

What Is the Journal?

Every write to a protected virtual machine is copied by Zerto Virtual Replication. The write continues to be processed normally on the protected site and the copy is sent asynchronously to the recovery site and written to a journal managed by a Virtual Replication Appliance (VRA). Each protected virtual machine has its own journal.

In addition to the writes, every few seconds all journals are updated with a checkpoint time-stamp. Checkpoints are used to ensure write order fidelity and crash-consistency. A recovery can be done to the last checkpoint or to a user-selected, crash-consistent checkpoint. This enables recovering the virtual machines, either to the last crash-consistent point-in-time or, for example, when the virtual machine is attacked by a virus, to a point-in-time before the virus attack.

Data and checkpoints are written to the journal until the specified journal history size is reached, which is the optimum situation. At this point, as new writes and checkpoints are written to the journal, the older writes are written to the recovery virtual machine's virtual disks. When specifying a checkpoint to recover to, the checkpoint must still be present in the journal. For example, if the value specified is 24 hours, then recovery can be made to any checkpoint up to 24 hours in the past. After the time specified, the mirror virtual disk volumes maintained by the VRA are updated.

During recovery, the virtual machines at the recovery site are created and the recovery disks for each virtual machine, managed by the VRA, are attached to the recovered virtual machines. Information in the journal is promoted to the virtual machines to bring them up to the date and time of the selected checkpoint. To improve the RTO during recovery, the virtual machines can be used even before the journal data has been fully promoted. Every request is analyzed and the response is returned from the virtual machine directly or, if the information in the journal is more up-to-date, it comes from the journal. This continues until the recovery site's virtual environment is fully restored to the selected checkpoint.

Each protected virtual machine has its own dedicated journal, consisting of one or more volumes. A dedicated journal enables journal data to be maintained, even when changing the host for the recovery. The default datastore, or storage, used for a journal when protecting to a vCenter Server or a SCVMM is the storage used for recovery of each virtual machine. Thus for example, if protected virtual machines in a VPG are configured with different recovery storage, the journal data is by default stored for each virtual machine on that virtual machine recovery storage. The default datastore used for a journal when protecting to a vCloud Director is the datastore with the most free space, that has either been defined as a journal datastore for the provider vDC, in the `Configure Provider vDCs` dialog or any datastore visible to the recovery host if the journal datastore was not defined in the `Configure Provider vDCs` dialog.

Defining the Journal

The journals are defined as part of the virtual protection group (VPG) definition. The definition includes VPG-level settings and VM-level settings.

VPG-Level Settings

The VPG-level journal settings are the default settings for all VMs within the VPG:

- **Journal History** - Determines the maximum amount of history that can be saved in a journal, assuming that the storage hosting the journal has enough space. Saving more history requires more space. The period of time can be between one hour and 14 days.
- **Default Journal Storage or Datastore** - The location of the journal volumes. By default, this is the target storage used for the recovery of each virtual machine itself, but can be any storage accessible by the recovery host.
- **Journal Size Hard Limit** - The maximum size that the journal can grow, either as a percentage or a fixed amount. The minimum journal size, set by Zerto Virtual Replication, is 8GB. You can specify the journal size. The maximum size can be either a percentage of the virtual machine volume size or can be a fixed maximum size.

Unlimited - The size of the journal is unlimited and it can grow to the size of the recovery datastore.

Size (GB) - The maximum journal size in GB.

- **Percentage** – The percentage of the virtual machine volume size the journal can grow to.
- **Journal Size Warning Threshold** parameter. The size of the journal that triggers a warning that the journal is nearing its hard limit.
- **Unlimited** – The size of the journal is unlimited and it can grow to the size of the recovery datastore.
- **Size (GB)** – The size in GB that will generate a warning.
- **Percentage** – The percentage of the virtual machine volume size that will generate a warning.

Zerto recommends that the journal storage should be accessible by all the recovery hosts and not just by one of the hosts.

Note: Changes to the VPG-level settings are not applied to virtual machines already defined in the VPG, only to virtual machines added to the VPG after the changes.

