means that it reads data from the backup files and re-hydrating them on the fly, or it reads directly from the source VM in case of replication. The WAN accelerator component will then process those 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 a 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 accelerated backup copy job log and statistics.

Analyzing Backup Job

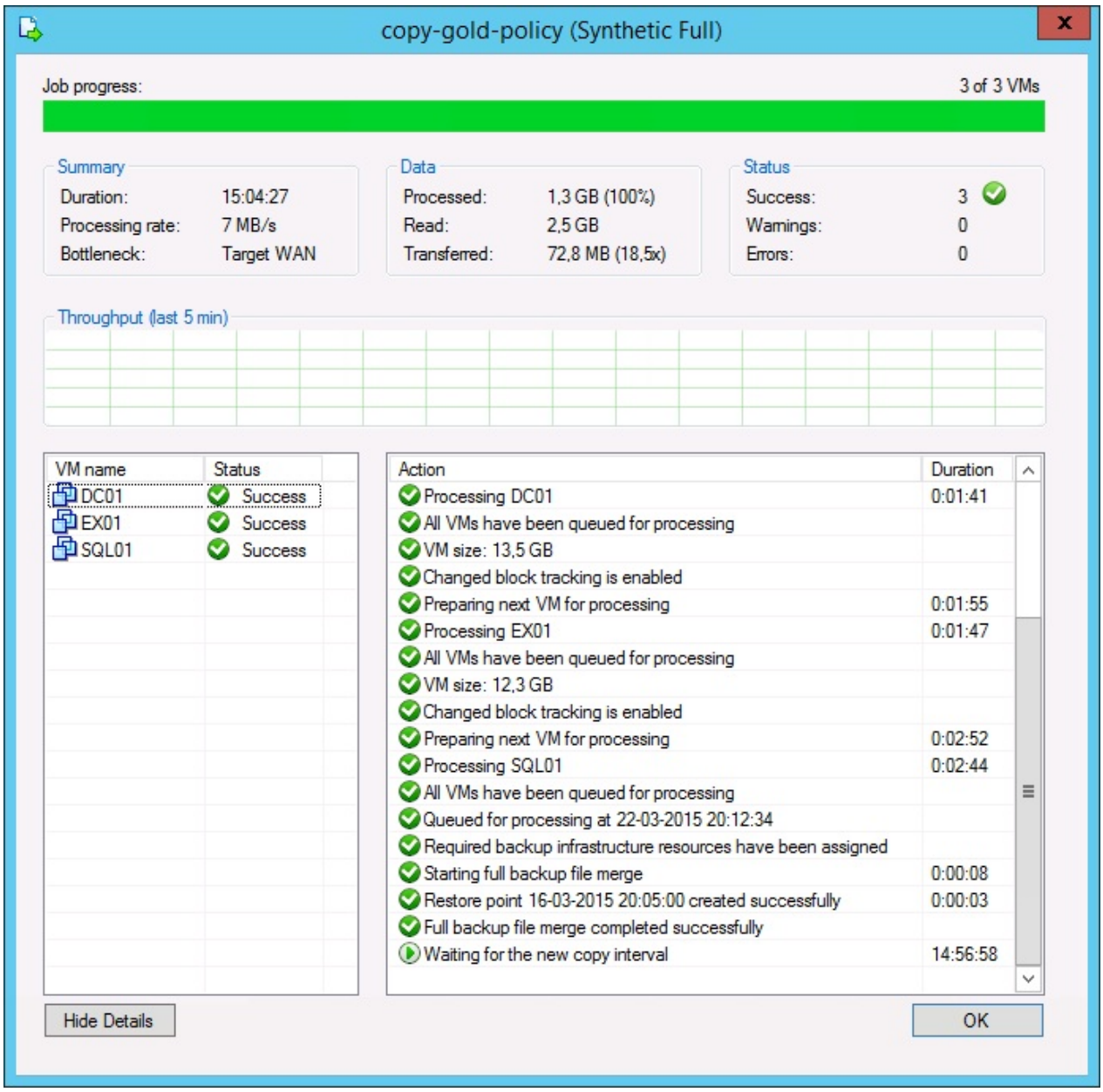
During both full and incremental job sessions, three metrics are displayed in the session data: **Processed**, **Read** and **Transferred**. To better understand the difference between direct data transfer and WAN accelerated mode, examine the **Read** and **Transferred** values:



- **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 works with a different block size compared to the fingerprinting algorithm and block size used by backup jobs that created the original backup file. For this 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

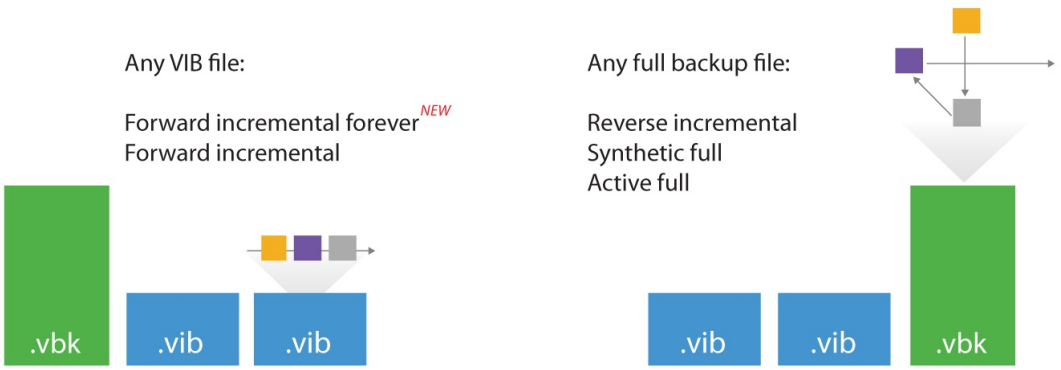
Consider that the savings rate (18.5x) displayed in the GUI is based on **Processed** data ("re-hydrated" data blocks). In the example above, 283 MB would have been transferred over the WAN link in Direct Transfer mode, while only 72.8 MB were transferred after enabling WAN acceleration. The actual savings 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 methods will make backup copy jobs work much faster, as read operations will be sequential rather than random. To avoid similar fragmentation and random I/O on forward incremental modes, keep backup storage maintenance enabled when possible.

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.

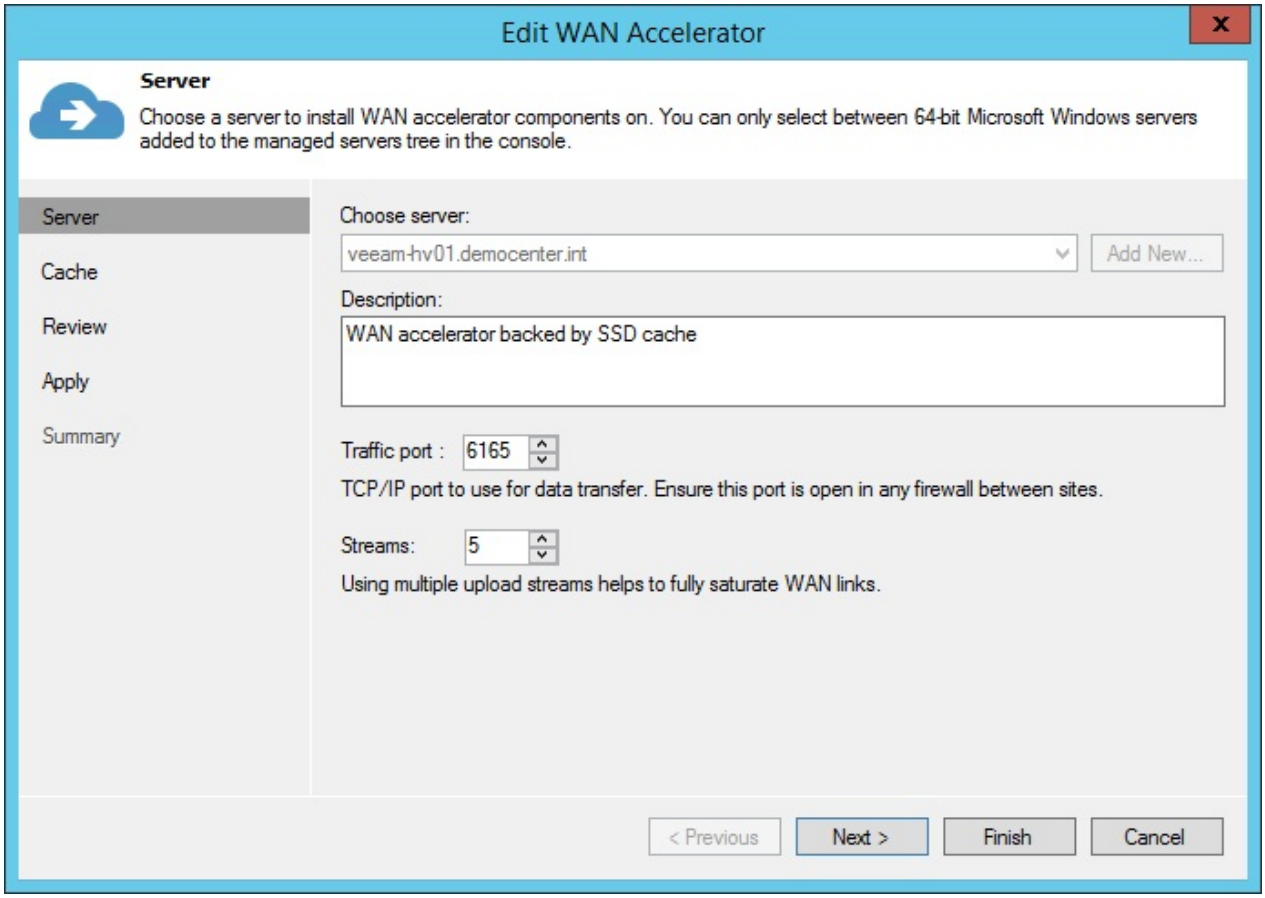
Configuration

Thanks to our friends at PernixData for helping with I/O analysis using PernixData Architect.

When configuring the WAN accelerator, not all configuration parameters affect both source and target WAN accelerators. In this section we will highlight what settings should be considered on each side.

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 only and is automatically replicated to the target WAN accelerator at the beginning of each job. This ensures different source WAN accelerators can have different settings even when using the same target WAN accelerator. The maximum setting is 100 simultaneous threads for throughput optimization and compensation for high latency or packet loss.



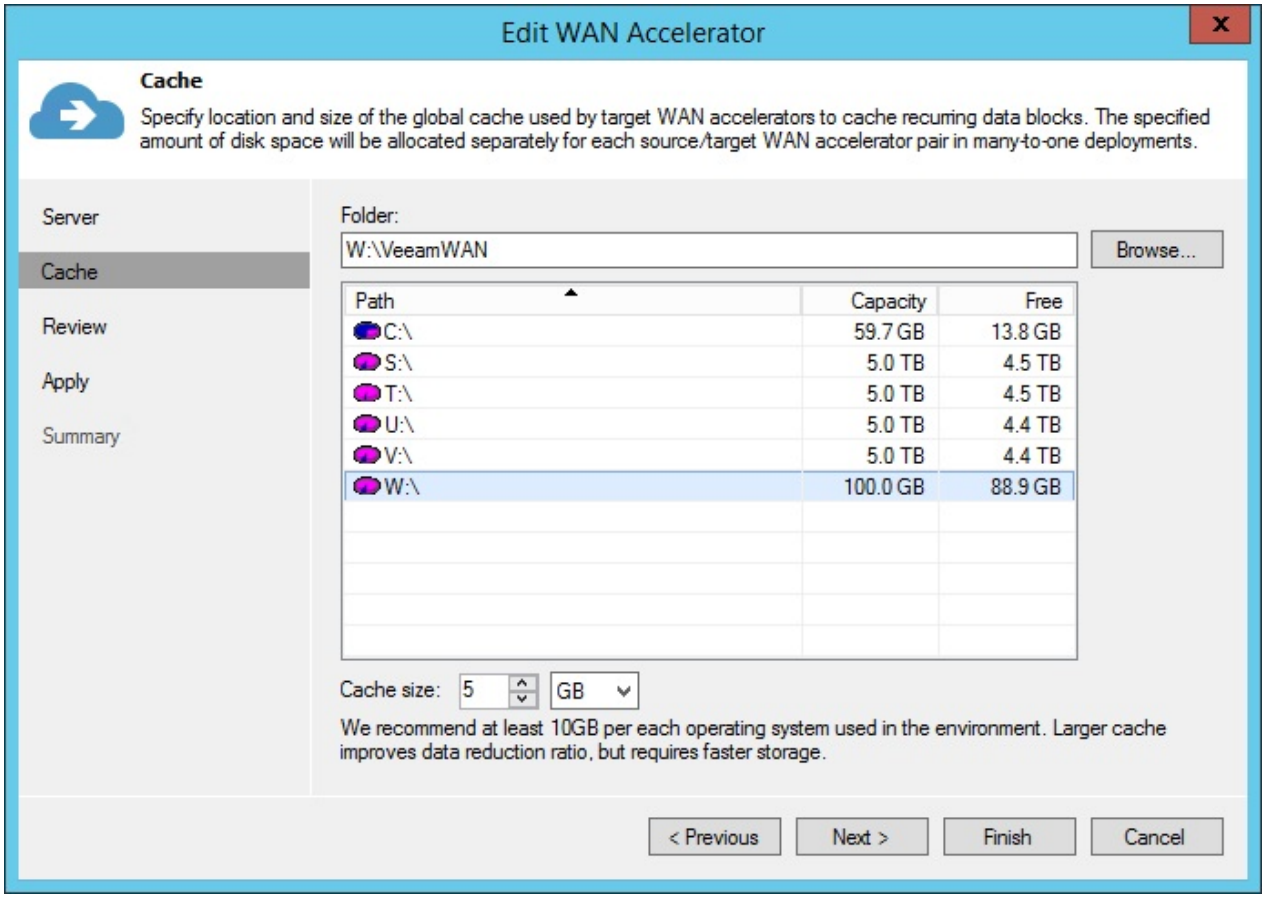
If the link has low latency and high bandwidth, the default setting (5 streams) may be enough to fully saturate it. If the link is still not saturated, the number of streams may be increased accordingly.

Testing shows that with high latency links, ***link speed x 1.5*** is a good best practice for estimating the number of streams required. Below is an example benchmark on a 10 Mbit/s WAN link with 100 milliseconds of latency.

Link (Mbit/s)	Latency (ms)	Packet loss (%)	Streams	Throughput (Mbps)
10	100	0	3	3.5
10	100	0	10	7.5
10	**100**	**0**	**15**	**10**
10	100	0	20	10

Increasing the number of streams to more than required for fully saturating the link will cause initialization of data transfers to slow down, as the data transfer will wait for all streams to initialize and stabilize before beginning transferring any data.

Tip: To test different scenarios in the lab before deploying WAN acceleration, you can use a WAN emulator (such as WANem).



When configuring the cache location for the source WAN accelerator, consider that the actual cache size on the source is irrelevant, as it is used only for digest files (where block hashes are stored). However, if a WAN accelerator will be used for bi-directional

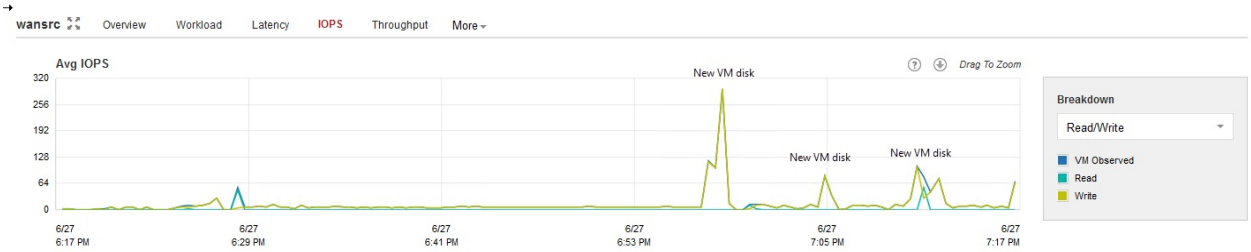
acceleration (act as both source and target), follow the guidelines provided in the "Target WAN Accelerator" section below.

Sizing

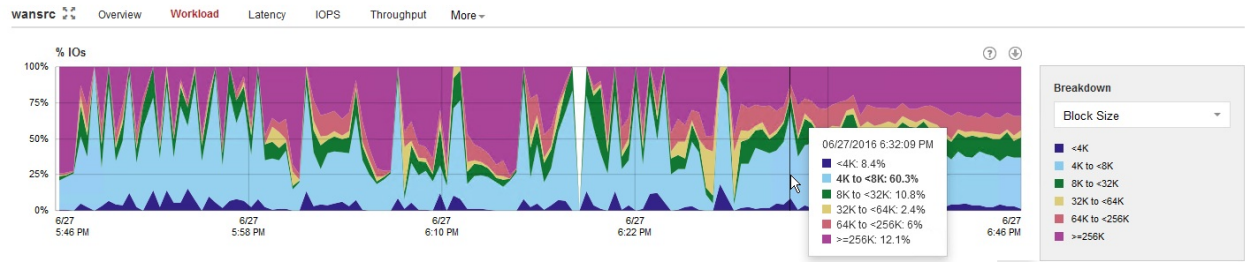
When configuring the WAN accelerator on the source side, consider that all VM disk data blocks are already in the source backup repository and they can simply be re-read from the source repository when needed. This is the reason why configuring the cache size on a source WAN accelerator does not matter. It is never used for caching any data. However, there are other files residing in the source WAN accelerator folder, and the file structure will be described in the following sections.

Hardware

The source WAN accelerator will consume a high amount of CPU whilst re-applying the WAN optimized compression algorithm. Recommended system configuration is 4 CPU cores and 8 GB RAM.



The I/O requirements for the source WAN accelerator spikes every time a new VM disk starts processing. Thus, it is recommended to deploy WAN accelerators on disk configurations with decent I/O performance.



The typical I/O pattern is made of many small blocks, so using high latency spinning disks is not recommended.

Disk Size

Each digest file consumes up to 2% of its source VM disk size. This means, for example, that a 2 TB VM disk file can produce a digests file up to 40 GB in size.

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

- Formula: `(<Source data size in GB> * 2%) + 10 GB`
- Example with 2 TB source data: `(2,000 GB * 2 %) + 10 GB = 50 GB`

For understanding how disk space is consumed, please see the following sections.

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.

VeeamWAN\GlobalCache\src

Only a `data.veeamdrf` file is located in the `\VeeamWAN\GlobalCache\src` folder. This file will be synchronized *from* 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

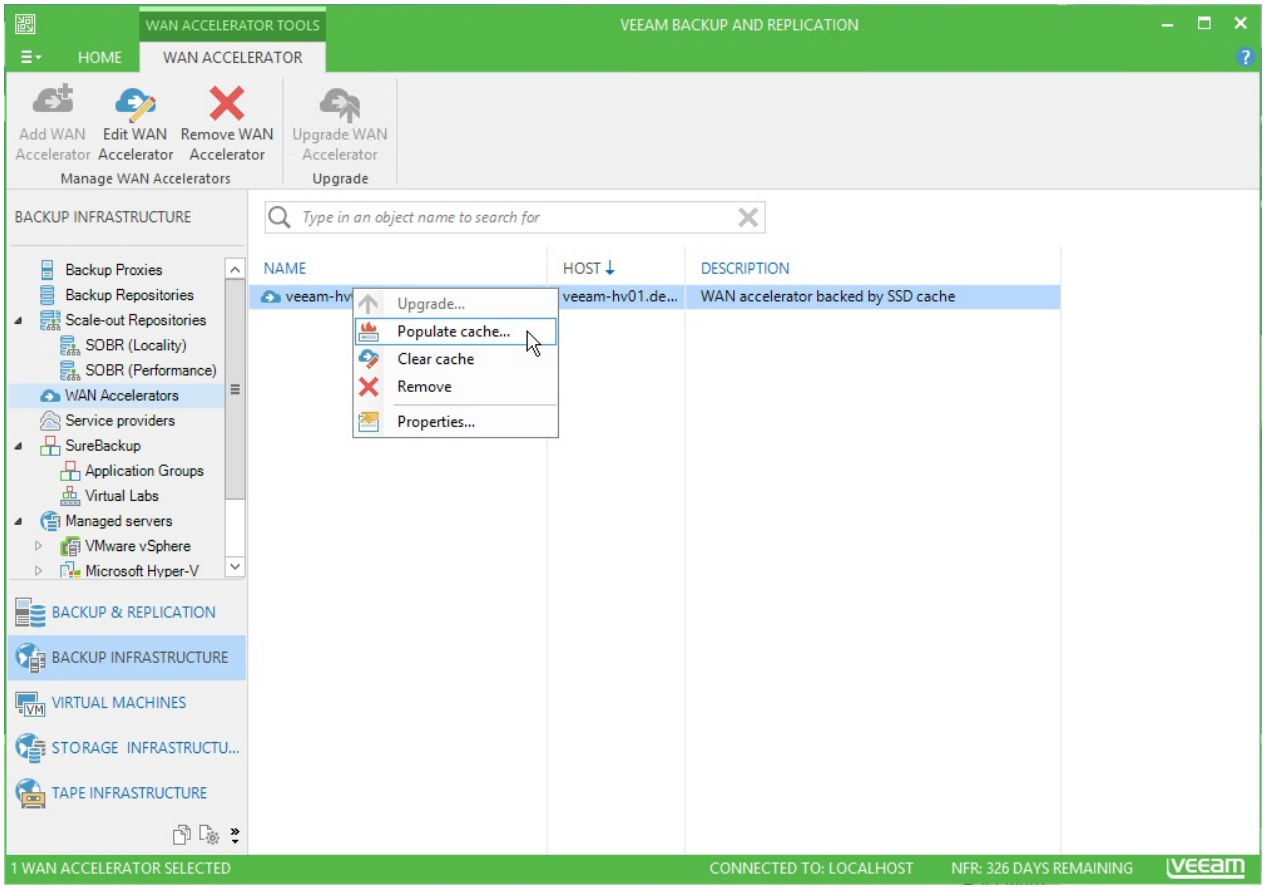
On the source WAN accelerator there are the VM disk digests that take up disk space. For each processed VM disk, a disk digest file is created and placed in `\VeeamWAN\Digests\<JobId>_<VMId>_<DiskId>_<RestorePointID>`.

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

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[1] configured for each target WAN accelerator.
- It is recommended to configure the cache size at 10 GB for each operating system[2] processed by the WAN accelerator.
- Once the target WAN accelerator is deployed, it is recommended to use the cache population feature (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.

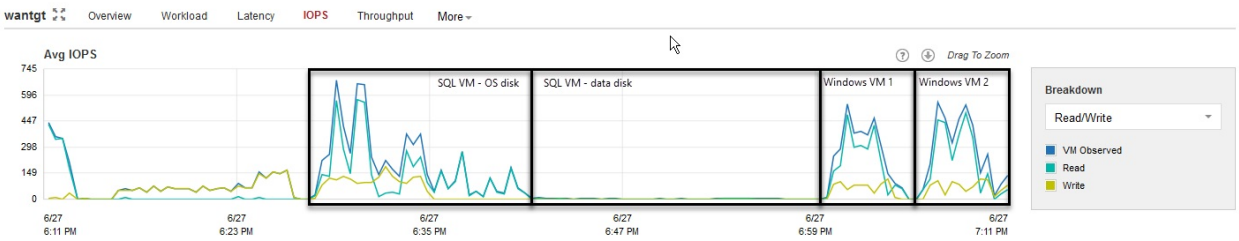


Sizing

Hardware

Although a target WAN accelerator will consume less CPU resources than the source, the I/O requirements for the target side are higher.

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). As described in the user guide, the cache is active on operating system data blocks, while other data blocks are being processed only with the WAN optimized data reduction algorithm (in-line compression).



Tests show that there are no significant performance differences in using spinning disk drives as storage for the target WAN accelerator cache rather than flash storage. However, when multiple source WAN accelerators are connected to a single target WAN accelerator

(many-to-one deployment), it is recommended to use SSD or equivalent storage for the target cache, as the I/O is now the sum of all the difference sources.

Disk Size

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

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

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

If the cache is pre-populated, an additional temporary cache is created. The temporary cache will be converted into being the cache used for the first connected source. Subsequently connected sources will duplicate the cache of the first pair. As caches are duplicated the configured cache size is considered **per pair** of WAN accelerators.

Formulas:

- Formula for configured cache size (insert this number in configuration wizard):
  - `(Number of operating systems * 10 GB) + 20 GB`
- Formula for used disk space:
  - `(Number of sources * <formula for configured cache size>)`

Examples:

- Example with one source and 2 operating systems:
  - Configured cache size: `(2 operating systems * 10 GB) + 20 GB = 40 GB`
  - Used disk space: `(1 source * 40 GB) = 40 GB`
- Example with five sources and 4 operating systems:
  - Configured cache size: `(4 operating systems * 10 GB) + 20 GB = 60 GB`
  - Used disk space: `(5 sources * 60 GB) = 300 GB`

For understanding how the disk space is consumed, please see the following sections.

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 `blob.bin` file containing the cache will be located. That file size corresponds to the setting configured in the management console.

Note: The `blob.bin` 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.

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.

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 to fully saturate the WAN link.