VM-Level Settings

You can modify VM-level settings to override the VPG-level settings for the journal for each virtual machine in the journal:

- **VM Recovery Datastore (vSphere) or Storage (Hyper-V)** – This is the virtual machine recovery datastore. It is also the default datastore used for the journal. The datastore must be accessible by the host specified in the `Recovery Host` parameter.
- **Journal Size Hard Limit**
- **Journal Size Warning Threshold**
- **Journal Datastore (vSphere) or Journal Storage (Hyper-V)** – The storage used by the journal for the virtual machine when not using the `VM Recovery Datastore (vSphere)` or `Storage (Hyper-V)`.

Journal Behavior

After defining a VPG, the protected virtual machine disks are synced with the recovery site. After initial synchronization, each write to a protected virtual machine is copied, asynchronously sent to the recovery site, and then written to the journal of the recovery virtual machine. Every few seconds all journals are updated with a checkpoint time-stamp. The last checkpoint written to the journal is used to establish the recovery point objective (RPO).

Checkpoints are marks in the journal history that enable recovery to a specific time. All recent checkpoints are displayed in the `Select Recovery Point` dialog; old checkpoints are available but the time between those displayed increases as they are older. The actual history stored is measured from the present to the oldest checkpoint.

Data and checkpoints are written to the journal until the specified journal history size is reached, which is the optimum situation. At this point, as new writes and checkpoints are written to the journal, the older writes are written to the virtual machines' recovery virtual disks.

Journal Size Changes

The journal is always thin-provisioned and the data passed from the protected site to the recovery site is saved in the journal in a compressed format. The provisioned journal size is the current size of all the journal volumes and the amount of space initially allocated is 16GB.

Note: The data is compressed when it is passed over the network to the recovery side and uncompressed when a recovery operation is required. This uses less space without impacting the RTO. Compression is enabled automatically when WAN compression is enabled for a VPG.

If the journal grows to approximately 80% of the provisioned journal size or less than 6GB remains free, a new volume is added to increase the journal size. Each new journal volume added is bigger than the previous volume. The journal size can increase up until a specified hard limit. If the hard limit of the journal is reduced in the VPG definition after new volumes have been added, these volumes are not removed and continue to be used if required. In this case, the journal size can be bigger than the defined hard limit and the reduced journal size definition is not applied, except to ensure that no new volumes are created if the new journal hard limit is reached or exceeded.

When the amount of the journal used is approximately 50% of the provisioned journal size, the biggest unused journal volume from the added volumes is marked for removal. This volume is then removed after the time equivalent to three times the amount specified for the journal history, or twenty-four hours, whichever is more if it is still not used.

In a vSphere environment:

- With a VMFS datastore and when the VRA is on a host ESXi that is version 5.1 or higher, the journal can also reclaim unused space on a volume.
- Reclaiming space on a volume does not change the provisioned journal size, which is the current size of all the journal volumes. Also, unused space is not reclaimed when using NFS datastores or any storage with a host with a version less than 5.1.

Journal Behavior When Nearing the Journal Size Hard Limit

When the size of a virtual machine journal approaches its specified hard limit, Zerto Virtual Replication starts to move data to the target disks. Once this begins, the maintained history begins to decrease. If the journal history falls below 75% of the value specified for the journal history, a warning alert is issued in the GUI. If the history falls below one hour, an error is issued. However, if the amount of history defined is only one hour, an error is issued if it is less than 45 minutes.

If the storage where the journal resides drops below 30GB or 15% of the total storage size, whichever is smaller, the storage itself is considered full and an error alert is issued and all writes to journal volumes on that storage are blocked. Replication is halted, but history is not lost. As such, the RPO begins to steadily increase until space is made available on the storage. Thus, with a large storage where 30GB is less than 15%, the alert will be triggered when the amount of free space is less than 30GB. With a small storage where 30GB is more than 15%, the alert will be triggered when the amount of free space is less than 15% of the total size.

If the journal is not big enough to store all the data for the time specified, as defined in the *Journal Size Hard Limit* parameter, the time frame for storing data is reduced. When the journal reaches the limit specified in the *Journal Size Warning Threshold* parameter, an alert is issued.