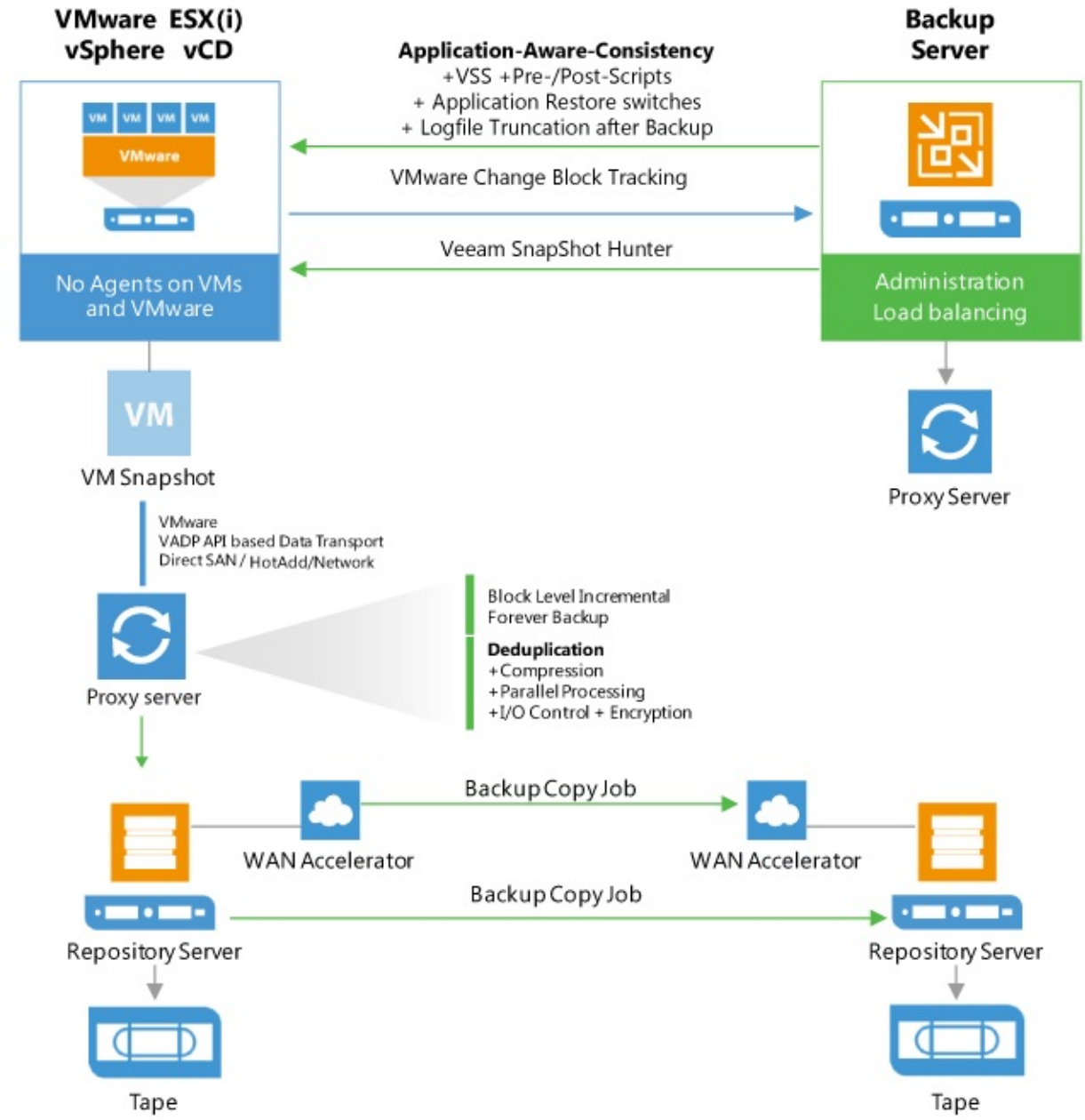
> [1]. A pair of WAN accelerators means any source WAN accelerator paired with the target WAN accelerator. ↩

> [2]. All Linux operating systems are considered as one in terms of WAN accelerator sizing. ↩

Tape Support

Overview

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



Tape Device Connection

The following configuration prerequisites must be met:

- All connection types require driver installation
- 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
- StorageTek ACSLS is not supported while a direct connection to the library is
- Dynamic Drive Sharing is not supported
- Library Partitioning is supported
- Multiple control paths are supported only when control path failover and MPIO is configured correctly. Please contact the vendor for more information

Connection Type	Compatibility
FC/SAS/SCSI/FCoE/Infiniband/iSCSI or other block technology to physical Tape Proxy	Supported with Windows driver as long as the tape vendor supports the connection. ("Unknown media changer" support for FC/SAS and VTLs)
FC/SAS redirect to VMware VM	Unsupported
FC/SAS redirect to Hyper-V VM	Unsupported
FC/SAS to iSCSI Converter/Bridge	Supported
Starwind Tape Redirector	Supported

Tape device support

While the system requirements dictates what tape devices are technically supported, there is a community validated list available on the Veeam forums: Unofficial tape device compatibility list

Supported

- LTO-3 or higher
- For VTLs, see the corresponding section under Deduplication Storage

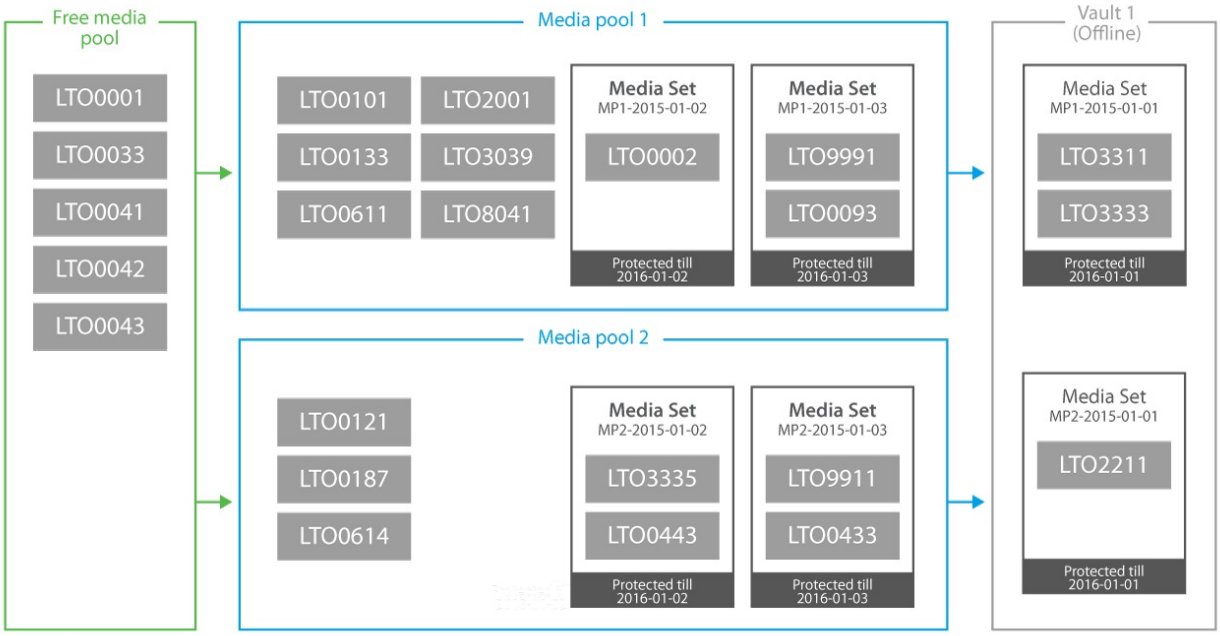
Not supported

- IBM "Jaguar" TS11x0 Enterprise tape drives
- StorageTek T10000 tape drives
- Older Tape drives like DLT or AIT

Drivers

- **IBM drivers**: use “non-exclusive” driver setup and start it with admininstrative rights.
- **HP drivers**: these are not installable with the downloadad install **.exe** file on a VM (for example, to use with VTL). As a solution, run the install **.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 Management



Media Pool

A media pool simply defines a group of tapes managed by Veeam Veeam Backup & Replication. There are three types of media pools:

- **Service media pools**. Created and managed automatically. It is not possible to modify their settings. They contains:
  - 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.

    - After inventory or cataloging, media with existing data is placed into the **Imported pool**. Review the contents and place such media into the **Free** pool for overwrite or leave in **Imported pool** to keep the data.
    - Exhausted or broken tapes are placed into the **Retired pool** and are not used further.

- **Media pools** 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.
    - Since v9 a (Custom) Tape Pool can be spanned over multiple tape libraries. The idea is to use the capacity and drives of multiple tape systems together and to failover to another tape library in case one library goes offline.
- **GFS media pools** are used to store weekly, monthly, quarterly and yearly backups on tape.
    - You can create as many GFS tape pools as needed.
    - Media can be assigned to a pool manually, or configured to be automatically assigned from the free pool. As well optional can define specific tapes for specific media sets (for example yearly backups).
    - 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). It is also possible to reuse the same media set forever. When 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.

Backup Modes

Backup jobs can create different backup types of backup file chains on disk depending on the backup mode used. Depending on backup mode, "Backup to Tape" jobs either copies files to tape or synthesize a full backup. The following rules apply:

- When archiving reverse incremental backups, the behavior varies on the type of media pool used:
  - **Standard Media Pool:** The tape job will always copy the full backup and ignore any rollback files (VRB)
  - **GFS Media Pool:** The tape job can synthesize a full backup from VRB files on specified day(s)
- When archiving forward incremental backups, *with* active or synthetic full scheduled, the backup chain on tape will be a copy of the backup chain on disk. The synthetic full option in tape job configuration is ignored.
- If you archive forward incremental backups without synthetic or active full enabled, or archive Backup Copy Jobs, the full files are synthesized from existing restore points on disk. The synthetic full backup schedule can be configured on the "Backup to Tape" job. For more

  ```
  information about virtual full to tape, please see
  [Veeam Help Center](https://helpcenter.veeam.com/backup/vsphere/virtual_full_backu
  p.html).
  ```

If the source backup job contains multiple chains, and the checkbox "Process latest full backup chain only" in advanced job settings is unchecked, you will be prompted for a decision, when creating a Backup to Tape job. You may choose to either only the last backup chain or all existing restore points.

Sizing

For the highest throughput, enabling parallel processing for the Backup to Tape is recommended. You need to size the servers and storage connection accordingly. It can be helpful to create multiple partitions with 2-4 tape drives and add these partitions to different tape servers. Adding these libraries to the media pool and enabling parallel processing will distribute the load across multiple drives and tape servers.

Note: Parallel processing for tape is unavailable for GFS media pools.

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

Perform a POC to test throughput of tape and disk. If you have no opportunity to test speed, assume that the lowest speed for backup to tape jobs with LTO5/6 is 50MB/s as a conservative estimate. We highly recommend to do a POC to evaluate real throughput to avoid additional hardware costs.

The GFS (Grandfather, Father, Son) tape job can help avoid a complex Backup to Tape job creation by handling weekly, monthly, quarterly and yearly backups in a single job.

For Backup to Tape jobs that use forward incremental (without synthetic or active fulls) jobs or Backup Copy Jobs as source of the data, it may be required to temporarily disable the job using pre- and post scripts, as the transform process of forever incremental forever will terminate the tape job. Another option is to increase the restore points of these jobs temporarily. By increasing the number of restore points for the source job, the Backups to Tape job will not be terminated by the merge process. However, please note this will increase the transform time significantly once the setting is reverted and is highly discouraged for large jobs. An example of this implementation can be found here: v9 GFS job - No more continuous?

Using 3rd party 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. This approach can also support enterprise customer "Segregation of duty" demands as 2 complete different teams can handle backups and tape

backups. No single person can delete by mistake or on purpose the primary and tape chain. Before having two backup solutions co-exist on the same server, please verify they do not conflict each other.

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 based encryption is not supported by the tape device, software based AES-256 encryption is used. Please note software based encryption may cause significant performance degradation, if not natively accelerated by the CPU of the tape server.

Hardware based encryption is typically available for LTO-4 or newer libraries, and while a license is often required, this is usually supported for free by the tape library vendor.

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 Protection in Enterprise Manager. See the 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.

For additional details on tape encryption, see the corresponding section of this guide > Encryption

Tips

- "Short Erase" all tapes before use with Veeam to avoid any problems cause by data from other backup software
- Install latest Windows Updates
- Install latest firmware on library, drives, HBA (verify interoperability)
- Install separate HBAs for tape is recommended, but not required
- A staging area for backup files is required when restoring from tape. Keep this in mind when sizing backup repositories.
- Tape compression should be disabled for tape jobs, when backup files are already compressed at the backup repository

- "File to Tape" engine is optimized for smaller amount of big files (e.g. backup files) only

Veeam Explorers

Veeam Explorers are tools included in all editions for item-level recovery from several application. As of v9, following Explorers are available:

- Veeam Explorer for Active Directory
- Veeam Explorer for SQL Server
- Veeam Explorer for Exchange
- Veeam Explorer for SharePoint
- Veeam Explorer for Oracle
- Veeam Explorer for Storage Snapshots

Each Explorer has a corresponding user guide available in Helpcenter: Veeam Backup Explorers User Guide. This section will focus primarily on the planning required for SQL Server and Oracle restores. For these Explorers, special attention should be paid to planning network connectivity between mount server and staging servers in restricted environments. Ports used for communication between them are listed in the corresponding section of the User Guide (https://helpcenter.veeam.com/backup/vsphere/used_ports.html#explorers).

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.

Explorer for Active Directory

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

Explorer for SQL Server

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.

One special case of custom settings that must be configured on staging server is encryption. When performing restore/export of encrypted database please refer to KB2006 for details on configuring the staging server.

Explorer for Oracle

Oracle restore uses a combination of executing commands via SSH or RPC depending on the platform, and using the RMAN client. VM disks are mounted to target server using iSCSI (Windows) or FUSE and loop device (Linux). Only database files will be restored, not instance files. Instance files may be recovered through file-level recovery if needed.

When backing up Oracle on Linux, the backup server is used for initiating connections, whereas a Guest Interaction Proxy will be selected for Oracle on Windows.

Ensure the account used to connect to target/staging server has enough permissions on operating system and database as described in the corresponding section of User Guide or Application-aware Image Processing section of this guide.

Note: When restoring to Linux ensure that account used to connect to restore target server has valid shell.

Restore workflow

When performing restore Veeam follows the following steps:

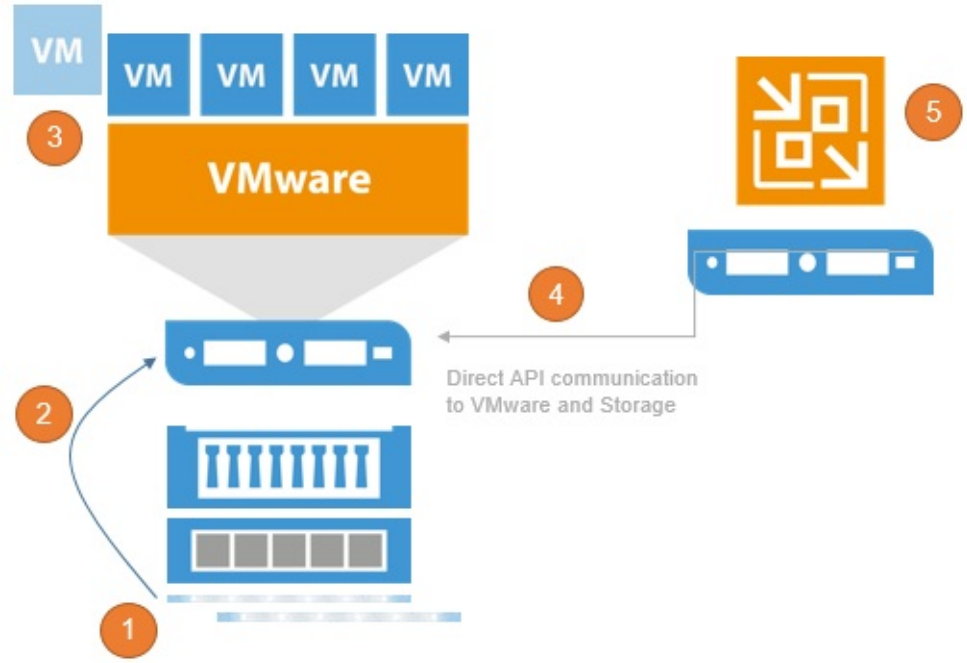
1. Oracle instance/database discovery is performed and information is collected, that includes path validation and sisk space availability checks.
2. VM disks are mounted.
3. Target database is shut down and dropped, configuration is cleaned (configuration and temporary instance files).
4. Database is started from the temporary location, if that fails another restore attempt is performed with safe set of parameters.
5. After successful test start from temporary location database is restored to proper location using automatically generated RMAN script.
6. Restore control files are restored after that. Database is updated to specific transaction prior to that in case point in time was selected for restore.
7. Fast Recovery Area parameters are restored and database is upgraded accordingly if restoring 32-bit instance to 64-bit.
8. To finalize restore mounted backup is removed from RMAN repository, restored database is restarted and new DB ID is generated. Remaining bits of the configuration are restored as well - parameter file is restored to proper path along with password file, DBNAME is changed if needed, logs are reset and online logs are recreated.

Explorer for Storage Snapshots

Veeam Explorer for Storage Snapshots (VESS) is included, but it is related to storage integrations with primary storage. This is explained in the Backup from Storage Snapshots section of this guide.

VESS is a very easy way to perform item-level recovery directly from storage snapshots. Veeam is able to use discover and mount any storage snapshot for restores. By combining the Veeam application consistent with crash consistent snapshots, the RPO for certain applications can be significantly reduced.

When opening VESS, the following workflow kicks off:



1. Creating a Clone of the Snapshot to make it writeable
2. In case of Block access (iSCSI, FC, FCoE) mount the new LUN to a proxy ESXi and register a temporary datastore, in case of NFS access the existing NFS datastore and look for the cloned VM
3. Register the temporary VM within the VMware inventory
4. Access the VM using the VMware API
5. Show the content as a Veeam Explorer to restore

After restoring and exiting VESS, the temporary datastore, VM and LUN clones will be rolled back and cleaned up.

Interaction with vSphere

Veeam Backup & Replication relies heavily on the vSphere infrastructure it is protecting. 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 must 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 what 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 limitations. Keep in mind, that the maximum file size include all snapshot files and the data disk in total. For example if you have an old VMFS version 3 the maximum file size (including snapshots) is 2TB and so your data disk should not be sized over 1.98TB to still be able to create snapshots. For details, see VMware KB article 1012384.

The default number of concurrently open snapshots per datastore in Veeam Backup & Replication 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

Please note that enabling Storage Latency Control will override the registry setting, as the snapshot threshold will instead adjust itself dynamically according to current storage latency.

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" for older versions and the same method as storage vMotion uses starting from vSphere 6.0u1 to minimize the impact and duration of the stun, as described below:

For vSphere 6u1 and newer: The host leverages the Storage vMotion Mirror driver to copy all needed data to the original data disks. When completed, a "Fast Suspend" and "Fast Resume" is performed (comparable with vMotion) to bring the original data files online.

For older vSphere Versions (Rolling Snapshot):

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:

- Upgrade to vSphere 6u1 or newer to use the new Storage vMotion based Snapshot commit processing.
- **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 can keep long-running jobs from running during production. See the corresponding setting on the **Schedule** tab of the job wizard
- **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.x 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 large datastores (up to 62 TB). You can increase the warning threshold in the backup server options, of 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: REG_DWORD
  - Default value (in GB): 2

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

- **Enable parallel processing.** Parallel processing tries to backup multiple VM disks that belong to a single VM at the same time. This reduces snapshot lifetime to the minimum. This option is enabled by default. Please note 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, when the **virtual appliance (hot-add)** transport mode is used. Hot-add is takes priority in the intelligent load balancer, when Backup from Storage Snapshots or Direct NFS are unavailable.

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 Direct NFS as part of Veeam Direct Storage Access processing modes and Veeam Backup from Storage Snapshots on NetApp NFS datastores. We highly recommend you to use one of these 2 backup modes to avoid problems.

If for what ever reason Direct NFS processing can not be used and HotAdd is configured, 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: REG_DWORD
- Default value: 0 *(disabled)*

  Value = 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.

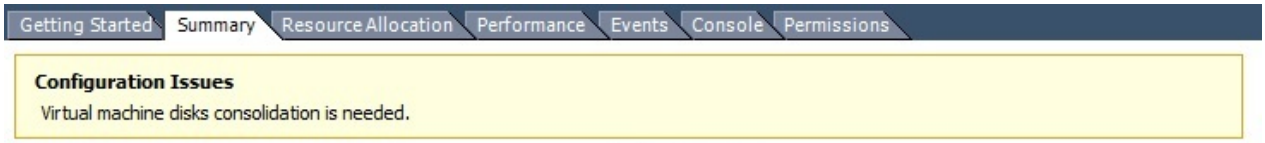
  Value = 2 - when proxy A is available on the same host, Veeam Backup & Replication will leverage it. If proxy A is busy, Veeam Backup & Replication will wait for its availability; if it becomes unreachable for some reason, Veeam Backup & Replication will switch to NBD mode.

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) if Direct NFS backup mode can not be used.

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 any VM 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 Storage Latency 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 Veeam Backup & 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.

Storage Latency Control

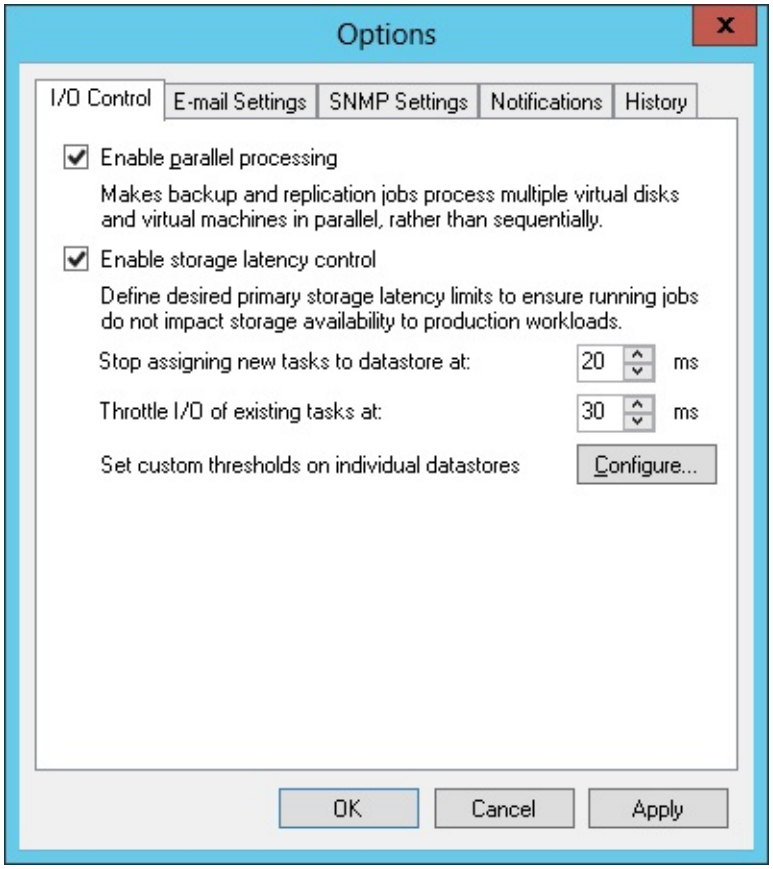
One question that often arises during the development of a solid availability design is how many proxy servers should be deployed. There must be a balance between the production infrastructure performance (as you must avoid overloading production storage), and completing backup jobs in time.

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. Storage Latency Control was implemented to help avoid this issue.

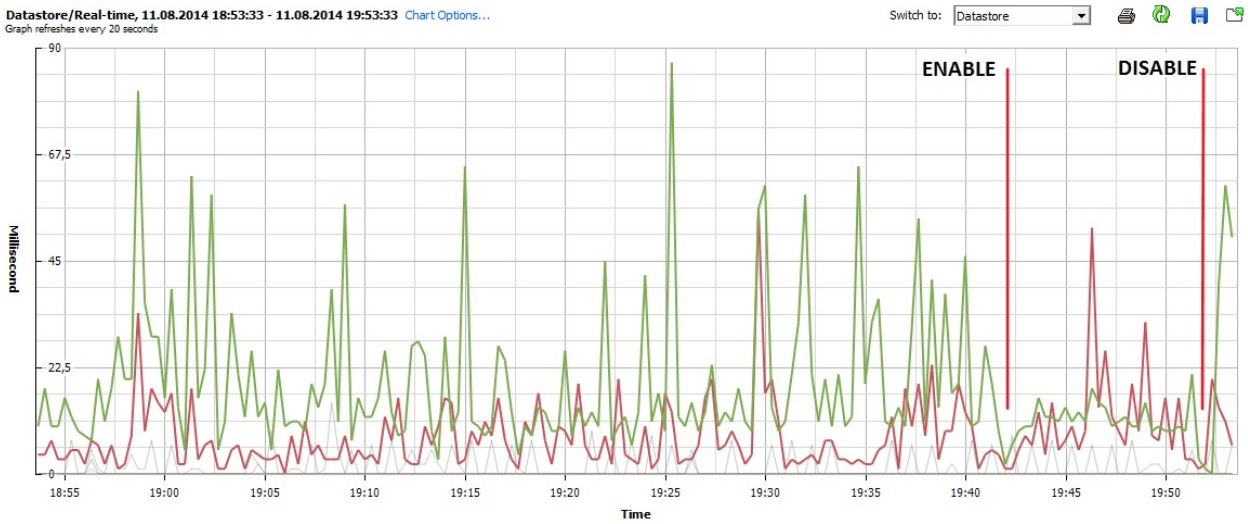
When Storage Latency Control is enabled, it monitors the storage read latency on the production datastores using real-time metrics from the hypervisor. By default, metrics from the hypervisor are collected every 20 seconds. These settings are inherited from vSphere.

The first Storage Latency 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, Storage Latency Control will throttle the throughput for existing tasks according to the second threshold **Throttle I/O of existing tasks at**.



The results of enabling Storage Latency Control are very easy to review using the vSphere Client.



When to Use?

Storage Latency Control provides a smart way to extend backup windows or even eliminate backup windows, and run data protection operations during production hours.

When Storage Latency 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 Storage Latency 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 Storage Latency Control.

Storage Latency 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 or if you use Network Transport mode (NBD) 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.

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.

You can also optimize the ESXi network (NBD) performance by increasing the NFC buffer size from 16384 to 32768 MB (or conservatively higher) and reducing the cache flush interval from 30s to 20s. For details how to do this, see VMware KB article 2052302. After increaing NFC buffer setting, you can increase the following Veeam Registry setting to add addition Veeam NBD connections:

- Path: `HKLM\SOFTWARE\VeeaM\Veeam Backup and Replication`
- Key: `ViHostConcurrentNfcConnections`
- Type: REG_DWORD
- Default value: 7 *(disabled)*

Be careful with this setting. If the buffer vs. NFC Connection ratio is too aggressive, jobs may fail.

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. Specifically if you access the vCenter over a WAN link such scoping can reduce the (management background) WAN traffic.

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

Job Configuration

In the following section, you will learn more about how configuration guidelines for different job types, and how to optimize both the user experience of using Backup & Replication, and the backend operations to get the most of the available infrastructure.

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.

For a graphical representation of the mentioned backup modes in this section, please see Veeam KB1799.

As a generic overview for I/O impact of the backup modes, please see this table:

Method	I/O impact on destination storage
Forward incremental	1x write I/O for incremental backup size
Forward incremental, active full	1x write I/O for total full backup size
Forward incremental, transform	2x I/O (1x read, 1x write) for incremental backup size
Forward incremental, synthetic full	2x I/O (1x read, 1x write) for entire backup chain
Reversed incremental	3x I/O (1x read, 2x write) for incremental backup size
Synthetic full with transform to rollbacks	4x I/O (2x read, 2x write) for entire backup chain

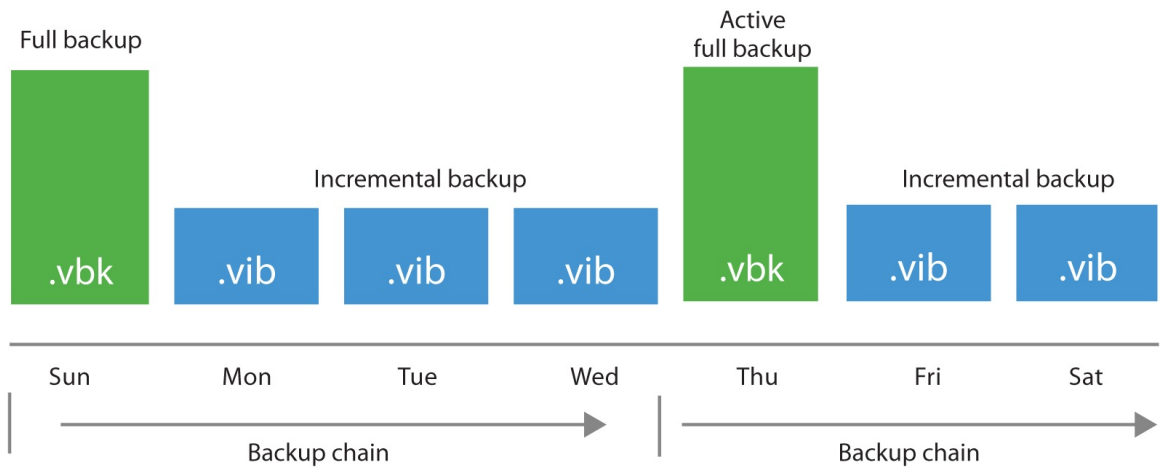
Forward Incremental

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 full backup file (.VBK).

Each time an active full is performed (either on schedule or by manually triggering the Active Full command), a new .VBK file is created by reading all data from the source VM. Incremental backups are stored in incremental backup files (.VIB).



When performing active full backups, all blocks are re-read from the source datastore. As opposed to incremental forever modes, this eliminates the recommendation for periodical health checks and compacting operations on the full backup file (VBK).

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 datastore, and this will increase the time a VM snapshot remains open (see also the "Impact Snapshot Operation" section of this guide). The snapshot lifetime can be reduced by leveraging Backup from Storage Snapshots.

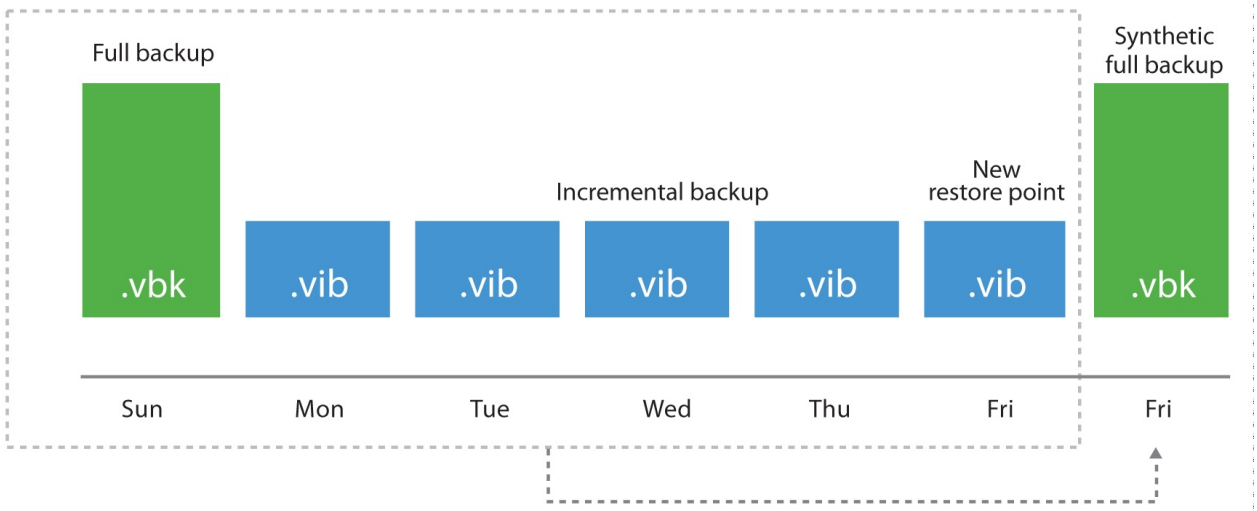
When to use

Forward incremental backup provides good performance with almost any storage and highest level of backup chain consistency since each new chain is populated by re-reading VM source data. Incremental backups are still processed using Changed Block Tracking (CBT). Active Full can be used in any case where plenty of repository space is available, the backup window allows enough time and network bandwidth is sufficient to support reading the source data in full.

Use	Don'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.

Synthetic Full

Synthetic full summarize the information stored in the most recent file chain (VBK and VIB files) to create a new full backup.



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 the incremental backup, the synthetic full generation is started. It reads the most recent version of each block for every VM in the job from the backup chain 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 or production networks during this time as I/O occurrs only in the repository.

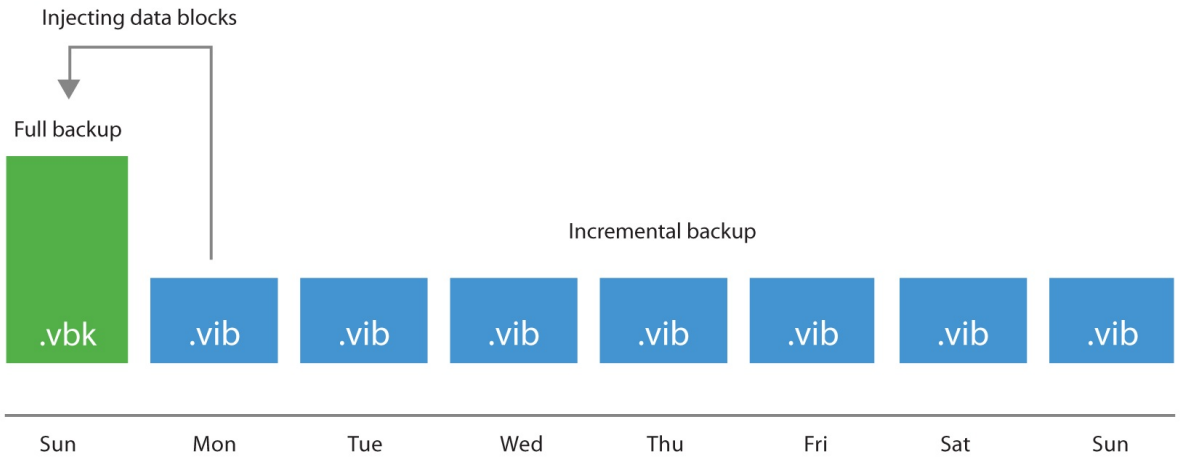
Recommendations on Usage

Due to the way this synthetic full process works, having many smaller backups jobs with fewer VMs 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.
Deduplication appliances that support offloading synthetic operations (DataDomain, StoreOnce and ExaGrid)	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 is recommended to have many smaller jobs with a limited number of VMs, which can significantly increase the performance of synthetic merge process. Very large jobs with more than 100 VMs can experience significant increase in time due to extra metadata processing. This may be remediated by combining forward incremental forever mode with per VM backup files.

Use	Don't Use
Repositories with good performing disk configuration and cache is recommended	Smaller backup repositories or NAS devices with limited spindles
Ideal for low change rate VMs	Large change rate jobs may take a long time to merge

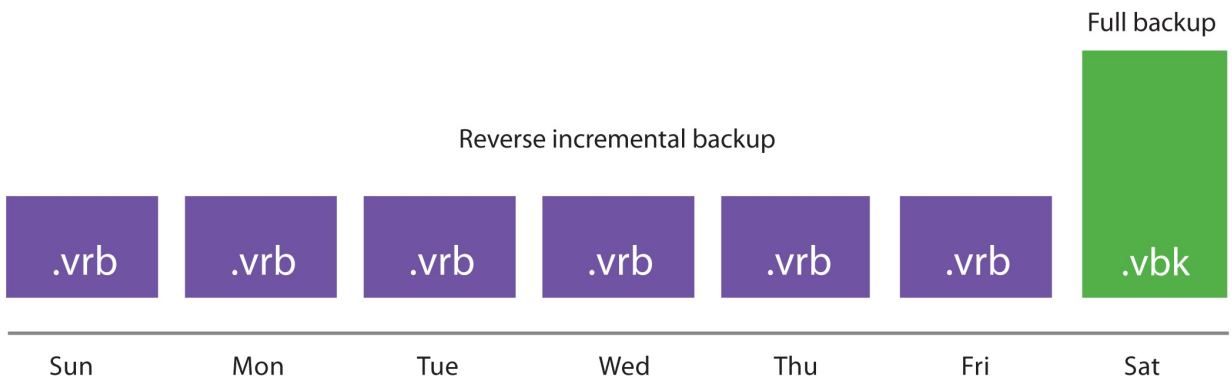
Reverse Incremental

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. During the incremental backup, changed blocks are written directly into the full backup, while replaced blocks are taken out and copied into a 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. Additionally, restore operations from the most recent point in time are faster, as the most recent point in time is always the full backup (VBK).

The disadvantage is that creation of rollback 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.

Over time, this also causes fragmentation of the VBK file. It is recommended to enable compacting on backup jobs running in reverse incremental mode without periodical active full backups enabled.



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 existing, 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 33/66 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	Don'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 the 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. If the password gets lost, the backup server will provide a challenge key for Enterprise Manager. Using asymmetric encryption with a public/private key pair, Enterprise Manager generates a response which the backup server can use for unlocking the backup file without having the password available.

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. Read more here > Data Encryption

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. If enabled, they backup file is also encrypted using an additional key generated by the Password Loss Protection mechanism.

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

Enabling encryption requires an active full backup.

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. 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 Password Loss Protection.
- Export a copy of the active keyset from Enterprise Manager (see User Guide for more information).
- 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 Password Loss Protection 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. 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 the 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 Password Loss Protection.
- 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 Password Loss Protection for these backups.

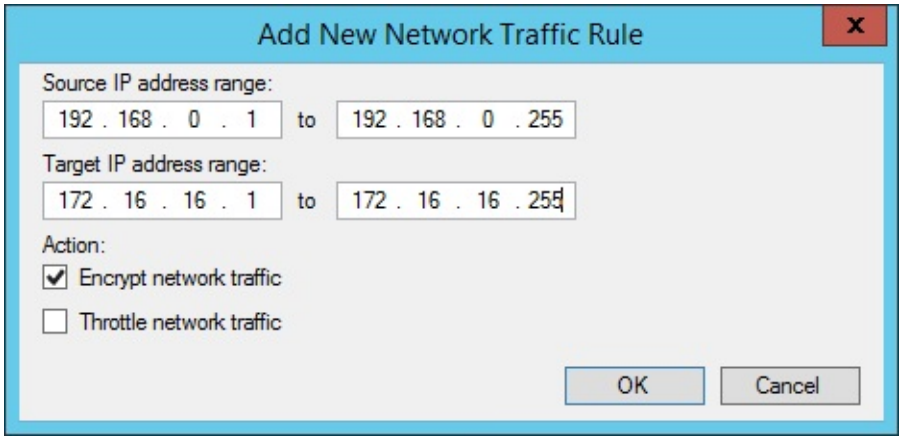
Network Transport Encryption

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, however Cloud Connect establishes a TLS 1.2 encrypted tunnel to the service 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[1].

  Note: .

- 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 following registry key to disable

automatic network-layer encryption.

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`
- Key: `DisablePublicIPTrafficEncryption`
- Type: REG_DWORD
- Value: 1 (*default: 0*)

[1]. Some newer CPU architectures can offload encryption and reduce the amount ↩

of CPU resources required. For Intel CPUs specifically, you may check your CPU model on Intel ARK and look for the "AES-NI" capability.

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 inside a single VM disk or 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 (>16 TB)** – 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 4 MB blocks.

Note: Local (>16TB) underlying block size has changed in v9, from 8 MB to 4 MB. If you upgrade to Veeam Backup & Replication 9.0 from the previous product version, this option will be displayed as "Local Target (legacy 8MB block size)" in the list and will still use blocks size of 8 MB. It is recommended that you switch to an option that uses a smaller block size and create an active full backup to apply the new setting.

When to use it?

Veeam deduplication should be enabled in almost all cases, *except* when backing up to deduplication devices. Disabling in-line deduplication in such cases significantly increases restore performance.

However, there are a few special cases where a user might consider disabling this option:

- **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.

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 full backups, 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.

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.
- 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).

Setting	Block Size	Maximum recommended job size
WAN	256 KB	4 TB of source data
LAN	512 KB	8 TB of source data
Local	1,024 KB	16 TB of source data
Local (>16 TB)	4,096 KB	64 TB of source data

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

Compression

What does it do?

The 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 the very simple RLE 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 `zlib` 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 `zlib` 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 **Dedupe-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 repository level by using the **Decompress backup data blocks before storing** advanced option in repository configuration.

Enabling compression at the job level, and decompressing once sent to the repository will reduce the traffic between proxy server and backup repository by approximately 50% on average. If proxy and repository runs on the same server, the compression engine is automatically bypassed to prevent spending CPU for applying compression. The uncompressed traffic is sent between local data movers using shared memory instead.

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 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.

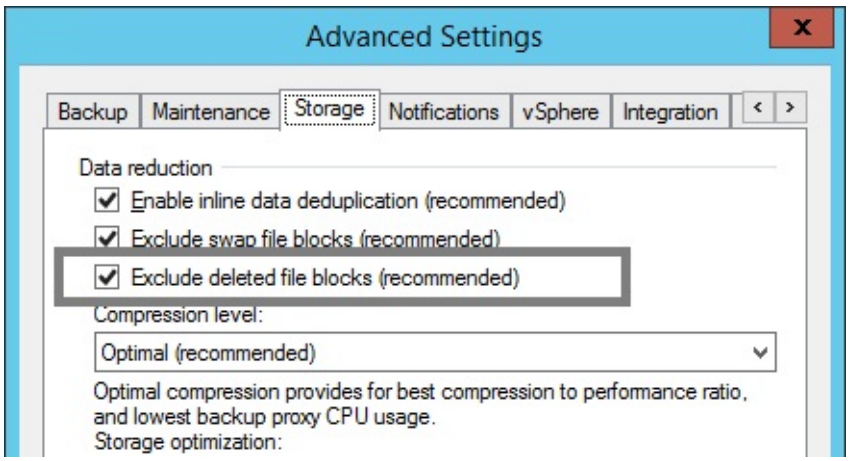
Compression settings can be changed on the job at any time and any new backup sessions will write new blocks with the new compression mode. Old blocks in already stored backups will remain in their existing compression level.

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.

BitLooker

The option "Exclude deleted file blocks" is the third configurable option in job settings. In several places you will see references to this feature under the name "BitLooker".



When enabled, the proxy server will perform inline analysis of the Master File Table (MFT) on NTFS file systems and automatically skip blocks that have been marked as deleted.

When upgrading from versions prior to 9.0, the setting is disabled for existing backup jobs. To enable it for existing jobs, use the following PowerShell commands.

```
Add-PSSnapIn VeeamPSSnapin;

Foreach ($job in Get-VBRJob) {
  $job.Options.ViSourceOptions.DirtyBlocksNullingEnabled = $true;
  $job.SetOptions($job.Options)
}
```

It is always recommended to leave BitLooker enabled, as it will reduce the amount of storage space required.

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.

Increasing Deduplication Rate

If the target repository is not configured to use per VM backup files, deduplication across all VMs within a single job is available. When using per VM backup files, deduplication is only available within a single VM backup chain, which reduces it's efficiency but still makes it relevant. The following recommendation applies to job level deduplication only.

Grouping VMs running the same operating system or deployed from similar templates into a single job will increase deduplication rate. Job sizing guidelines still apply, and it is recommended to monitor the backup window and the size of the job for manageability.

Container based jobs

Adding resource pools, folders, datastores, or vSphere Tags (vSphere 5.5 and higher) to backup jobs makes backup management easier. New machines that are member of such constructs or containers are automatically included in the backup job.

When creating jobs based on groups or constructs, ensure that the configured constructs do not overlap. Overlapping constructs may cause undesired results. For instance, when creating jobs based on datastores, VMs with disks residing on multiple datastores included in more than one backup job will cause the VM to be backed up in each job.

Tags

Tags are very convenient for a policy driven approach to data protection. However, it is recommended to follow these guidelines:

- Monitor the number of VMs automatically added to the job to avoid too many VMs being backed up within a single job
- Only one tag can be used to include a VM in a job
- Using tags, you can classify VMs by service levels, using different backup jobs for different service levels
- Veeam ONE Business View (OBV) is a very convenient tool for managing vSphere Tags. OBV allows for creating classification rules and update corresponding tags in vCenter. Classifications can be defined from CPU, RAM, VM naming convention, folder, resource pool, datastore etc. OBV can also import VM/host/datastore descriptions from a CSV file. This feature can be useful when refreshing VMware tags, for example, to update CMDB.

Exclusions

It is recommended to limit the number of exclusions in backup jobs. While exclusions can be very useful, the virtual infrastructure is dynamic and changes rapidly. It is quite possible that a VM gets moved to a folder or resource pool that is excluded which makes it unprotected. Monitoring Protected VMs with Veeam ONE is highly recommended.

Compression and Storage Optimization

Detailed descriptions of compression and storage optimization settings and their influence on the backup infrastructure is provided in the Deduplication and Compression section of this guide. 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.

When using a deduplication appliance for storing backups, please see the Deduplication Appliances section of this guide for a detailed description of compression and storage optimization settings.

Encryption

A detailed description of encryption settings and its influence on the backup infrastructure is provided in the Encryption section above in this document.

For general guidelines about encryption, refer to the Veeam User Guide: Encryption keys.

Storage maintenance

While data amount is growing and backup window is decreasing, forward incremental forever backups have become increasingly important in any backup solution. Backup jobs with no scheduled synthetic or active full backups are becoming more widely adopted. Forward incremental with weekly synthetic full backups is however still the default setting.

The two main objections towards using a forward incremental forever backup mode are the following:

The first one is **full backup file fragmentation**, leading to undesired VBK file growth over time, and degradation of performance due to fragmentation. Previously it was recommended to perform periodical active full backups in order to create a new VBK file and backup chain. This would mitigate issues of fragmentation and remove whitespace left by deleted data blocks.

The second objection is **silent storage corruption**. If ever a file or block in the chain got corrupted by a storage related issue, all subsequent consolidations or restores from this could be affected.

To address both objections, following features are available under the "Maintenance" tab, in the Advanced settings of a backup job.

Full backup file maintenance - "Defragment and compacting"

Full backup file maintenance will address two issues: VBK file fragmentation caused by transforms (forward incremental forever, or reverse incremental), and left over whitespace from deleted data blocks. These issues are mitigated by synthesizing a new full backup file on the backup repository i.e. copy blocks from the existing VBK file into a new VBK file, and subsequently deleting the original file. This process may also be referred to as "compacting".

How does it work? During VBK compacting, a new VBK file is created. Existing blocks are copied from the previous VBK, requiring free space equivalent to the size of an additional full backup in the repository. In the Restore Point Simulator, this space is part of the "Work space" parameter. When using Scale-out Backup Repository in Performance Mode, the compacting process may utilize several extents and significantly speed up the compacting process.

When to use? For every backup job with full transforms. Defragmentation will benefit the most jobs that are configured to generate a single chain per job, keeping files smaller and restore speed optimal over time.

When to avoid? When using deduplication storage, it is recommended to disable the "Defragment and compact". As deduplication appliances are fragmented by their very nature, and have very poor support for random I/O workloads, the compacting feature will not enhance backup or restore performance.

Storage-level corruption guard

In addition to using SureBackup for restore validation, storage-level corruption guard was introduced to provide a greater level of confidence for the total dataset.

How does it work? When a job has finished, storage-level corruption guard will perform a CRC verification for the entire backup chain. It will validate whether the contents of the backup chain blocks match the contents described within the backup file metadata. If a mismatch is discovered, it will attempt to repair the data block from production storage, assuming the block still exists and has not been overwritten. If it exists, the backup file will be repaired. If not, storage-level corruption guard will fail and make the user aware that a new full backup is required, and that the backup chain must be recovered from a secondary copy of the backup.

When to use? It is recommended to use storage-level corruption guard for any backup job with no active full backups scheduled. Synthetic full backups are still "incremental forever" and may suffer from corruption over time.

When to avoid? It is highly discouraged to use storage-level corruption guard on any storage that performs native "scrubbing" to detect silent data corruptions. Such storage will automatically heal silent data corruptions from parity disks or using erasure coding. This is the case for most deduplication appliances.

For more information, please see Veeam Helpcenter: Health Check for Backup Files.

Job 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 entire job chain delivers poor backup success rate.

A most common way to handle multiple jobs is to let the built-in Intelligent Load Balancing (ILB) handle the proxy/repository resources by starting multiple jobs in parallel by using all available proxy/repository resources. This will typically allow the shortest the backup window.

Load Balancing

When planning jobs schedule, you should consider 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 Storage Latency Control (or Backup I/O Control) settings.

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 Resource scheduling.

Binding Jobs to Specific Proxies

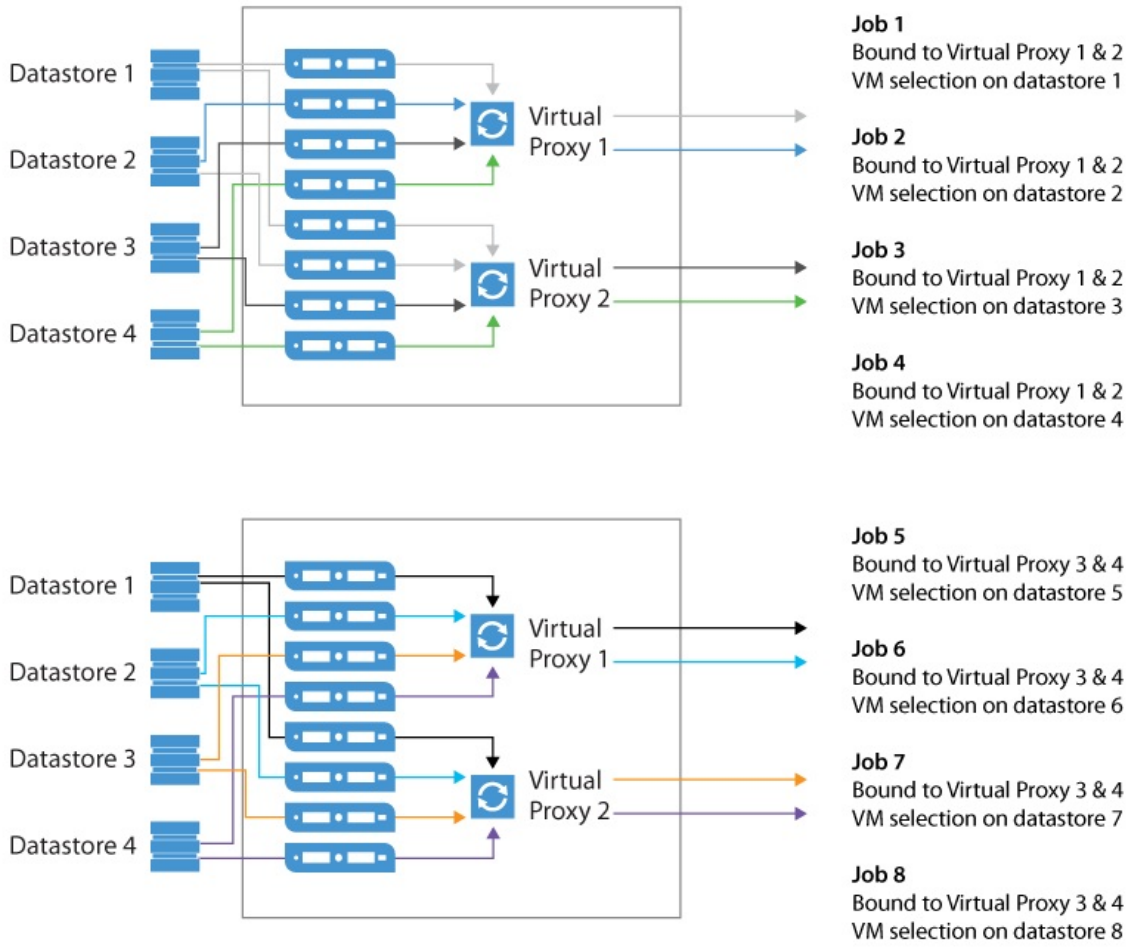
Refer to the User Guide in order to examine the advanced deployment scenario with multiple proxies: Advanced deployments.

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 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 destination. The second chain is independent from the first chain and adds therefore an additional level of protection. You can store VMs from multiple backup jobs in the same backup copy job, or you can select a subset of VMs from a bigger backup job as source if you do not want to backup all VMs to the backup copy job destination.

Every backup copy job creates its own folder on the target backup repository and stores its data in this location. The folder has the same name as the backup copy job.

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.