If a virtual machine needs a larger journal size than the hard limit specified to accommodate the journal history, you must manually adjust the size of the *Journal Size Hard Limit* parameter.

When are Checkpoints in a Journal Not Usable?

The following actions cause the journal history to be reset:

- Removing a virtual machine from the VPG
- Adding a volume to a virtual machine in a VPG
- Removing a volume from a virtual machine in a VPG

The journal is rebuilt starting with no history.

Note: Changing any of the following is performed transparently, using VMware Storage vMotion or Hyper-V Quick Storage Migration, without requiring that the journal be rebuilt:

- The recovery storage of a volume.
- The VPG journal storage.
- The virtual machine recovery storage.

Determining Where to Store the Journal

Each protected virtual machine has its own dedicated journal, consisting of one or more volumes. A dedicated journal enables journal data to be maintained even when changing the target host used for recovery.

The journals for the protected virtual machines are defined as part of the VPG and, by default, are defined to reside on the same storage as the virtual machine. This can be overridden at the VPG level to allow storage tiering.

The following table shows the different journal storage options and the consequences of each.

	Allows Storage Tiering	Notes
Default Journal	No	The journal is located on the virtual machine's recovery storage. By default, the recovery storage for each virtual machine is the same as the virtual machine recovery storage.
Journal Datastore/ Storage for Each VPG	Yes	Specify a journal storage for each VPG. All journals for the virtual machines in the VPG are stored in this storage.
Journal Datastore/ Storage for Multiple VPGs	Yes	Enables the use of advanced settings such as storage IO controls etc., to provide individualized service to customers by grouping VPGs by customer and assigning each group to a specific storage. This option is recommended for cloud providers.

Journal Sizing

In general, assuming a 10% change rate per day and four hours of journal history saved, a 15GB limit is large enough to support a virtual machine with 1TB of storage. Use the Journal Sizing Tool to make better journal hard limit size estimates for each virtual machine.

Note: Assuming a constant rate of change rate per day, a four hour history requires half the space of an eight hour history. Larger histories also require longer promotion times for moves or failover operations, impacting performance. If more space is required over time than is available, warnings and errors are issued. In this case you should increase the journal's hard limit size.

When defining a VPG, estimate the amount of storage required at the recovery site for the journal. Use the Journal Sizing Tool to estimate the size limit. The journal is thin-provisioned, therefore the actual size of the journal increases up to the maximum in order to accommodate the increasing size of the actual history.

Testing Considerations When Determining Journal Size

When a VPG is tested, either during a failover test or before committing a Move or Failover operation, a scratch volume is created for each virtual machine being tested. The scratch volume created uses the same size limit defined for the journal for each virtual machine.

The size limit of the scratch volume determines the length of time that you can test. Larger limits enable longer testing times if the rate of change is constant. If a small hard limit size is set for this amount of history, for example 2 - 3 hours, the scratch volume created for testing will also be small, thus limiting the time available for testing.

Note: The limit for the scratch volume cannot be increased during testing.

Thus, when considering the journal size limit you must also consider the length of time required to test the VPG. You must specify a limit for the journal accordingly, or specify that the size is unlimited, in which case it is only limited by the storage size.

Estimating a Virtual Machine Journal Size

Use the Zerto Journal Sizing Tool to:

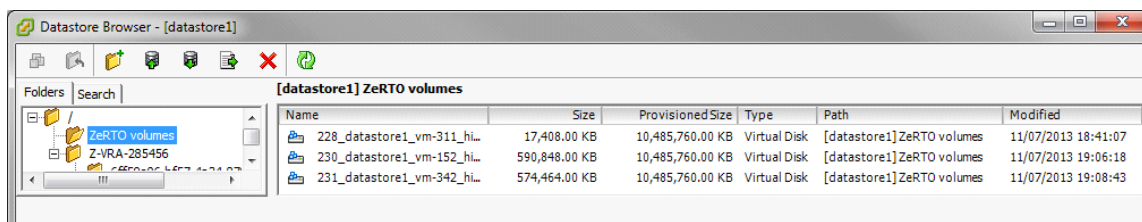
- Estimate the journal size hard limit that is required.
- Estimate the total storage size that is required for each storage used on the recovery site.