Backup copy jobs file chains layout will depend on the repository option: "Per VM backup files" will generate one file chain per each VM, otherwise a chain will be generated per each job.

If a backup copy job cannot process all requested VMs during an incremental execution 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. The backup copy process will be paused and resume during the next synchronization interval.

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

Important Note: only Cloud Connect jobs or Backup Copy Jobs with WAN acceleration enabled are serial, in all other scenarios backup copy job has parallel processing.

Backup Copy Job Scheduling

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

A copy job restarts at the defined **Copy every** interval setting (default is 12:00 AM daily) and looks for new restore points of the selected VMs in the specified sources. On the **Schedule** tab, it is possible to schedule the when data transfers are allowed. This is especially helpful, when transferring multiple times per day (e.g. hourly synchronization interval).

The concept of the "interval" is used to define two parameters: how often it should be looking for new points, and for daily intervals at what time it should start looking for points. If you set an interval of 1 day, that equals to instruct the backup copy job that once a day, starting at the selected time, it should begin looking for new restore points. When the restore point is found, the copy job will copy it. However, once a single point is copied, another point for that VM will not be copied until the next interval starts.

The synchronization interval is implemented to provide a policy driven approach to offsite copies. Since the copy job can contain multiple source backup jobs, and most source backup jobs neither start nor complete at the same time, the synchronization interval is helpful in defining a SLA for when it should look for restore points across the included source jobs.

The reason for this design is that you may run local backups more often (for example, hourly), but you may only want to copy data offsite only daily 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. **Pre-job activity** — if enabled, the pre-job scripts are executed at the very beginning of a copy interval.
2. **Health check** — if scheduled, backup file integrity is verified before the next copy is initiated.
3. **Data transfer (synchronization) phase** — during this phase, the backup copy job checks for a new restore point in the source, creates a file for a new restore point at the target and starts copying the state of the latest 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 Synchronization Intervals). 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 Backup Copy Window).

4. **Transform phase** — copy jobs are by nature running in "forever forward incremental" mode, and performs transform operations on the target backup repository accordingly. Additionally, it is possible to schedule health checks or backup file compacting as described in the Backup Job section. The transform phase begins when all VMs are successfully copied to the target, or if the synchronization interval expires.

   Note: the transform process itself puts additional pressure on the target repository. In large environments with deduplication storage appliances used as backup repositories or with backup copy jobs processing a large number of VMs or big VMs, the transform process can take a significant amount of time. For non-integrated deduplication appliances, it is recommended to use the "Read entire restore point..." option. This forces the Backup Copy Job to running forward incremental with periodical full backups copied entirely from the source backup repository rather than being synthesized from existing data.

5. **Compact full backups** — if enabled, the recent full backup file is re-created to avoid fragmentation.

6. **Post-job activity** — if enabled, several post-job activities are executed before the job enters the idle phase, such as post-job scripts and sending e-mail reports.

7. **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 the synchronization interval expires, a new interval starts at 1.

For more information, refer to the corresponding section of the User Guide > Backup Copy Job.

Job Layout and Object Selection

Source Object Container

- **Select from infrastructure**: this selects specific VMs or containers from the infrastructure. The scheduler will look for the most recent restore point containing the VMs within the synchronization interval. The scheduler will look for restore points in all backups, regardless which job generated the restore point. If the restore point is locked (e.g. the backup job is running), the backup copy job waits for the restore point to be unlocked and then start copying state of the VM 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 job you want to copy. The job container will dynamically protect all VMs in the selected

source job(s).

- **Select from backup**: this method is equivalent to the **Select from infrastructure** method, but allows for selecting specific VMs inside specific backups. This is helpful, when only certain critical VMs should be copied offsite.

Backup Copy and Tags

As you can select any VM to be copied from multiple backups, you can plan for policy-based configurations. 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 to apply tags.

Initial synchronization

When creating the initial copy to the secondary repository, it is recommended to use backup seeding (see Creating Seed for Backup Copy Job) whenever possible. Especially when transferring large amounts of data over less performant WAN links, the seeding approach can help mitigating initial synchronization issues.

While Backup Copy Jobs were designed for WAN resiliency, the initial copy is more error prone, as it is typically transferring data outside the datacenter over less reliable links (high latency, or packet loss).

The most frequent synchronization issues are described in the User Guide > Handling Backup Copy Job Issues.

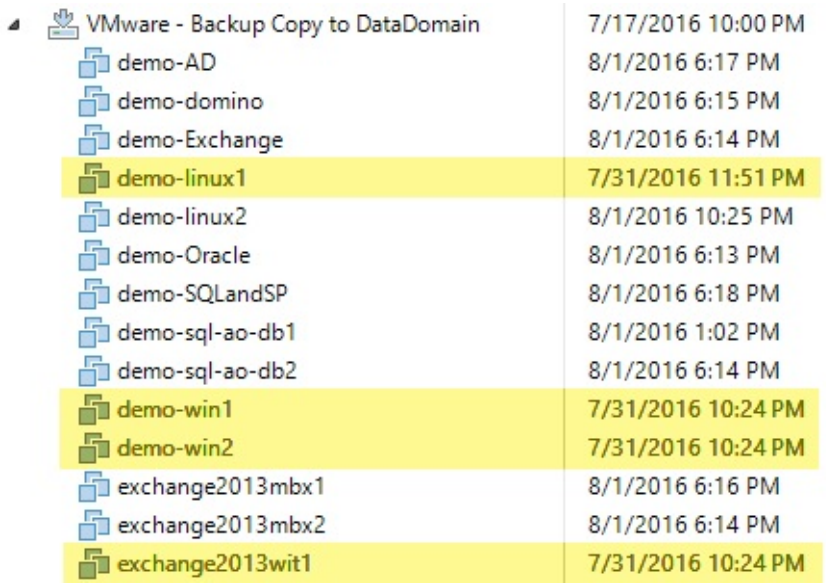
Additional Options

Restore Point Lookup

By default, after a restart of the job interval (the **Copy every** setting), a backup copy job analyzes the VM list it has to protect, and searches *backwards in time* for newer restore point states. If the state of the restore point in the target repository is older than the state in the source repository, the new state is transferred.

For example, if the backup job is scheduled to run at 10:20 PM, and the backup copy job uses the default schedule of copying the latest restore point state every day at 10:00 PM, the state copied by the backup copy job is typically one day behind. In the image below, you

can see some VMs affected by this behavior.



To change this behavior, it is possible to use the `BackupCopyLookForward` registry key as described below. Reevaluating the example above, using this registry key, the backup copy job will instead start searching at 10:00 PM, but will instead wait for the new restore point state created after this point in time.

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`
- Key: `BackupCopyLookForward`
- Type: REG_DWORD
- Value: 1

The following forum thread provides a very good explanation of the backup copy scheduler and the LookForward registry key > Veeam Community Forums - Backup Copy Intervals

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 used to send data remotely. If it is necessary to send data over a slow link, you can seed the backup copy job by taking the following steps:

1. Create a "local" backup copy job and target it at a removable device used as a backup repository, or copy the backup files afterwards. Run the created backup copy job to create a full backup set on this device. Note that also the .vbm file has to be moved.
2. Once the backup copy job is over, delete the local backup copy job from the Veeam

console.

3. Transport the removable device with the created backup files to the destination site.
4. Copy backup file to the target backup repository.
5. Import the backup on the target. If already imported, perform a rescan.
6. Create the final backup copy job on the Veeam console. On the **Target** step of the **Backup copy job** wizard, use the **Map backup** link and select the transported 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: https://helpcenter.veeam.com/backup/vsphere/wan_populate_cache.html.

Note: Only the initial first run of a reverse incremental chain can be used with seeding (but any forward incremental chain can be used).

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.

1. 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 (hotadd) transport mode for writing replica data to target. In case of NFS datastore at target, you can as well use Direct Storage access mode (Direct NFS) to write the data.

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 https://helpcenter.veeam.com/backup/vsphere/virtual_appliance.html.

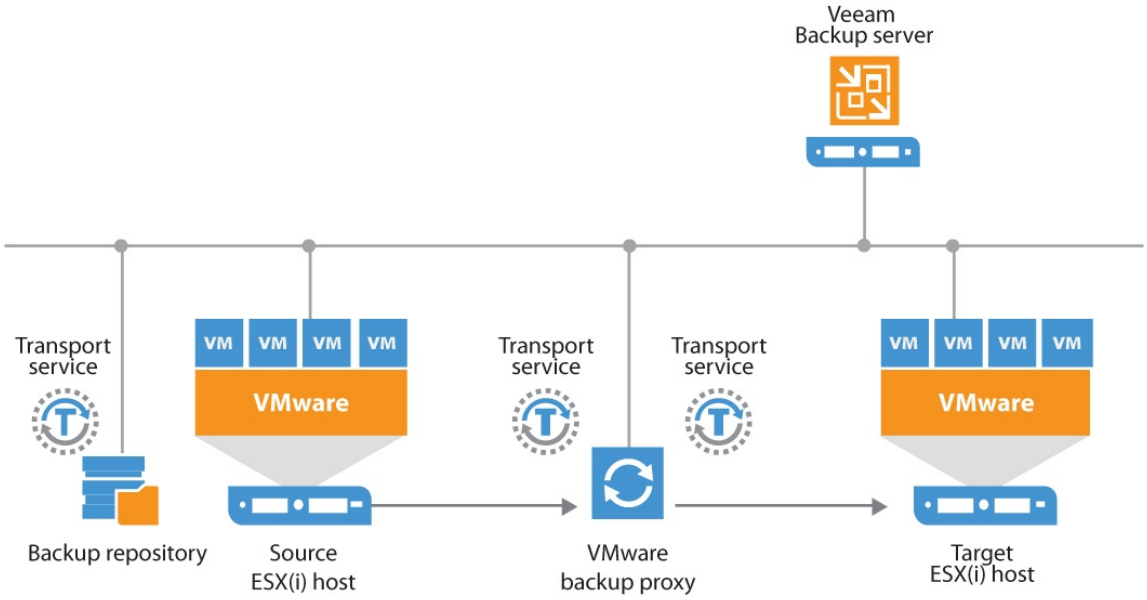
If the backup proxy is deployed on a physical server, or the Virtual Appliance or Direct NFS mode cannot be used for other reasons, Veeam Backup & Replication will use the Network transport mode to populate replica disk files. For information on the Network mode, see

https://helpcenter.veeam.com/backup/vsphere/network_mode.html.

The Direct SAN mode (as part of Direct Storage Access) 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.



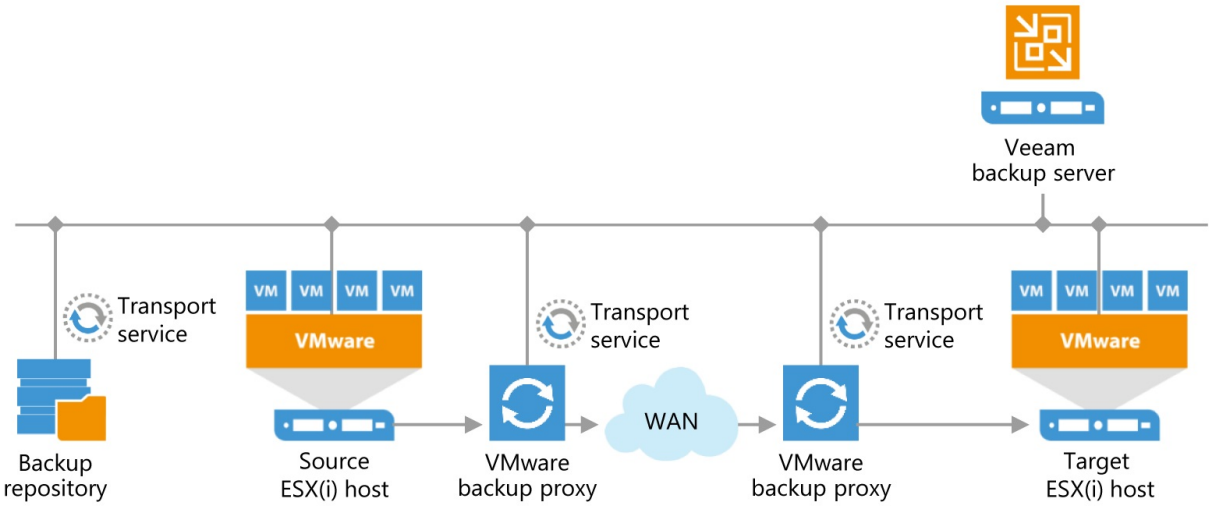
Offsite Replication

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.

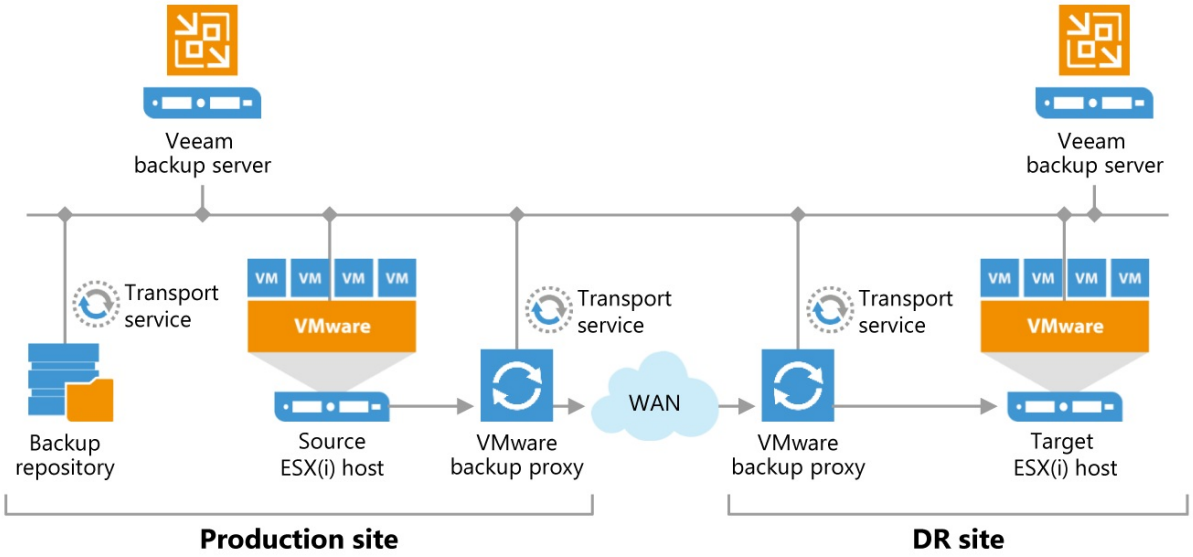
For offsite replication, open the connections between the Veeam backup components:

- 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.

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 from the replica target side (only) directly to Veeam Backup & Replication as managed servers and to perform replication without vCenter Server on the target side. In this szenario a failover an be performed from the Veeam console without an working vCenter Server itself (for example to failover the vCenter Server).

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 replication 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: https://helpcenter.veeam.com/backup/vsphere/setting_network_traffic_throttling.html.

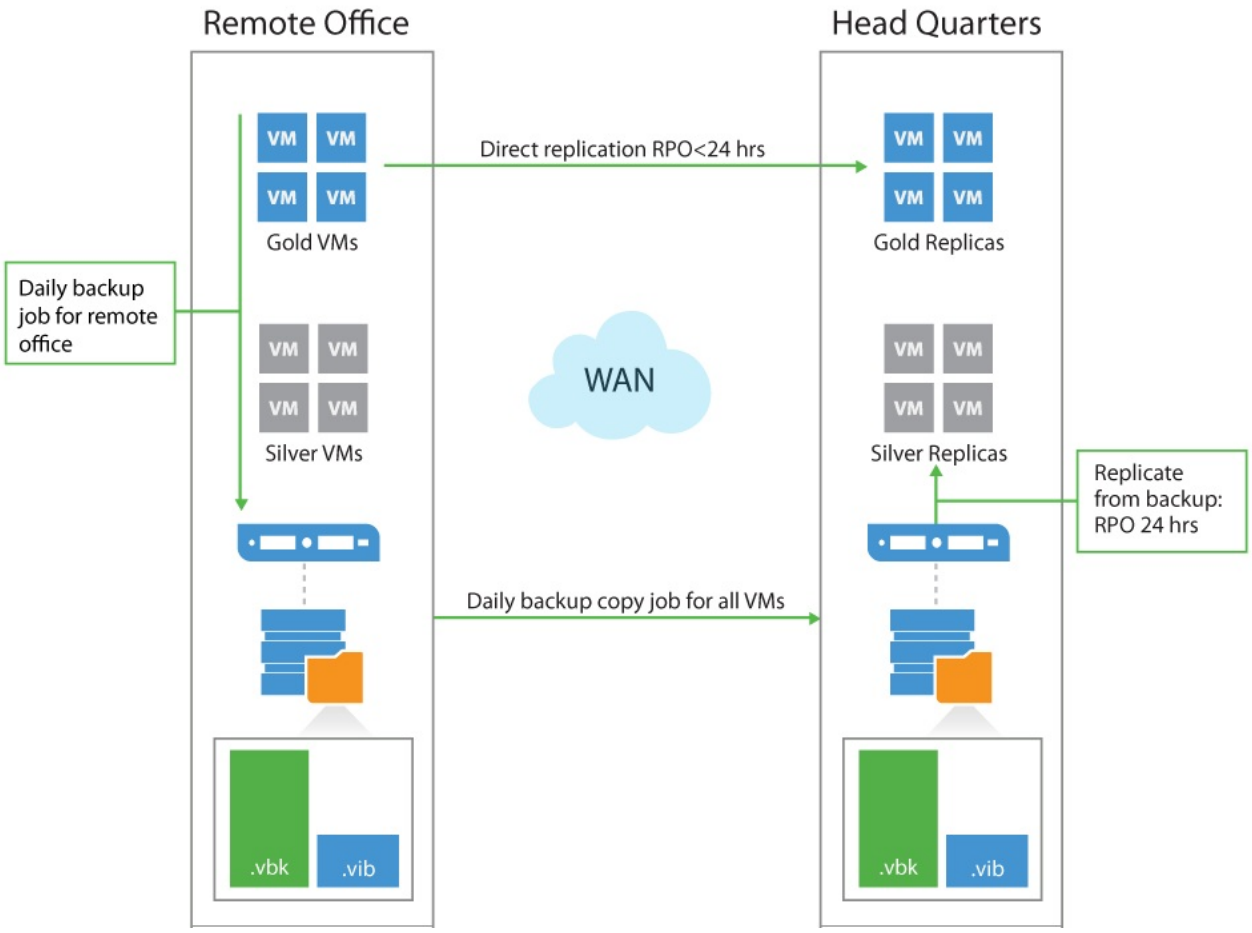
Tip: Replication can leverage WAN acceleration allowing a more effective use of the link between the source and remote sites. For more information, see the User Guide https://helpcenter.veeam.com/backup/vsphere/wan_acceleration.html or the present document (the “WAN Acceleration“ section above).

Replication from Backups

When using replication from backup, the target VM updates directly from the backup files created by a backup or backup copy job.

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 beside the usage of proactive VM restores, 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: https://helpcenter.veeam.com/backup/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.

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's proprietary application-aware image processing (using Microsoft VSS or Linux scripts). Key features of both methods are illustrated by the following table:

Feature	VMware Tools Quiescence	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

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) or using native Oracle commands (for Oracle databases on Linux).

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 (https://helpcenter.veeam.com/backup/vsphere/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)	Linu appli
Guest OS processing is applicable	Y	Y	Y	Y	Y
Use VMware Tools quiescence	N	Y	N	N	N
VMware Tools quiescence with VMware Script processing	Y	N	N	N	N
Enable Veeam Application-Aware Image Processing	N	N	Y	N	N
Enable Veeam Application-Aware Image Processing and InGuest Scripts	N	N	N	Y	N
Disable Veeam Application-Aware Image Processing	N	N	N	N	Y

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).

Guest Interaction Proxy

Depending on the guest VM operating system and/or Veeam Backup and Replication Edition different servers may be selected to perform guest processing step and initiate connection to a VM as per the table below.

Edition	Windows	Linux
Standard	Backup server	Backup server
Enterprise	Guest interaction proxy	Backup server
Enterprise Plus	Guest interaction proxy	Backup server

Any Windows server managed by Veeam Backup and Replication can be selected to act as guest interaction proxy but the preference would be given to the server that has IP address in the same subnet as subject VM. This functionality allows for having only small limited range of ports to allow through the firewalls in restricted environments and for that reason it is recommended to have guest interaction proxies in all VM subnets that are not supposed to be directly accessible from the network where Veeam backup server resides.

For details on network configuration refer to the section "Required ports" below.

Tip: If the backup server has no network connection to the VMs and deploying additional guest interaction proxies is not practical/possible (for example, service provider environments), order in which backup server or guest interaction proxy tries to communicate to a VM can be changed using the following registry key:

- Path: `HKEY_LOCAL_MACHINE\SOFTWARE\Veeam\Veeam Backup and Replication`
- Key: `InverseVssProtocolOrder`
- Type: REG_DWORD
- Value: **0** - try connection through RPC, failover to VIX (default)
- Value: **1** - try connection through VIX, failover to RPC

RPC connections means injecting the file via the "ADMIN$" share on the target VM. See Veeam Knowledge Base article at http://www.veeam.com/kb1230 for more information. Consider that this is a global setting that will be applied on the Veeam backup server level

and affects all jobs with application-aware image processing.

Guest Access Credentials

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.

Tip: To verify the credentials you supplied on the Guest Processing step of the job wizard, click **Test Now** button.

Windows OS

Application-Aware Image Processing (AAIP)	VMware Tools Quiescence	Veeam via VIX	Veeam via RPC	Disabled (crash-consistent)
Membership in the local Administrators group	User account not needed	No	Yes	Not needed
Enter username as *<servername>\ Administrator* or *<domain>\Administrator*	No	Yes[1]	No	No
UAC can be enabled	Yes	Yes[2]	Yes	Yes
VMware Tools must be installed and up to date	Yes	Yes	Yes	No

Linux OS

Linux guest OS processing	VMware Tools Quiescence	Veeam via SSH	Disabled (crash-consistent)
Root user account	No	Yes	No
User requires `sudo` rights	No	Yes	No
Certificate-based authentication available	No	Yes	No
VMware Tools must be installed and up to date	Yes	Yes	No

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 (https://helpcenter.veeam.com/backup/vsphere/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.

File exclusions

Another operation Veeam Backup can do on guest OS level (NTFS only) is excluding certain files or folders from the backup. Alternatively the job can be configured to include only specified files or folders in the backup.

This functionality operates very similarly and shares a lot of characteristics with excluding Windows page file and deleted file blocks. It may help reduce size of the backup files or implement additional data protection strategies for specific data. Backups for which this option was enabled remain image-level and hypervisor APIs are used to retrieve VM data. File exclusion feature uses a combination of NTFS MFT data and guest file system indexes collected by in-guest coordination process to determine which virtual disk blocks belong to the excluded files and thus should not be included in the backup.

Full file/folder paths, environment variables or file masks can be used to define exclusions. For more details on configuring exclusions and its limitations refer to the corresponding User Guide section.

Note: Generic file exclusions (defined for high level folders) are most effective. File masks exclusions require guest file system indexes and generating indexes may put additional stress on guest VM and will increase backup time. For this reason it is recommended to

avoid using file system masks especially on fileservers with large number (thousands) of small files and use high level folder exclusions instead. When using include filters, file exclusions are created for everything else and can take significant time.

How file exclusion works

For each VM in a job that has exclusions enabled Veeam Backup and Replication performs the following operations:

1. Virtual machine NTFS MFT is read into the memory cache on the backup proxy, data blocks that store excluded files are marked as deleted.
2. When sending data blocks to target repository data is read both from the VM snapshot and memory cache on the backup proxy. Target repository reconstructs VM disks without excluded VM blocks.
3. Virtual machine NTFS is modified using the data in the cache on the proxy and information about excluded data blocks is saved in the backup file or replica metadata. This information is necessary as CBT is not aware of which blocks were excluded and is used to determine which blocks should be processed during the next backup session.

[1] Only this account is able to bypass the UAC prompt for launching processes with administrative privileges. If not applicable, see [2].

[2] When performing application-aware image processing on Windows via VIX, UAC must be entirely disabled, unless the user account is the local administrator account (SID S-...-500).

Application Specific Configuration

Microsoft SQL Server

When backing up AlwaysOn availability group make sure all cluster nodes are processed by the same backup job for transaction logs processing and restores to work properly.

Oracle

Refer to the corresponding section of the User Guide (https://helpcenter.veeam.com/backup/vsphere/backup_job_vss_oracle_vm.html) for details on configuring Oracle database backup and transaction logs processing.

Only databases in ARCHIVELOG mode will be backed up online, databases in NOARCHIVELOG mode will be shut down which will cause **database availability disruption**. Before the backup the database (in ARCHIVELOG mode only) is put into backup mode, this has to be taken into consideration when performing restore - restoring database server VM is not enough for restoring the service, database has to be put out of backup mode:

```
ALTER DATABASE END BACKUP
```

Note: 32-bit Oracle instances on 64-bit Linux are not supported.

Tip: Avoid using aggressive logs truncation settings for databases protected with Data Guard as it may affect logs synchronization to secondary server. Data Guard should have enough time to transport logs remotely before they are truncated thus generally having "Delete logs older than" option less than 24 hours is not recommended.

Logs are stored temporarily on the guest filesystem before they are shipped for processing. This may cause undesired behavior if there is no enough space available in default location and changing temporary location from default is recommended as per KB 2093.

Permissions

Certain level of access is expected from the user account configured for performing Oracle backup. Refer to the corresponding section of the User Guide for details (https://helpcenter.veeam.com/backup/explorers/veo_connection_to_source_server.html).

When processing Linux instances, the same user account specified for application awareness is used to process the Oracle backup. For Windows instances, you may specify two separate accounts.

Note: It is not possible to use different accounts to access different Oracle instances running on the same VM, make sure specified credentials can be used to access all instances on a VM in those cases.

Windows OS

User account used to connect to a VM should have local administrator privileges on guest VM and read/write access to database files on filesystem level.

In addition this account or separate Oracle account in case it is different should have SYSDBA rights, this can be achieved by adding it to **ora_dba** local group.

Linux OS

Root account or account elevated to root should be used to connect to a VM. Automatic adding to **sudoers** can be enabled for the account but note that **sudoers** file entry will not be removed automatically. Persistent **sudoers** file entry with *NOPASSWD: ALL* option can be added manually, for example:

```
oraclebackup ALL=(ALL) NOPASSWD: ALL
```

This account should be included in the **oinstall**[1] group to access Oracle database files hierarchy, and to **asmadmin** group (where applies).

In addition this account or separate Oracle account in case it is different should have SYSDBA rights, this can be achieved by adding it to **dba** local group.

Oracle on Linux backup workflow

1. Coordination component which will perform all the necessary steps is injected into the guest VM. This component is the same as the one used for Linux application-aware image processing in general.
2. Perform application discovery. This is done using native OS methods, coordination component queries /etc/oraInst.loc and reads inventory.xml which is then compared to /etc/oratab information.
3. Status and version of instance(s) is fetched.
4. Disk group information is retrieved for ASM instances.
5. Log mode is identified, this information will later be used for decisions on how exactly

the database has to be processed. Database files, CDB (Oracle 12 only) and current DBID information is retrieved.
6. At this step archive log necessary information was collected and Veeam will start doing actual backup, modifying database state - current archive log is archived and all archive log information is retrieved.
7. PFILE backup is created and archived into the backup metadata.
8. Additional information is collected and recorded (current DBID, SCN, Sequence IDs, database unique name, domain, recovery file destination, basic listener information and current archive log).
9. Coordination component is shut down and then restarted again to finalize the backup: database is put into backup mode and database snapshot is created.

Oracle on Windows backup workflow

Behavior on Windows depends on the state of VSS writer, Oracle version and database type.

	VSS enabled	VSS disabled	Pluggable database
Oracle 11	Oracle VSS writer is engaged, NOARCHIVELOG databases are shut down and excluded from VSS processing	Same worflow as for Linux	N/A
Oracle 12	Oracle VSS writer is engaged, NOARCHIVELOG databases are shut down and excluded from VSS processing	Same worflow as for Linux	Same workflow as for Linux, VSS writer is skipped

Other applications

It is possible to ensure data safety and transactional consistency for applications not supported by Veeam Backup and Replication natively using pre-freeze and post-thaw scripts that will execute inside of the virtual machine. Subject application has to provide the way to prepare itself appropriately.

Generally speaking pre-freeze and post-thaw scripts have to (depending on the capabilities of the application):

- Pre-freeze - freeze transactions or create application-level consistent snapshot of its data. Alternatively application services can be shut down but this involved short user service downtime and thus is not desirable.
- Post-thaw - unfreeze transactions or delete snapshot created by pre-freeze (where

applies). In case services were shutdown they should be started again.

Certain applications do not require these steps as they include self-healing mechanics or maintain transactional consistency by other means, application documentation has to be checked and/or application vendor has to be contacted for specifics on achieving this.

Note that in addition to configuring application consistency for such applications, restore process has to be properly planned as additional steps would have to be followed to restore them as well.

Domino

Backup and restore of IBM Lotus Domino is covered in this Veeam webinar: https://www.veeam.com/videos/backing-up-non-vss-aware-applications-ibm-lotus-domino-4867.html

SAP HANA

Pre-freeze scripts can be used to create HANA snapshot before the backup starts. This snapshot can be used as transactionally consistent state of database after restoring HANA VM.

An example of ensuring database consistency for SAP HANA is described on Veeam community forums: https://forums.veeam.com/veeam-backup-replication-f2/sap-b1-hana-support-t32514.html.

> [1]. Unless installed from rpm/deb, the group name may differ. To find correct values, search this file: `$ORACLE_HOME/rdbms/lib/config.c`. ↩

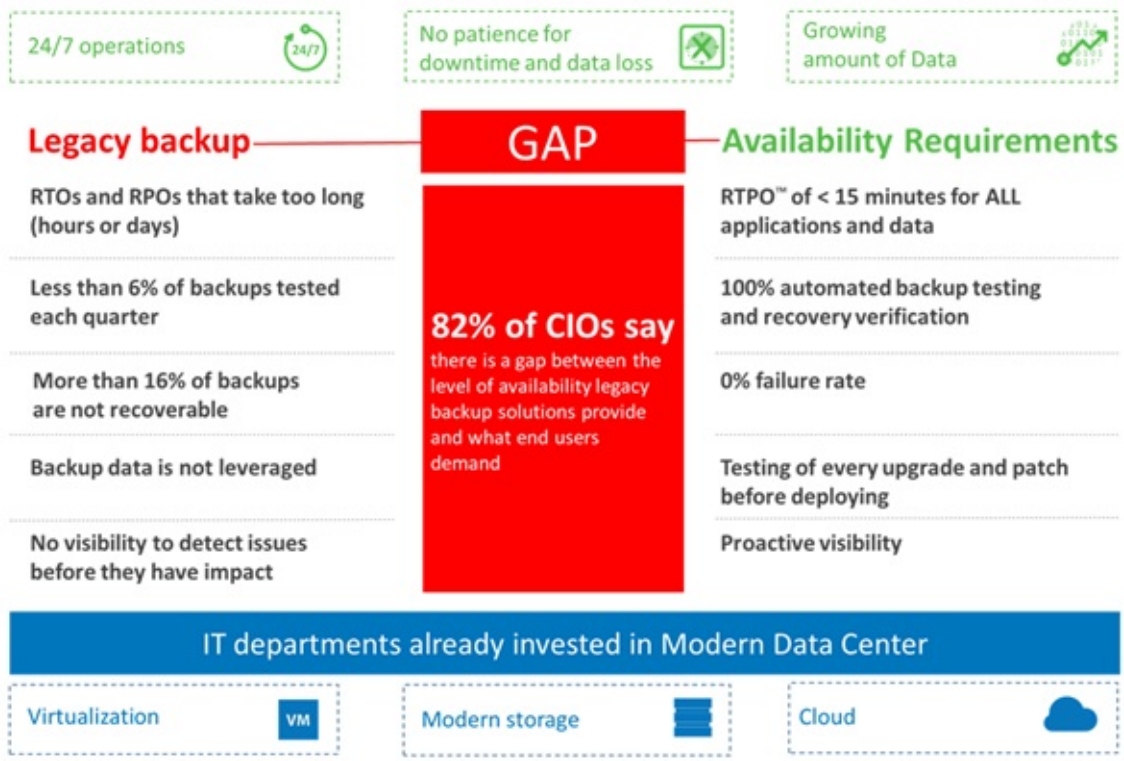
POC Guide

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 1,000 virtual machines, Veeam Availability Suite is applicable to any infrastructure size.

Assessment

Before starting a project, it is very important to understand customers' needs, vision and the IT environment. While the first two can 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

Alerts tab

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.

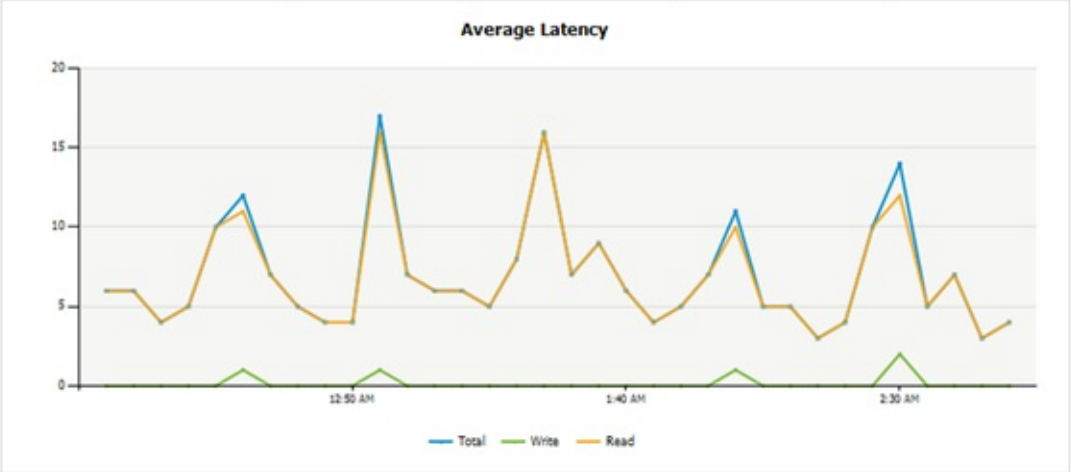
Veeam ONE Reporter

Storage Latency

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

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.

Change Rate Estimation

Summary

Top 5 VMs with Largest Change Rate (GB)

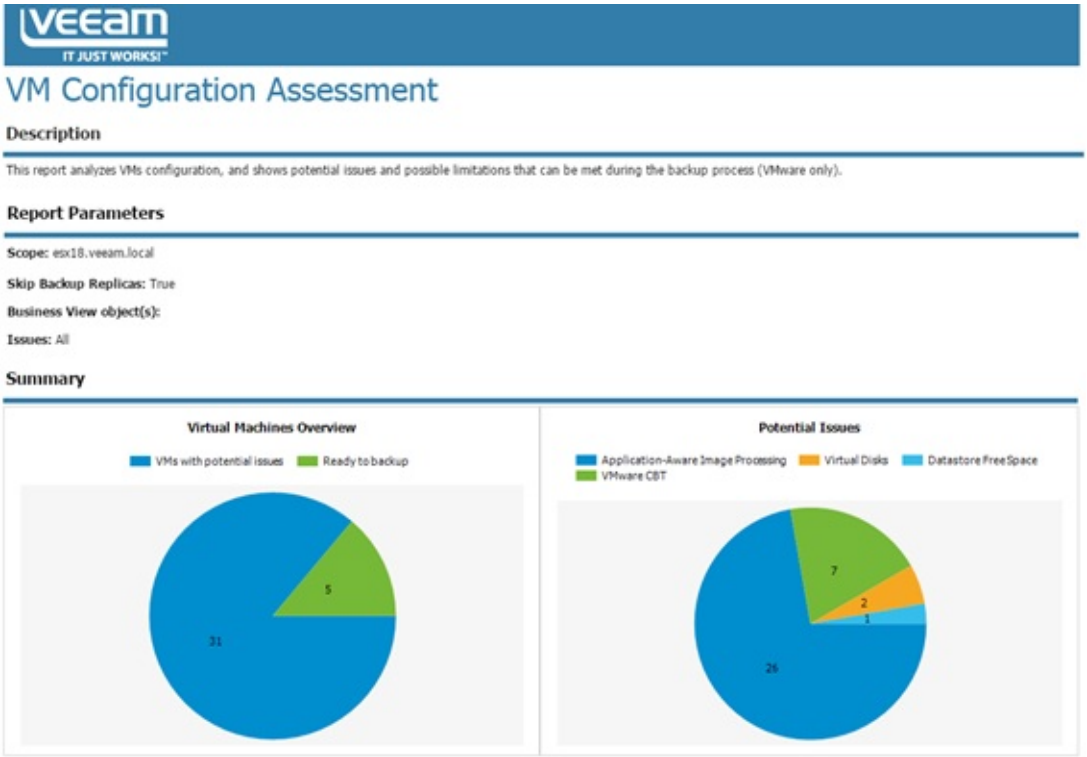
Top 5 VMs with Least Change Rate (GB)

Details

Scope	VM	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
⊟ clu0w8r2	10	57.82 GB	12.81 GB	11.03 GB	123.50 GB	8.73 GB	9.07 GB	10.87 GB	233.82 GB
	alba02	47.70 GB	2.68 GB	< 1 GB	114.59 GB	< 1 GB	< 1 GB	1.85 GB	166.96 GB
	scvmm0w8r2	3.86 GB	4.02 GB	4.11 GB	3.96 GB	4.17 GB	4.39 GB	3.96 GB	28.47 GB
	win2012r2	2.55 GB	2.29 GB	2.73 GB	1.06 GB	< 1 GB	< 1 GB	1.15 GB	11.37 GB
	pd2008x64	1.44 GB	1.59 GB	1.66 GB	1.44 GB	1.80 GB	1.43 GB	1.81 GB	11.17 GB
	oracle	0.99 GB	1.02 GB	1.03 GB	1.03 GB	< 1 GB	1.08 GB	1.04 GB	7.00 GB
	mediaserver07	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	4.32 GB
	proxy0w8r2	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	1.33 GB
	win2012	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	1.29 GB
	mediaserver02	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	1.19 GB
	virtual_repository	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB	< 1 GB
Total:		57.82 GB	12.81 GB	11.03 GB	123.50 GB	8.73 GB	9.07 GB	10.87 GB	233.82 GB

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 Configuration Assessment



VEEAM
IT JUST WORKS!™

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: esx18.veeam.local

Skip Backup Replicas: True

Business View object(s):

Issues: All

Summary

Virtual Machines Overview

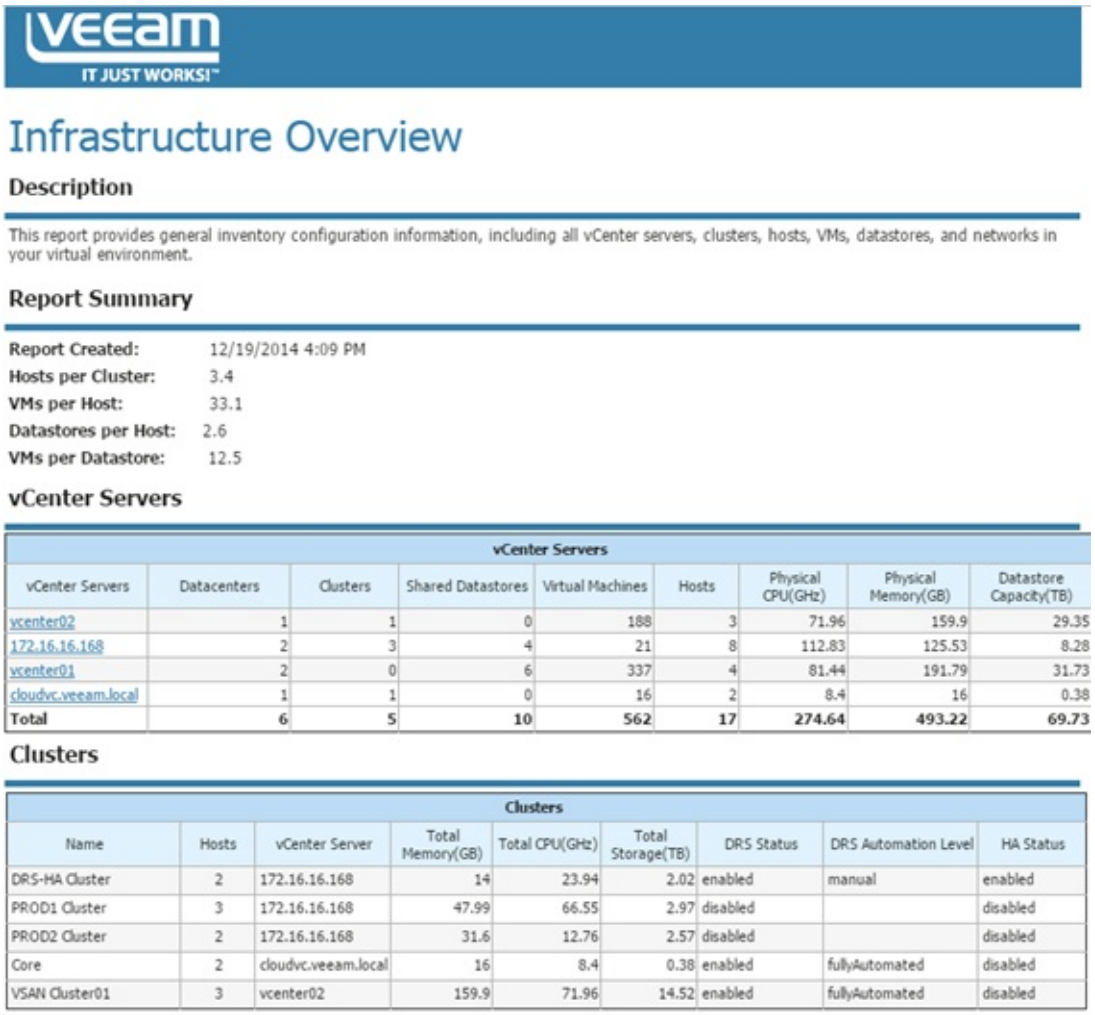
VMs with potential issues
Ready to backup

Potential Issues

Application-Aware Image Processing
Virtual Disks
Datastore Free Space
VMware CBT

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.

Infrastructure Overview



VEEAM
IT JUST WORKS!™

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

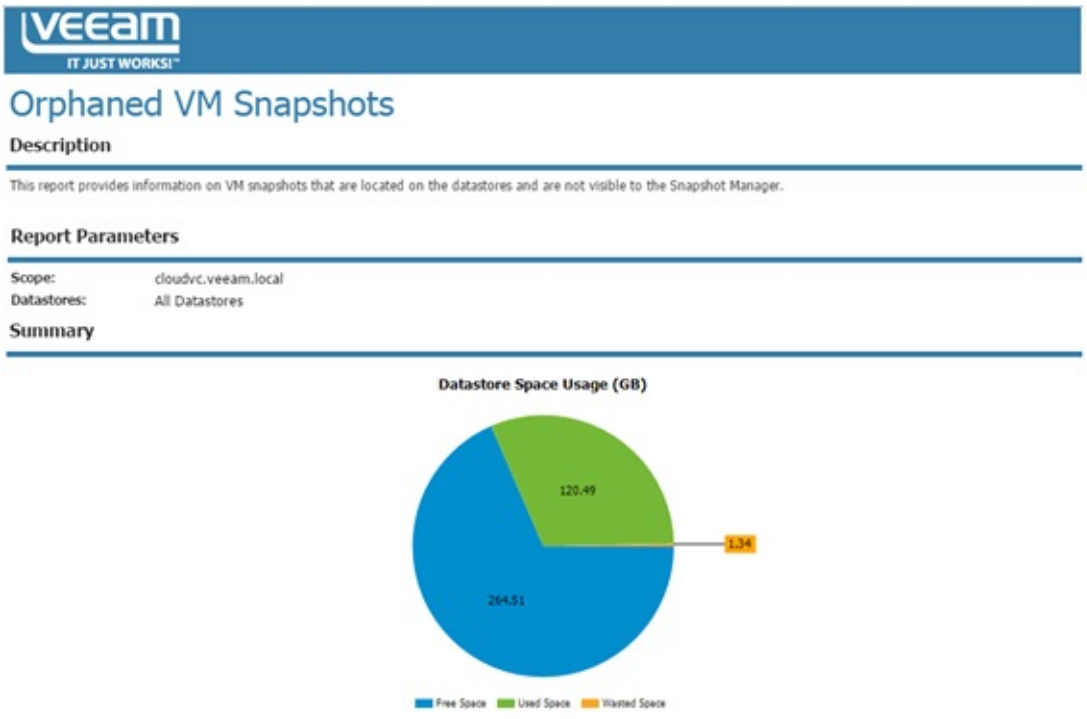
Clusters

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

Active Snapshots

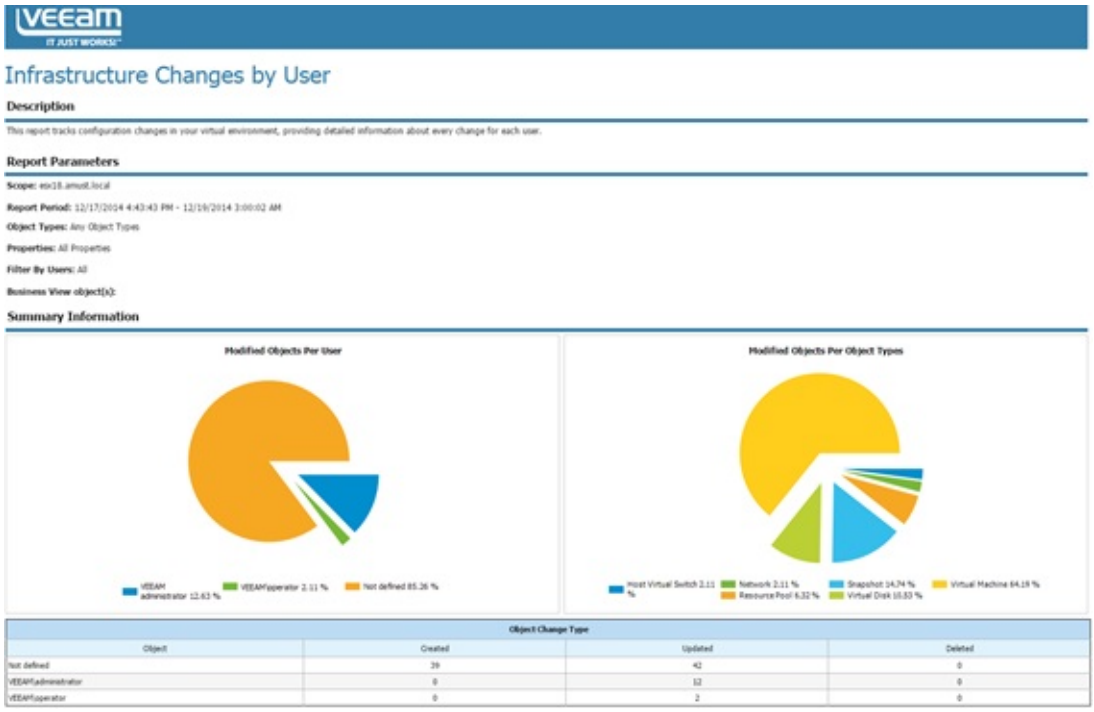
VMware snapshots are often done to save a specific state of the VM for some time. While they are created very easily, administrators forget to delete them over time. Together with administrators, you can release all snapshots that are not needed anymore. This will help prevent datastore downtimes because of snapshots filling up the whole physical storage.

Orphaned Snapshots



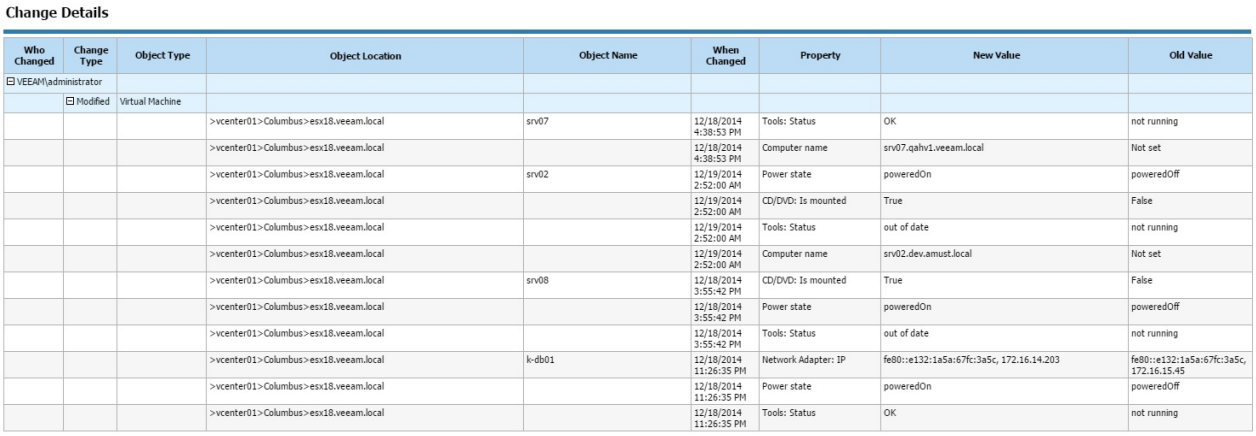
This report detects VM snapshots that are still active on datastores but do not show up in the VMware Snapshot Manager. Veeam Backup & Replication and its Snapshot Hunter will correct this situation by consolidating these snapshots, which can bring extra load at the first backup POC. We strongly recommend that you tune the VMware environment and consolidate all orphaned snapshots before you start a Backup & Replication project.

Infrastructure Changes by User



In the later POC phase, create a separate account for a VMware user and use this account for all authentication operations in Veeam Backup & Replication. With the Infrastructure Changes by User report, you can track and document all changes done by this user.

Inventory



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>esx18.veeam.local	srv07	12/18/2014 4:38:53 PM	Tools: Status	OK	not running
			>vcenter01>Columbus>esx18.veeam.local		12/18/2014 4:38:53 PM	Computer name	srv07.qahv1.veeam.local	Not set
			>vcenter01>Columbus>esx18.veeam.local	srv02	12/19/2014 2:52:00 AM	Power state	poweredOn	poweredOff
			>vcenter01>Columbus>esx18.veeam.local		12/19/2014 2:52:00 AM	CD/DVD: Is mounted	True	False
			>vcenter01>Columbus>esx18.veeam.local		12/19/2014 2:52:00 AM	Tools: Status	out of date	not running
			>vcenter01>Columbus>esx18.veeam.local		12/19/2014 2:52:00 AM	Computer name	srv02.dev.amust.local	Not set
			>vcenter01>Columbus>esx18.veeam.local	srv08	12/18/2014 3:55:42 PM	CD/DVD: Is mounted	True	False
			>vcenter01>Columbus>esx18.veeam.local		12/18/2014 3:55:42 PM	Power state	poweredOn	poweredOff
			>vcenter01>Columbus>esx18.veeam.local		12/18/2014 3:55:42 PM	Tools: Status	out of date	not running
			>vcenter01>Columbus>esx18.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>esx18.veeam.local		12/18/2014 11:26:35 PM	Power state	poweredOn	poweredOff
			>vcenter01>Columbus>esx18.veeam.local		12/18/2014 11:26:35 PM	Tools: Status	OK	not running

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.

There are many additional interesting reports in the Veeam Availability Suite.

Check out the VMware Optimization or Hyper-V Optimization sections of Veeam ONE Reporter. A good example is the Garbage Files Report that can identify possible wasted space on datastores. In some cases, it helped to free up 10 TB+ of space on the tier 1 storage.

Accelerated Evaluation

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

It 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:

- VMware vSphere environments
- Microsoft Hyper-V environments

Enhanced Evaluation

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 Enhanced Evaluation 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.

Enhanced Evaluation - Workshop Example

This section describes how to conduct 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.