To estimate the journal size with the Journal Sizing Tool:

1. Open the `Zerto Journal Sizing Tool.xlsx` file.
2. Enter the data in the VMs sheet row by row for each protected virtual machine. If a virtual machine has more than one volume, use a separate row for each volume defined for the virtual machine. The Journal Sizing Tool's data columns are:
 - VM** - The name of the protected virtual machine.
 - VPG** - The name of the VPG that includes the virtual machine.
 - Size (GB)** - The size of the virtual machine volumes.
 - Recovery Datastore/Storage** - The name of the storage used to recover/replicate the source storage.
 - Journal Datastore/Storage** - The name of the storage used for the journal.
 - Daily Change Rate (%)** - The expected change rate over a 24 hour period.
 - History (hours)** - The amount of history, in hours, required to be saved in the journal.
 - Required Journal (GB)** - The required journal size in GB, which is calculated automatically. This is the amount of space required for the journal. For space requirements for testing, continue with the next steps.
3. Fill in the testing considerations in the Testing Simulator sheet, row by row for each VPG, to determine additional space requirements for testing the VPG either during a failover test or before committing a move or failover operation. This sheet enables you to simulate a testing policy, testing all the protected virtual machines in parallel or one VPG at a time, or something between these two extremes.
4. Fill in the data in the Test Scratch Volume Definitions sheet with the testing change rate estimated over 24 hours and the time expected to be used for testing.
 - Test Change Rate (daily)** - The expected change rate during the test over a 24 hour period.
 - Test Duration (hours)** - The time allotted to carry out the testing.
5. View the results in the Datastore Usage sheet to assess the size requirements for each datastore/storage you use, both for recovery of protected data and for the additional space required when a journal resides on the same storage.

vSphere Environments Only: Monitoring Journal VMDK Usage

The actual size of the thin journal VMDKs created for each VM can be obtained from the datastore browser by looking at the `ZeRTO volumes` folder on the datastore. When using ESXi 5.1 and later, the used size of this VMDK displays the current journal usage.



Name	Size	Provisioned Size	Type	Path	Modified
228_datastore1_vm-311_hi...	17,408.00 KB	10,485,760.00 KB	Virtual Disk	[datastore1]ZeRTO volumes	11/07/2013 18:41:07
230_datastore1_vm-152_hi...	590,848.00 KB	10,485,760.00 KB	Virtual Disk	[datastore1]ZeRTO volumes	11/07/2013 19:06:18
231_datastore1_vm-342_hi...	574,464.00 KB	10,485,760.00 KB	Virtual Disk	[datastore1]ZeRTO volumes	11/07/2013 19:08:43

Impact on a Journal of Resynchronizing a VPG

When a VPG is resynchronized, for example after the WAN or the recovery site host is down, older checkpoints in the journal are removed. During the resynchronization, the delta changes in the source/protection site are added to the journal and older data in the journal is moved to the target/recovery virtual disk managed by the VRA for the virtual machine.

As the resynchronization continues and more old data is moved out of the journal, the checkpoints associated with the data are also removed from the journal **but new checkpoints are not added to it**. If the time to complete synchronization is so long that all the existing checkpoints would be removed, the last few checkpoints are kept so that recovery operations can still be performed.

Looking at the contents of a journal file after a long synchronization has completed can include a few checkpoints from before the synchronization started followed by a gap without any checkpoints and then a normal journal with checkpoints added every few seconds.

Note: Synchronizations do not usually last so long that all the checkpoints are removed from the journal.

Feedback

Zerto is constantly looking for ways to improve all aspects of its business and your feedback is very important to us. Please send any comments you have about this document, to help us to improve it, to support@zerto.com.

ABOUT ZERTO

Zerto is committed to keeping enterprise and cloud IT running 24/7 by providing scalable business continuity software solutions. Through the Zerto Cloud Continuity Platform, organizations seamlessly move and protect virtualized workloads between public, private and hybrid clouds. The company's flagship product, Zerto Virtual Replication, is the standard for protection of applications in cloud and virtualized datacenters.

www.zerto.com

For further assistance using Zerto Virtual Replication, contact Zerto support at support@zerto.com.