This section describes how to conduct 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:
    i. Virtualization platform and version
    ii. Main storage system, type, connection
    iii. Is storage virtualization used (between the storage arrays and hypervisor)?
2. Depict the second main customer datacenter (if available)
    i. Are there any storage replication/mirroring involved?
    ii. 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:
    i. Is there 10 GbE LAN?
    ii. Is there a WAN connection between the 2 datacenters?
    iii. What is the VMKernel Interface physical link speed?
    iv. Is vCenter Server physical or virtual? Where is it located?

    This is necessary to know if you plan to use the Virtual Appliance or Network backup mode.

10GbE gives you faster processing for the Network mode. To learn more, see the “Backup Proxy” chapter.

4. Define the amount of production data:
    i. Number of VMs (this can help to design jobs)
    ii. Used data (this can help to define the backup target and configure jobs settings)
    iii. Number of ESXi hosts and number of used sockets (this regards Veeam licensing).
    iv. Number of clusters
    v. Other information
5. Create the first Veeam implementation draft/sample scenario:
    i. 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.
    ii. 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.
    iii. 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.)
    iv. 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:
    i. Create a list of used operating systems.
    ii. 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.
    i. Define the number of VMs that are business critical.
    ii. 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

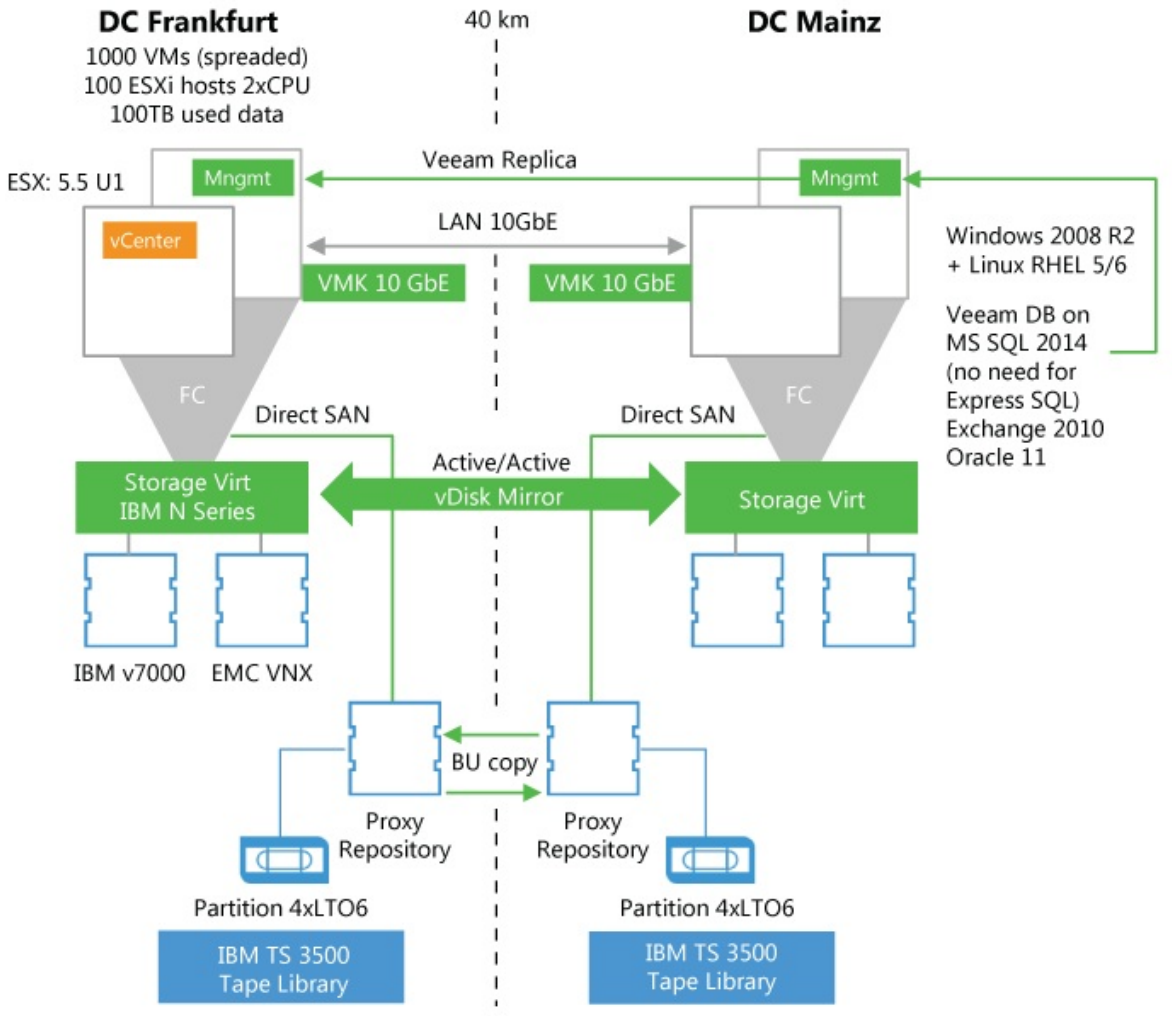
Veeam Availability Suite is very flexible and lets you implement different backup infrastructure schemes. Firewalls can be used between all backup infrastructure components. The only exception is RPC inspection functionality: it can cause delays in connections, and Veeam Backup & Replication can run 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

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.

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 Evaluation - Preparation

After having agreed and discussed the points in the Workshop Example section, 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 steps

1. Prepare for the POC planning with the Veeam User Guide and this document.
2. Check out the necessary 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: Perform firewall planning very carefully: if something is misconfigured, this may block the entire POC. In most cases, it is not easy to detect problems and introduce 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) 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.
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 and Veeam KB1999.
7. Ask the customer to prepare a decent performing storage system for the POC. Avoid

low-end NAS appliances for enhanced evaluations.

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.

Command line

Following operations are managed through the Windows command line:

- Installation - Link to Help Center
- Updates - Link to Help Center

PowerShell

Operations in Veeam Backup & Replication can be automated with Veeam PowerShell snap-in in the following areas:

- 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: 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 > Veeam Help Center - PowerShell Reference.

You can find help for the scripts in the Veeam Community Forums - PowerShell section. If you need some examples, refer to the following thread: Getting Started and Code Examples

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.

Here is a list of external resources:

- Veeam Help Center - RESTful API Reference
- Veeam Community Forums
- Veeam Help Center - Beginner Example

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.

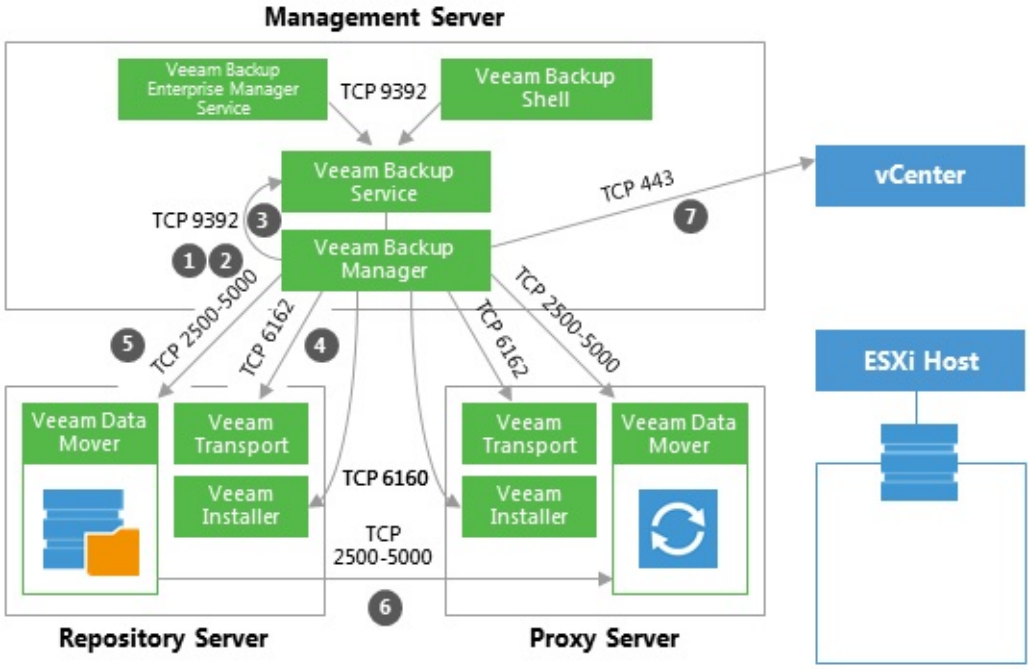
1. Initialization Phase

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.

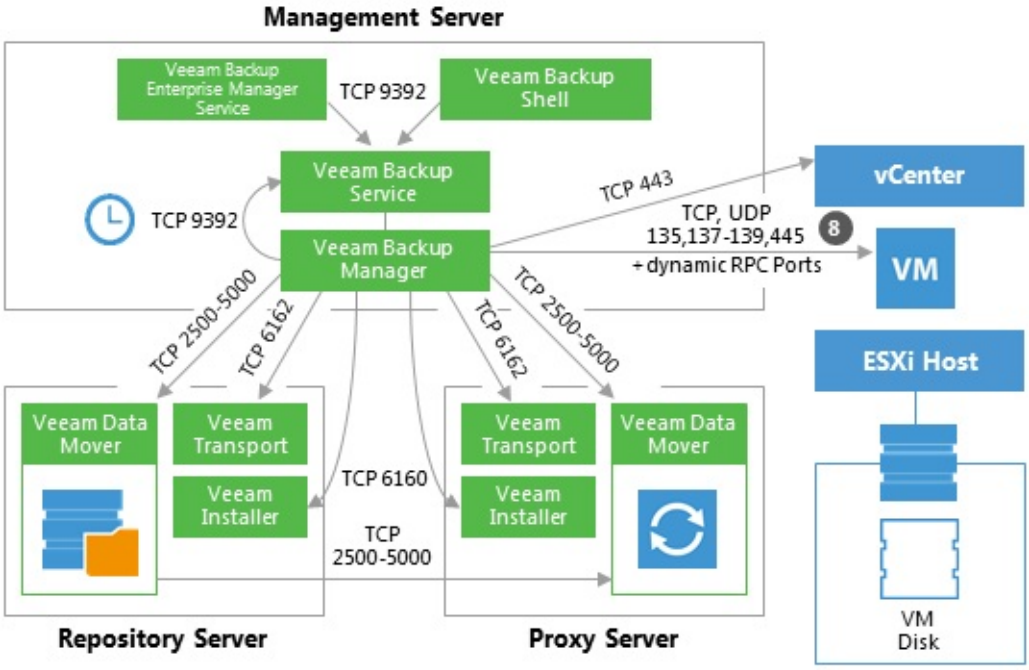
7. Veeam Backup Manager connects to the vCenter Server or ESXi host and gathers metadata about VMs and hosts engaged in the backup process. At this step, no connection between the Veeam backup server and VM guest networks is established.



2a. Guest Processing for Windows-Based VMs

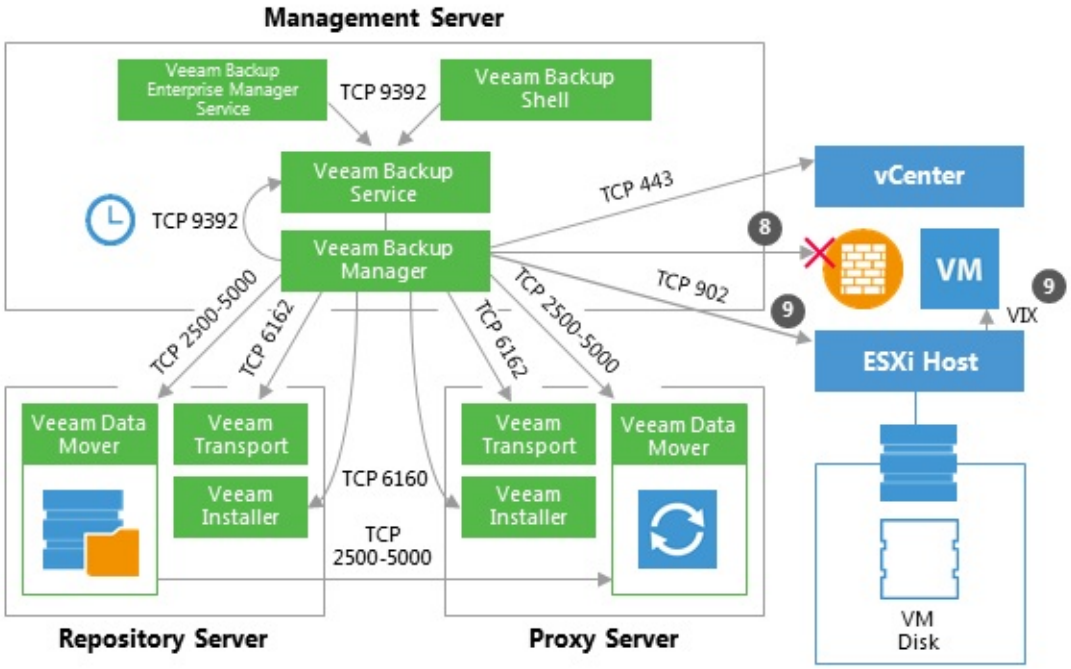
For VMs with Microsoft Windows guest OS, Veeam Backup & Replication obtains information about the guest's IP addresses from VMware Tools. Veeam uses these IP addresses to connect to the guest OS and perform in-guest processing tasks (if application-aware image processing is enabled).

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

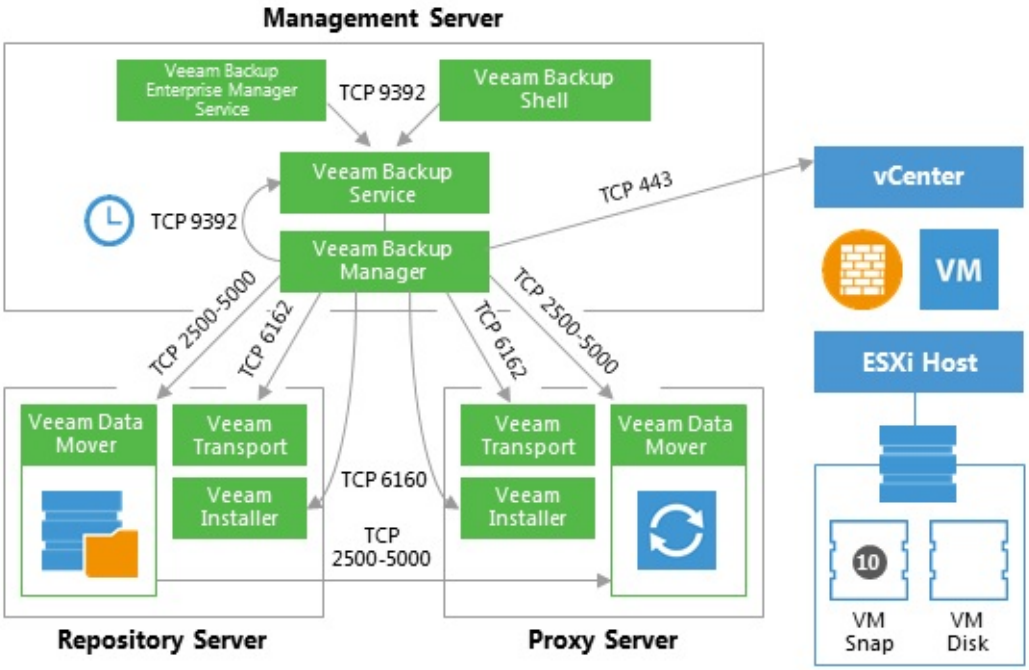
If pre-freeze and post-thaw scripts are enabled in the backup job properties, Veeam Backup & Replication obtains information about the guest's IP address from VMware Tools. Veeam uses this IP address to connect to the guest network over SSH and perform in-guest processing tasks. Scripts reside on the 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

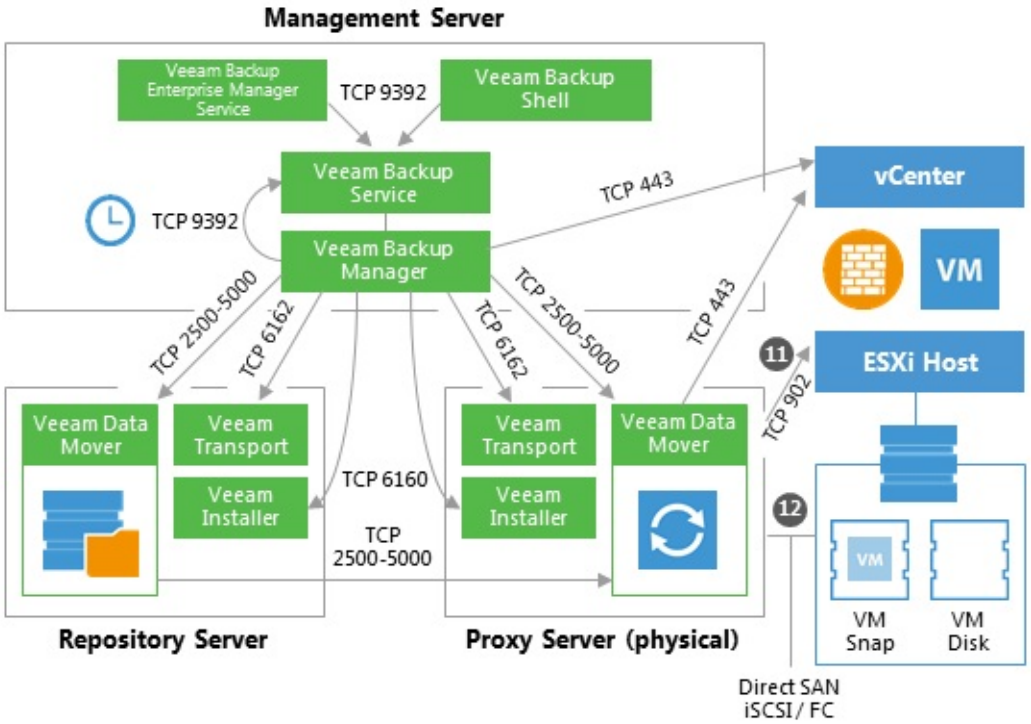
To read and transfer data from the VM snapshot, Veeam Backup & Replication can use one of the following transport modes:

- 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.

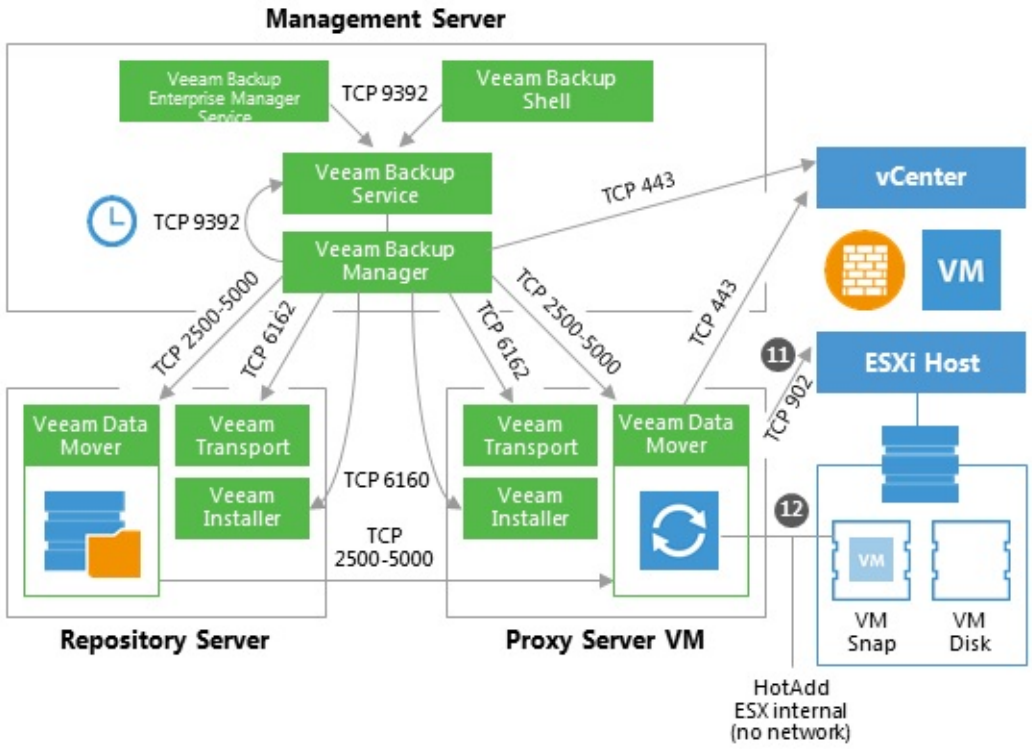
5a. Direct SAN Access Data Transport Mode

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.



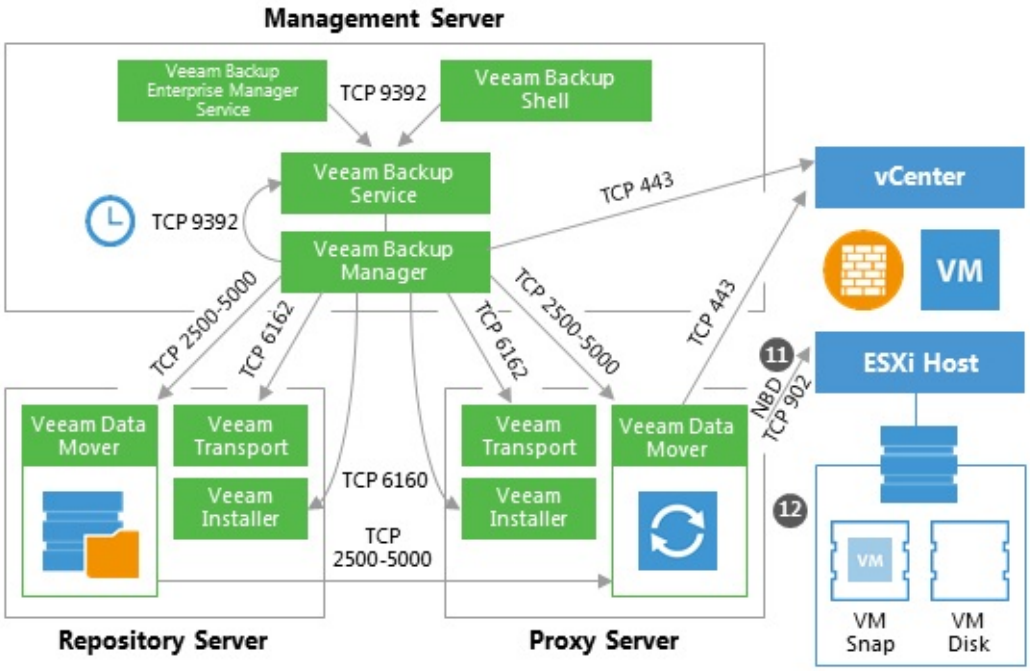
5b. Virtual Appliance Data Transport Mode

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.



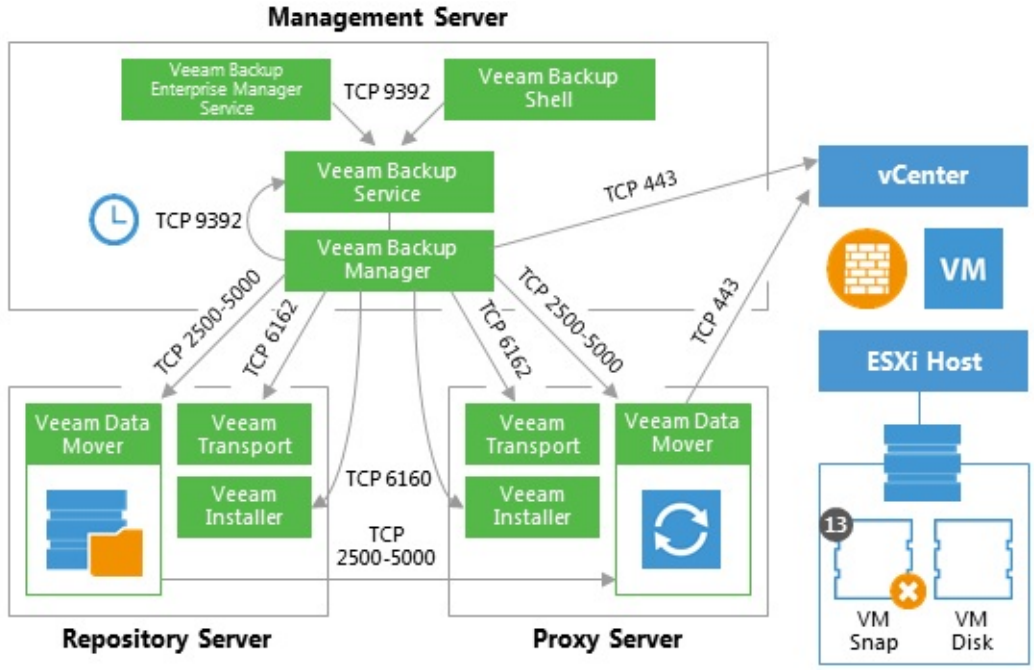
5c. Network Data Transport Mode

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.



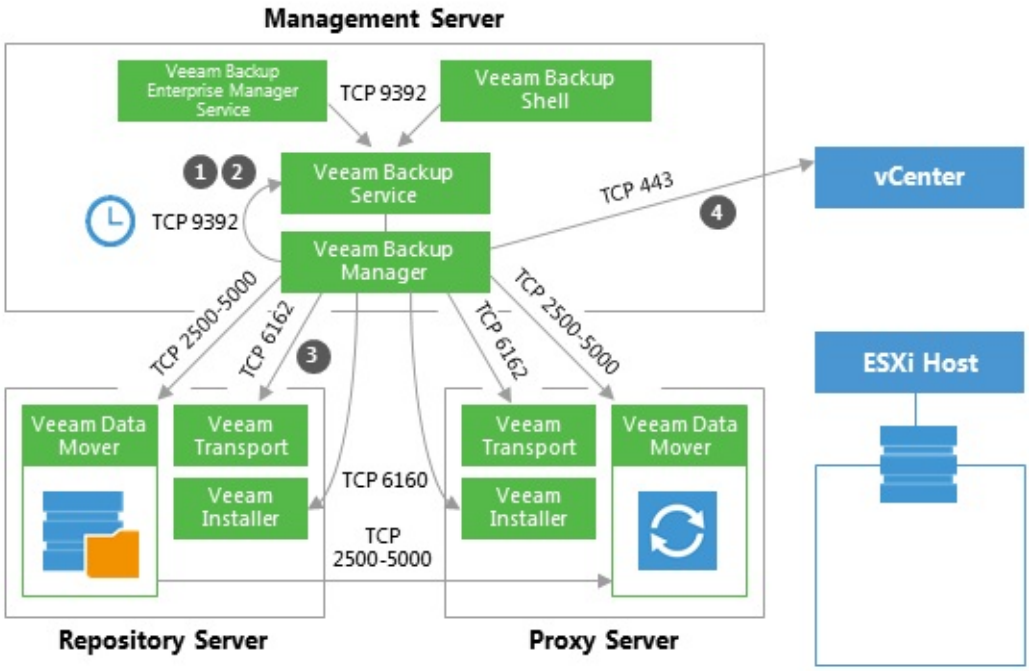
VM Restore

This section provides a step-by-step description of a full virtual machine restore process implemented in Veeam Backup & Replication.

1. Initialization Phase

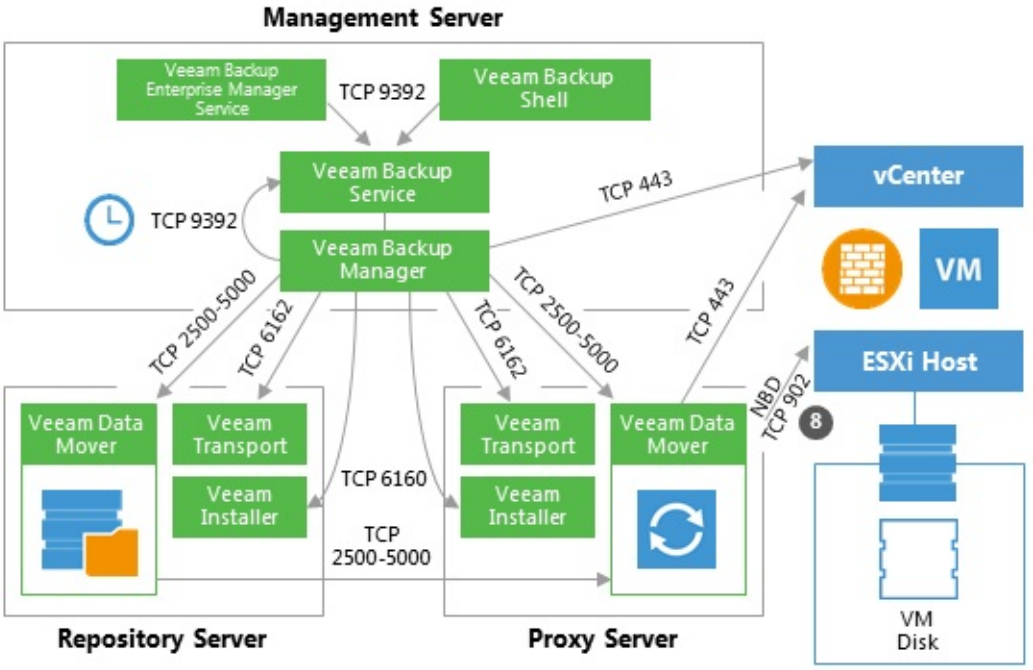
In the initialization phase, Veeam Backup & Replication prepares the resources necessary for full VM recovery. It performs the following steps:

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

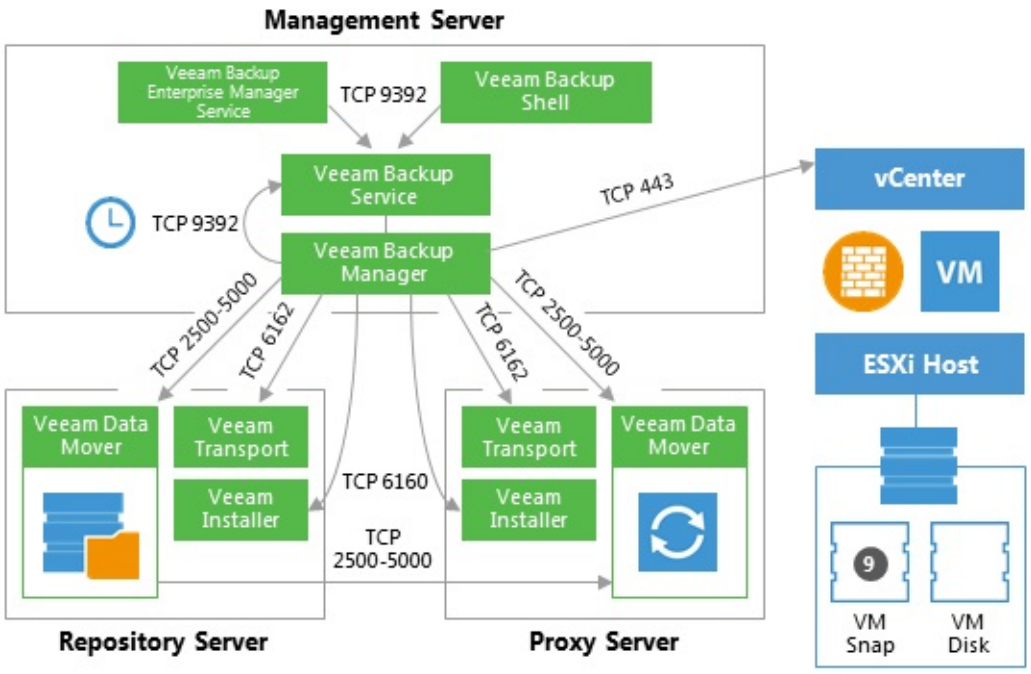
Veeam Backup & Replication retrieves VM configuration data from the backup and restores it on the chosen ESXi host/datastore. Next, it instructs VMware vSphere to register the restored VM on the host. If a user selects to change VM configuration (for example, disk format or network settings) during restore, 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:

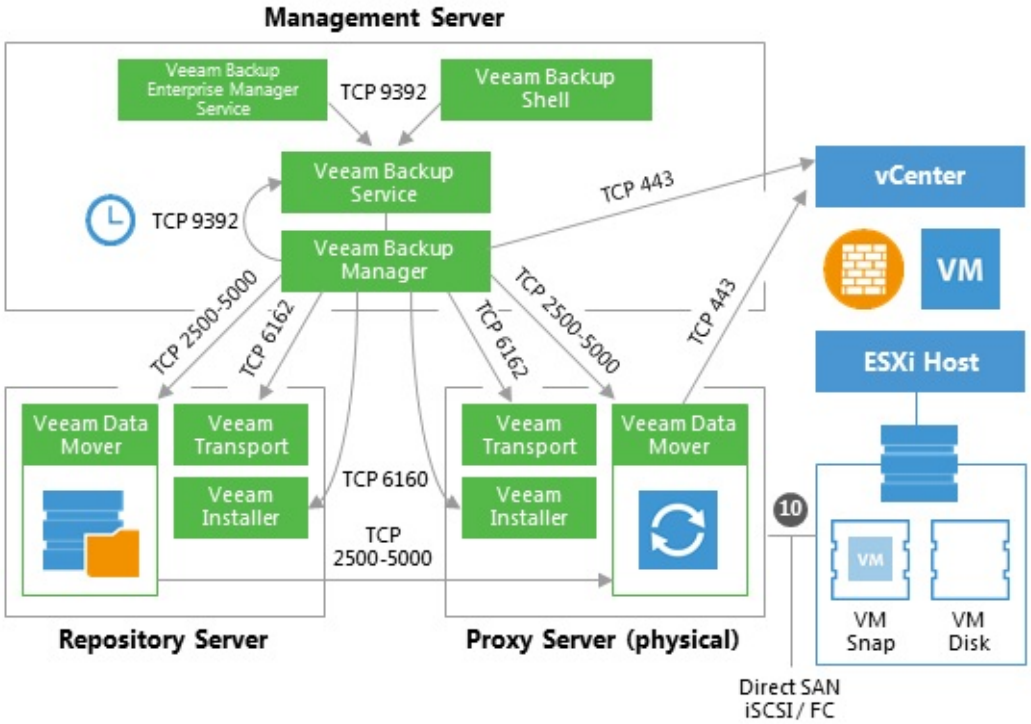
- Direct SAN Access
- Virtual Applicance (HotAdd)
- Network (NBD)

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

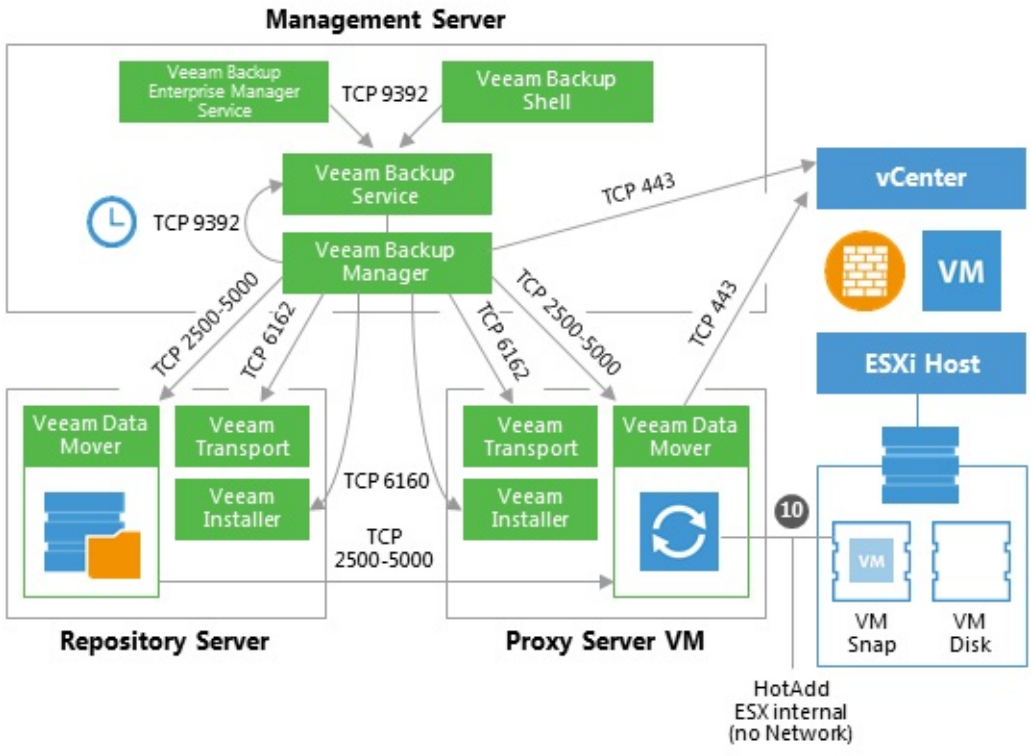
This mode is available only for VMs that have all disks in thick provisioning.

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.



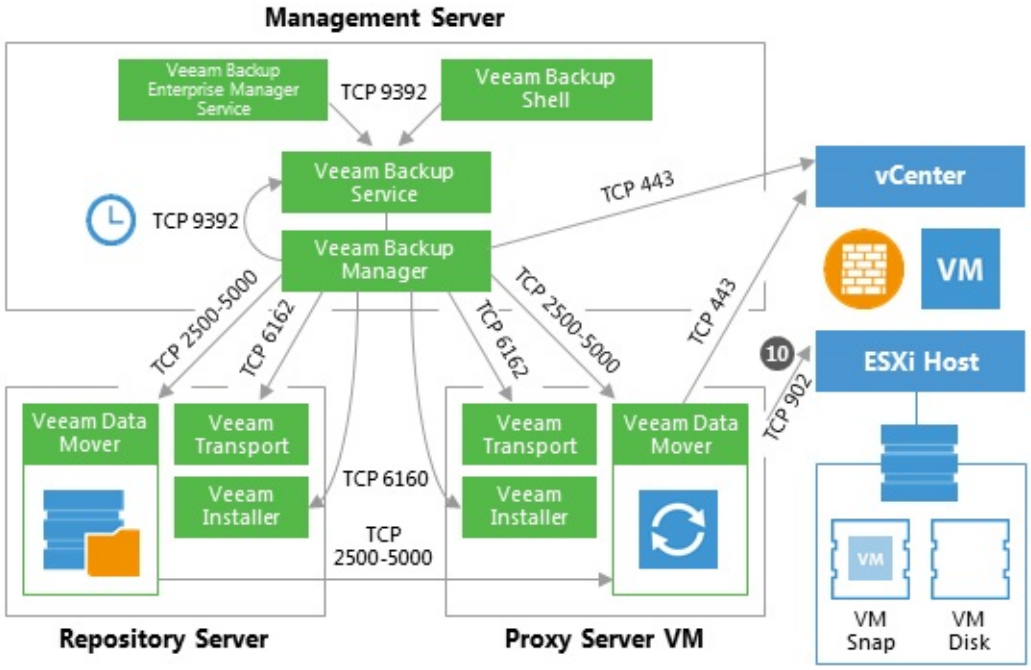
4b. Virtual Appliance Data Transport Mode

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.



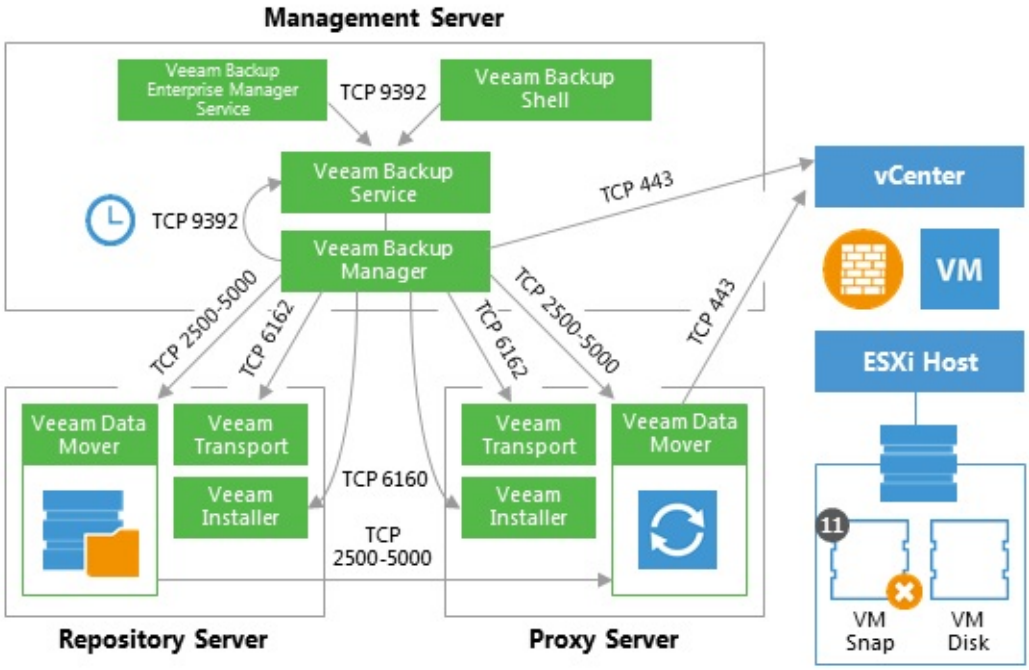
4c. Network Data Transport Mode

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.


Management Server
Veeam Backup Enterprise Manager Service
TCP 9392
Veeam Backup Shell
Veeam Backup Service
TCP 9392
Veeam Backup Manager
TCP 443
vCenter
VM
ESXi Host
TCP 2500-5000
TCP 6162
TCP 6162
TCP 2500-5000
TCP 443
Veeam Data Mover
Veeam Transport
Veeam Installer
TCP 6160
TCP 2500-5000
Veeam Transport
Veeam Installer
Veeam Data Mover
11
VM Snap
VM Disk
Repository Server
Proxy Server

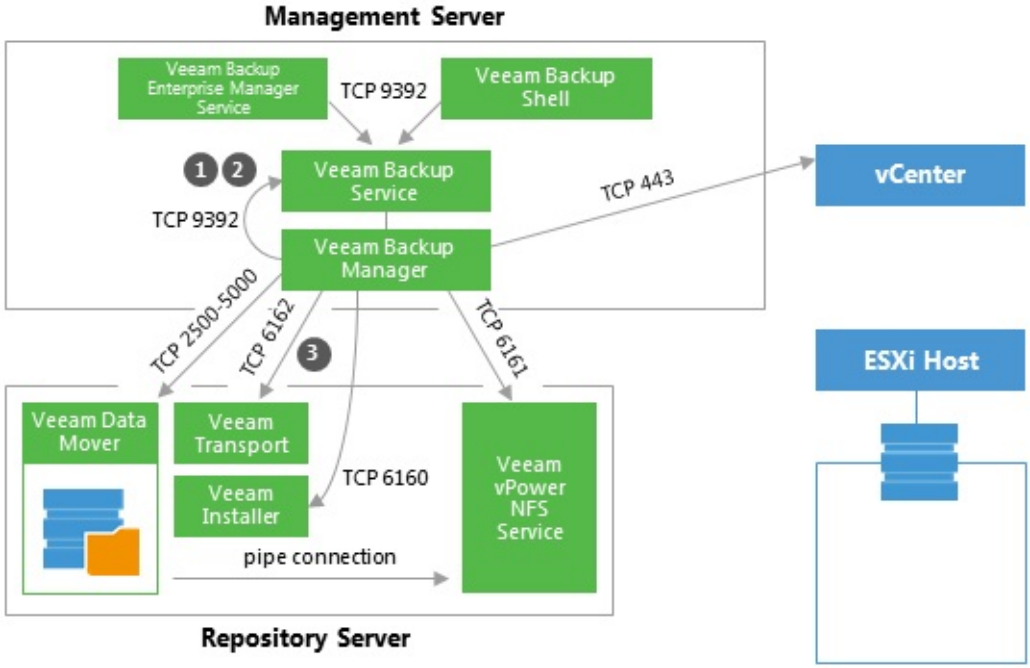
Instant VM Recovery

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:

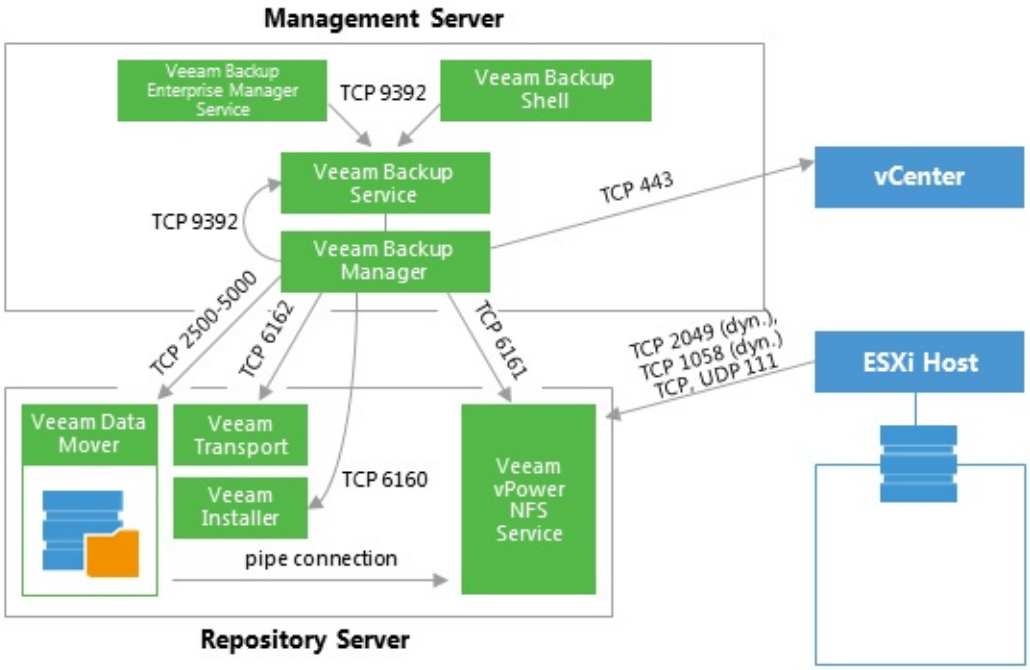
1. Starts the Veeam Backup Manager process on the Veeam backup server.
2. Checks with the Veeam Backup Service whether the necessary backup infrastructure resources are available for instant VM Recovery.
3. Communicates with the Transport Service on the backup repository to start Veeam Data Mover.



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.

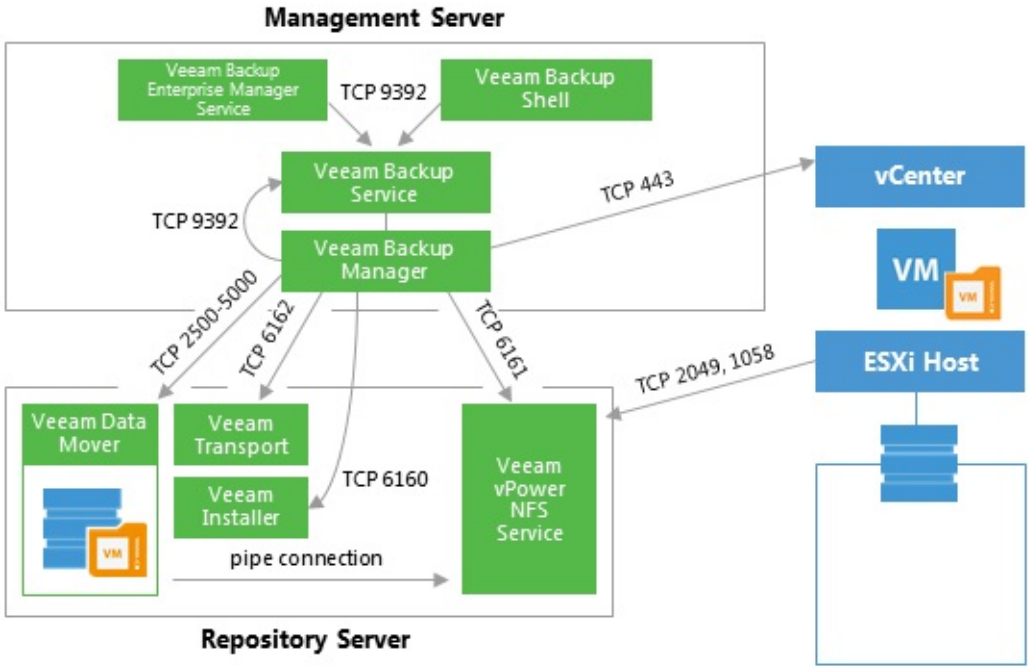
Next, Veeam Backup & Replication creates in the Veeam NFS datastore VM configuration files and links to virtual disk files. Virtual disk files remain in the backup on the repository, while all changes to these files are written to the cache file.



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.


Management Server
Veeam Backup Enterprise Manager Service
TCP 9392
Veeam Backup Shell
Veeam Backup Service
TCP 9392
Veeam Backup Manager
TCP 443
vCenter
VM
TCP 2500-5000
TCP 6162
TCP 6161
TCP 2049, 1058
ESXi Host
Veeam Data Mover
Veeam Transport
Veeam Installer
TCP 6160
Veeam vPower NFS Service
pipe connection
Repository Server

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

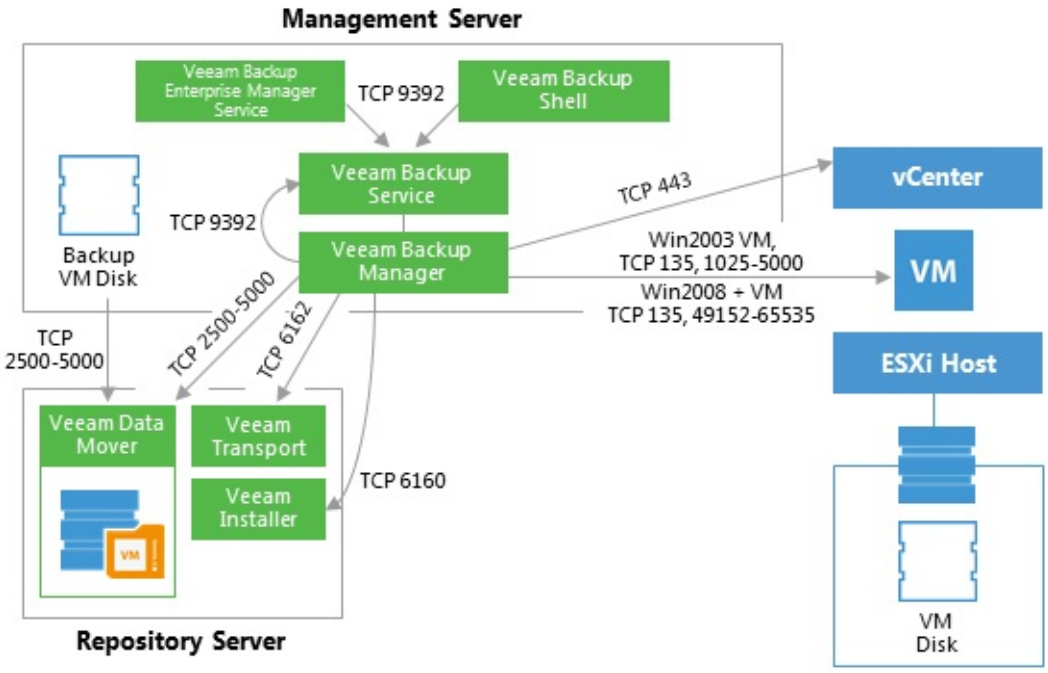
In the initialization phase, Veeam Backup & Replication prepares resources necessary for Microsoft Windows file-level restore. It performs the following steps:

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:\VeeamFLR* folder (can be changed via registry key).

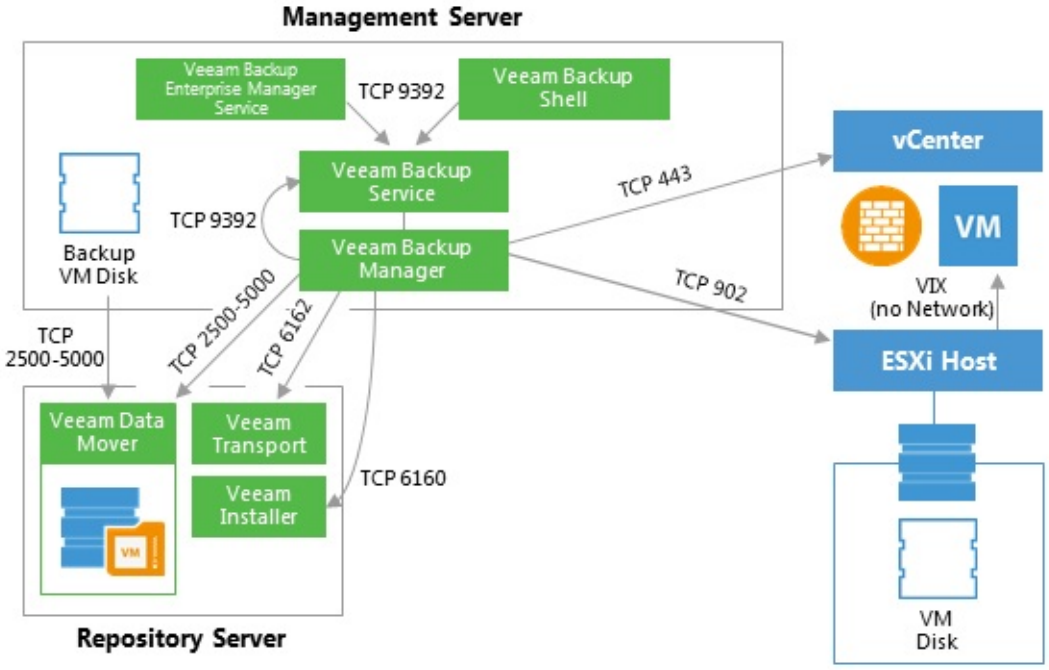
2a. Restoring Windows Guest OS Files (Network-Based)

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.

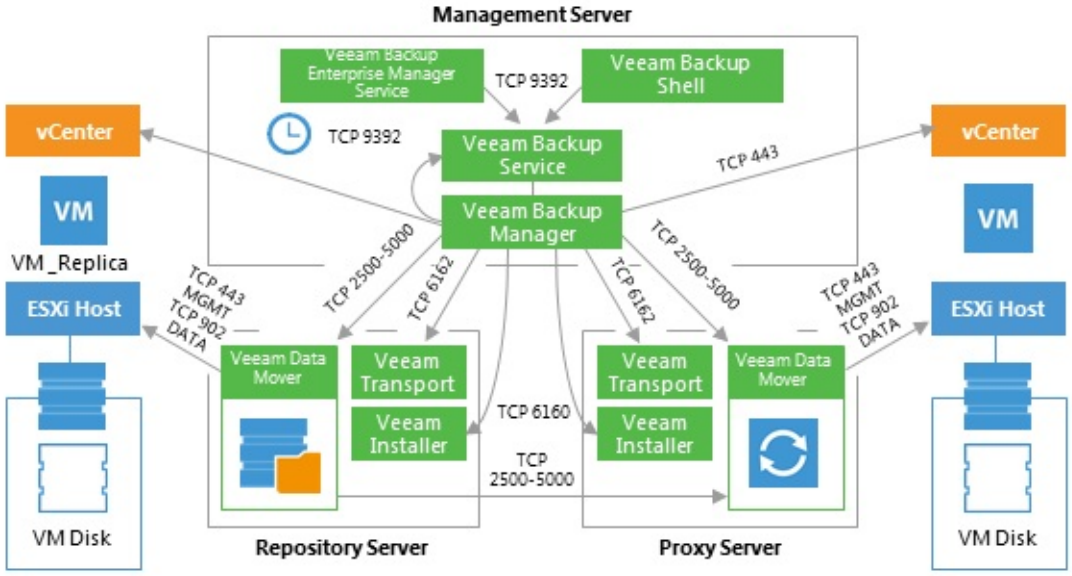
Replication

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.

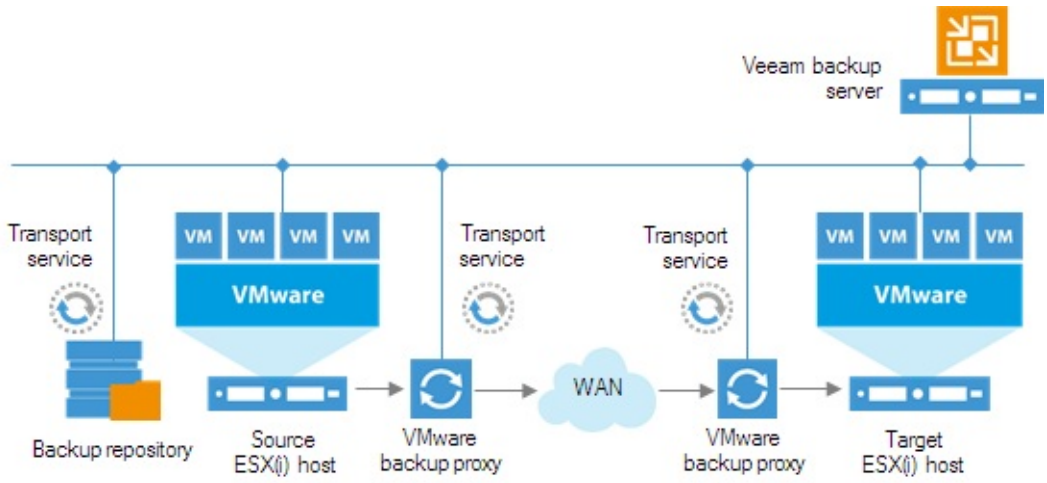
Next, Veeam Backup & Replication performs in-guest processing tasks, triggers VM snapshot creation, registers a replica VM on the target host and performs data transfer from the source host and datastore to the target host and datastore. The source and target proxies can use one of 3 available data transport 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/NFS Access and HotAdd transport modes, see the “Backup Anatomy” section above in this Appendix.



Note that Veeam uses backup repository to store replica metadata.

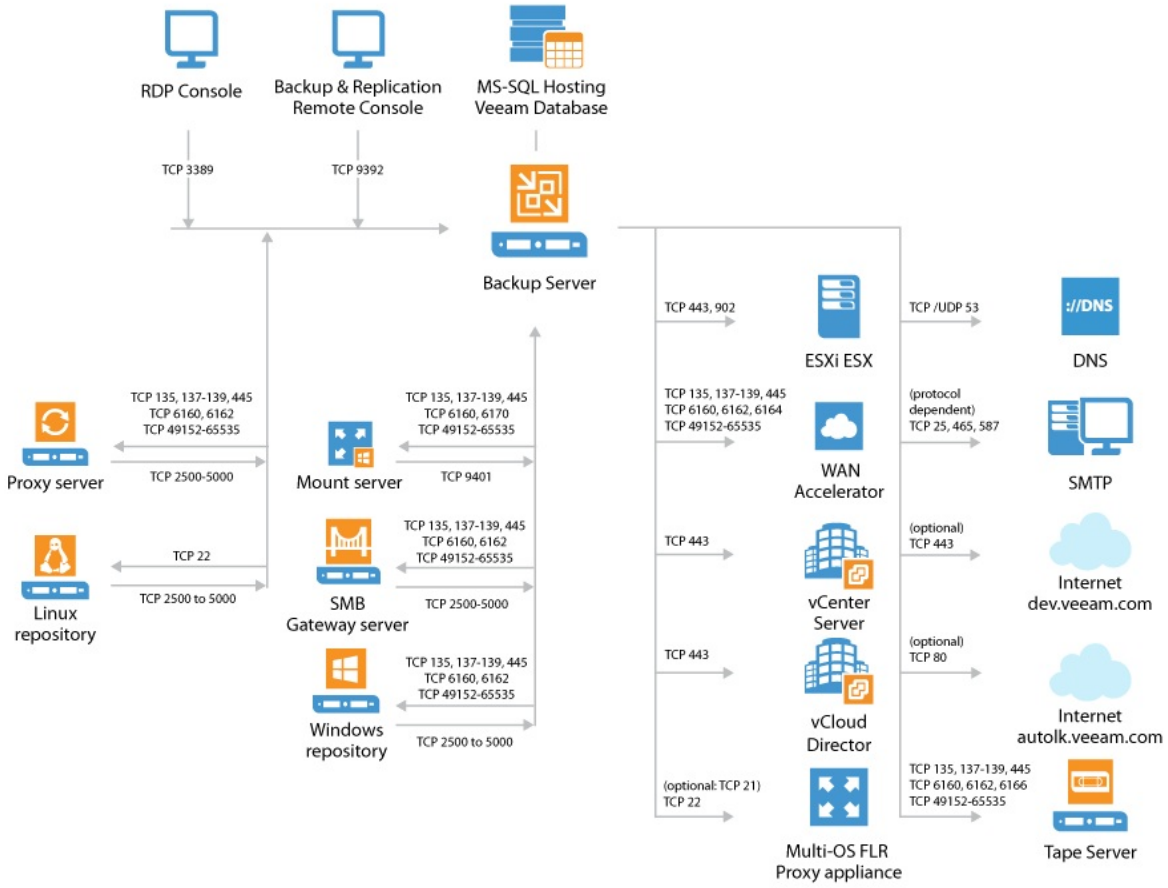
The following diagram illustrates a possible placement of the Veeam Backup & Replication components in a distributed environment, with a WAN link between the production and DR sites.


Veeam backup server
Transport service
VM
VM
VM
VM
VMware
Transport service
Transport service
VM
VM
VM
VM
VMware
WAN
Backup repository
Source ESX(i) host
VMware backup proxy
VMware backup proxy
Target ESX(i) host

Networking Diagrams

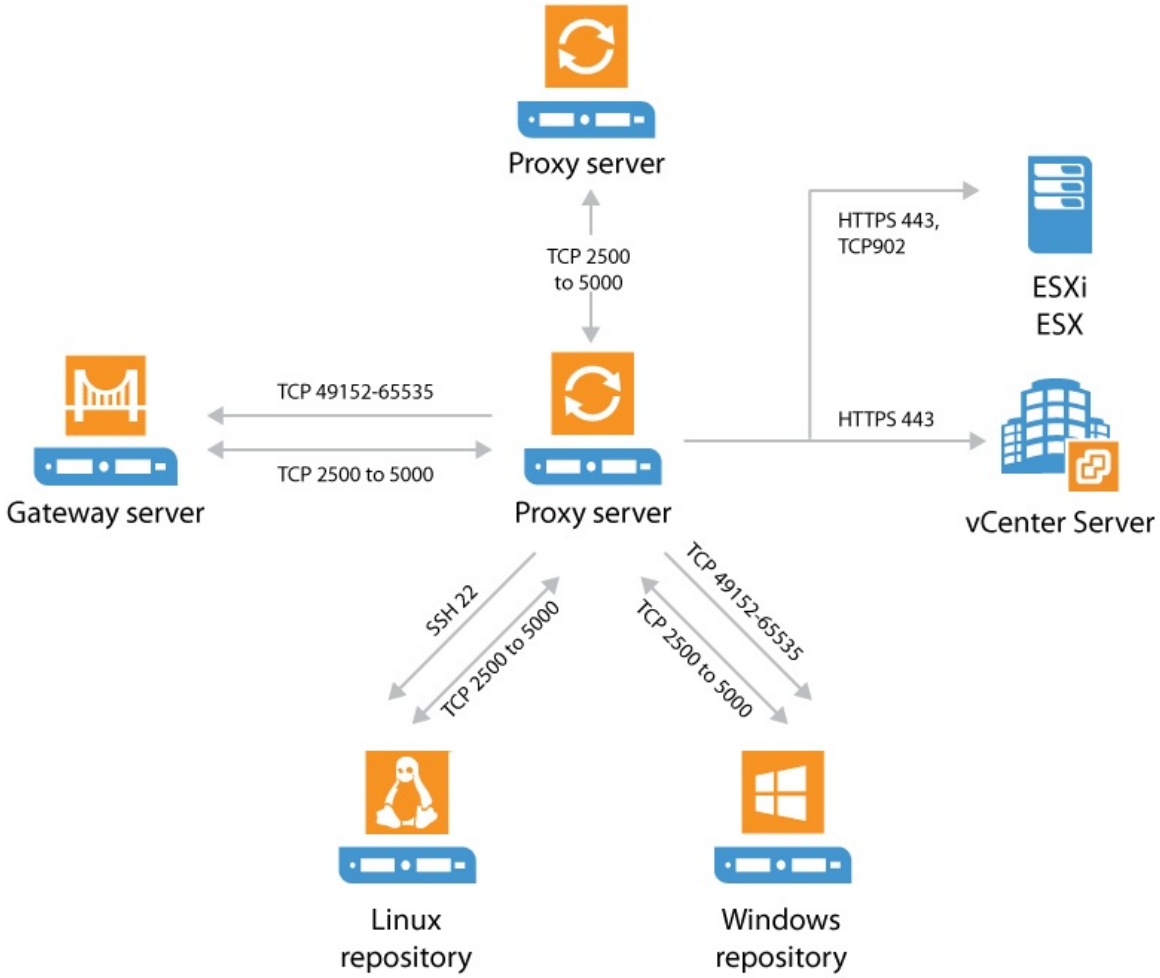
There is a detailed list of ports used by Veeam Backup & Replication available in the User Guide, but sometimes a more visual approach is helpful – you can use the diagrams below for that purpose.

Backup Server



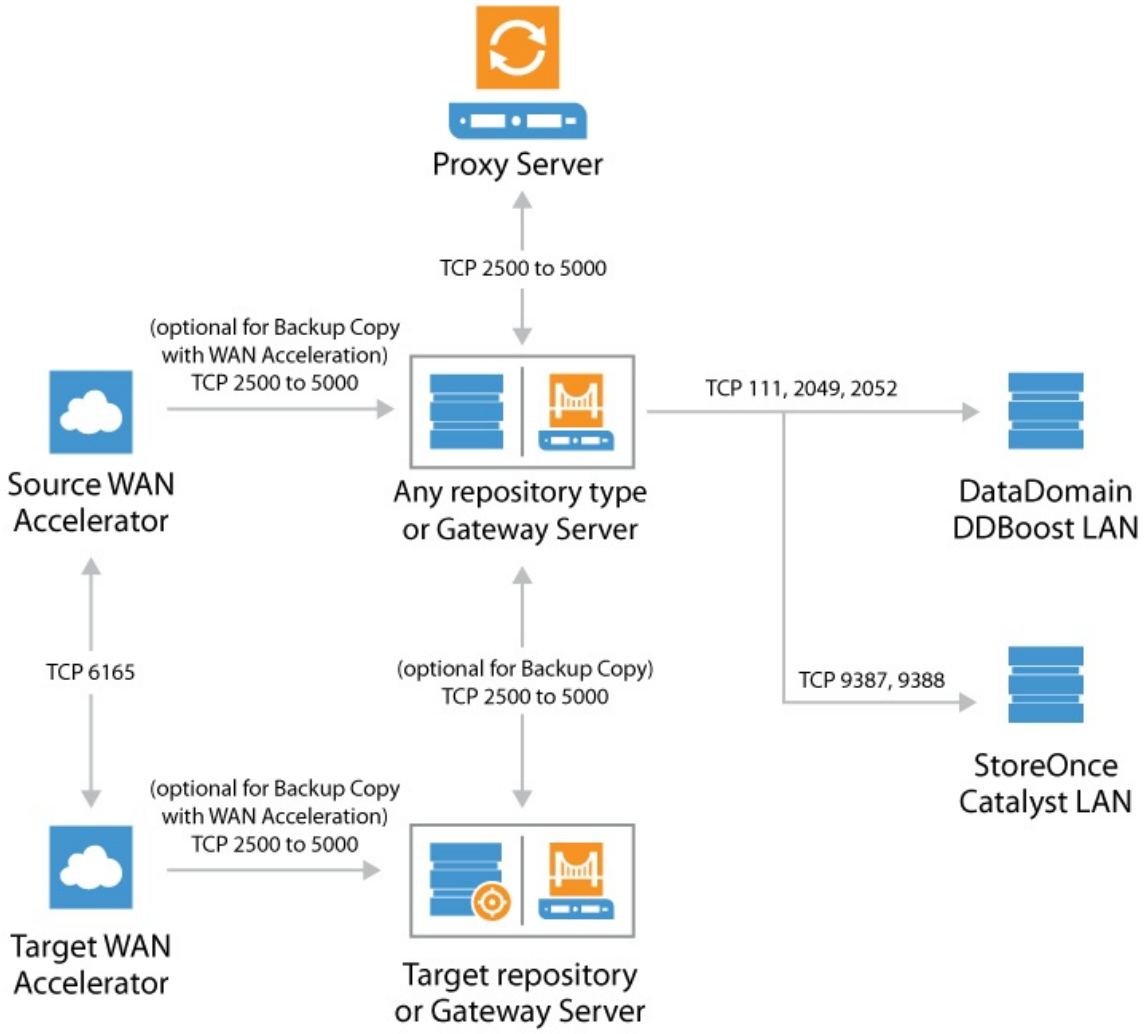
Proxy Server

The following ports are required for the proxy server.

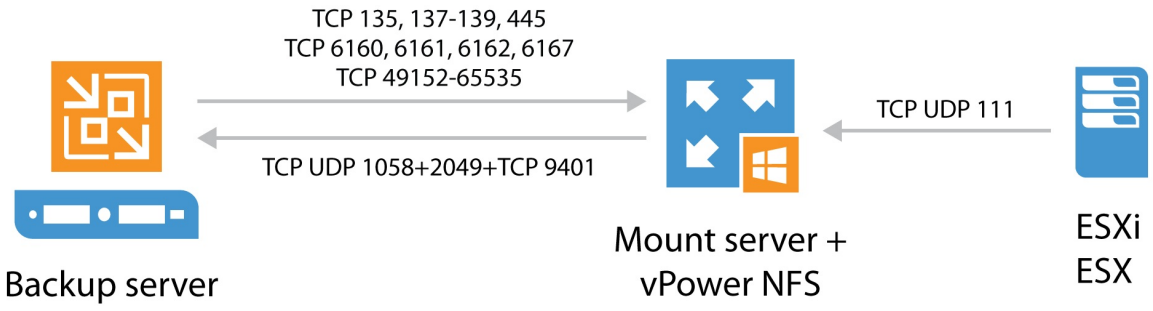


Repository Server

The following ports are required for the repository server.

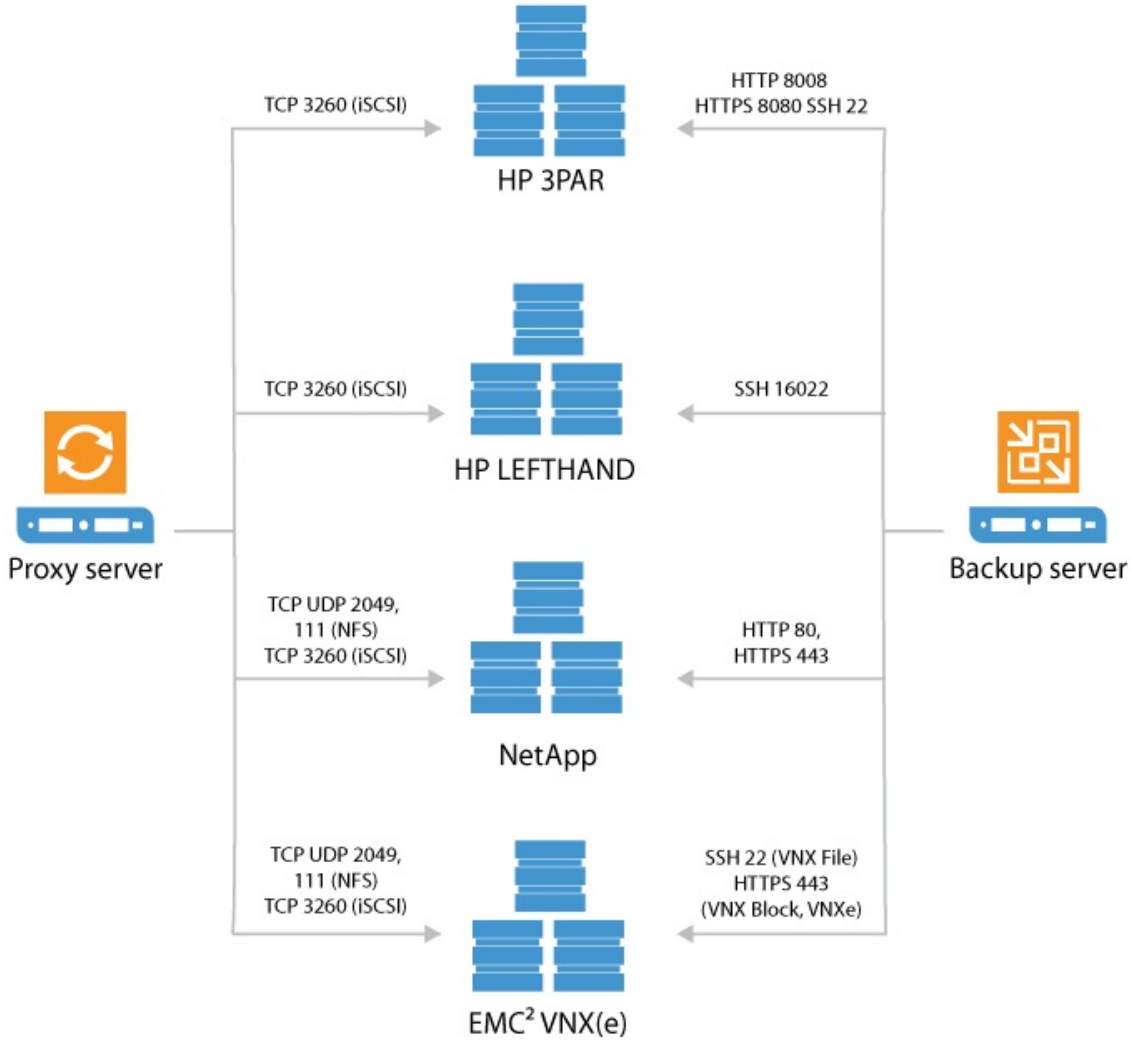


The following ports are required for vPower NFS.



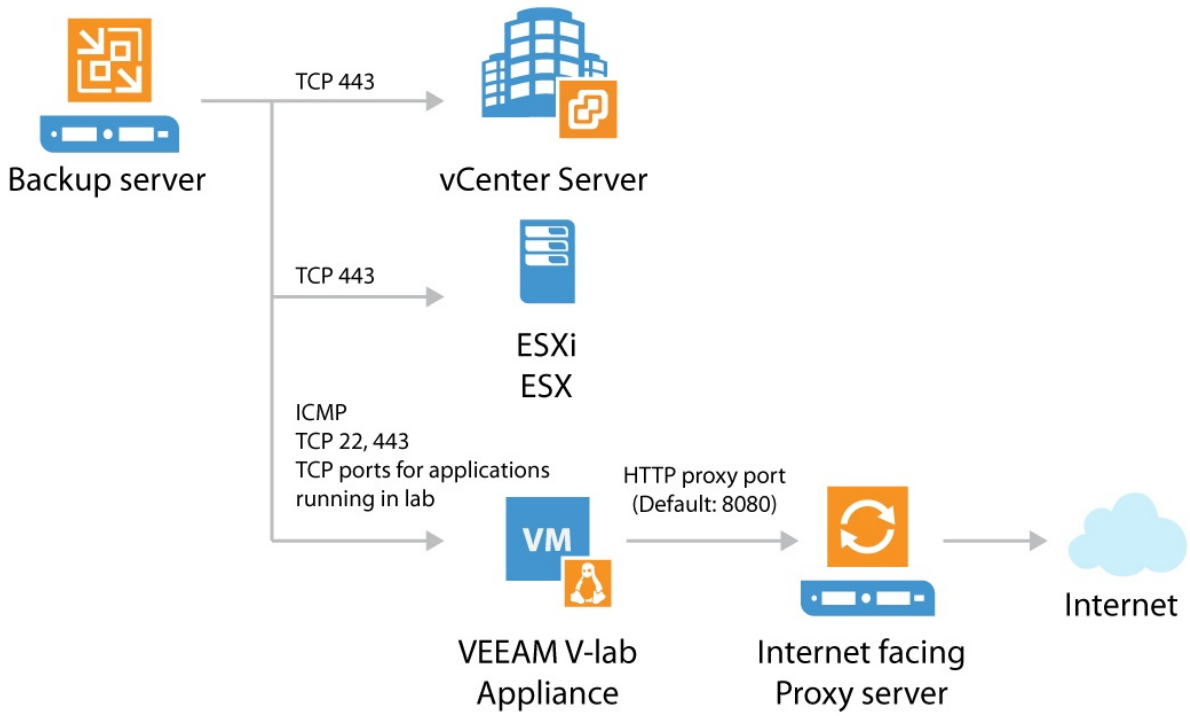
Storage Integrations

The following ports are required for integrated storage.



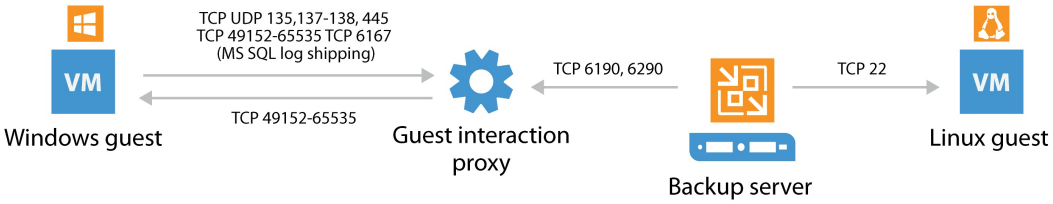
Data Validation

The following ports are required, when using SureBackup, SureReplica, or On-demand Sandbox from Storage Snapshots.



Application-aware Image Processing

The following ports are required for application-aware image processing over the network. If network ports are not available, the backup server will failover to using VIX via VMware Tools.



Enterprise Manager

The following ports are required for Enterprise Manager