



Red Hat Gluster Storage 3

3.0 Update 4 Release Notes

Release Notes for Red Hat Storage - 3.0 Update 4

Red Hat Gluster Storage 3 3.0 Update 4 Release Notes

Release Notes for Red Hat Storage - 3.0 Update 4

Pavithra Srinivasan

Red Hat Engineering Content Services

psriniva@redhat.com

Legal Notice

Copyright © 2014-2015 Red Hat Inc.

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](#). If you distribute this document, or a modified version of it, you must provide attribution to Red Hat, Inc. and provide a link to the original. If the document is modified, all Red Hat trademarks must be removed.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux ® is the registered trademark of Linus Torvalds in the United States and other countries.

Java ® is a registered trademark of Oracle and/or its affiliates.

XFS ® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL ® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Abstract

This Release Notes provides high-level coverage of the improvements and additions that have been implemented in Red Hat Storage 3.0 Update 4.

Table of Contents

CHAPTER 1. INTRODUCTION	3
CHAPTER 2. WHAT'S NEW IN THIS RELEASE?	4
CHAPTER 3. KNOWN ISSUES	5
3.1. RED HAT STORAGE	5
3.2. RED HAT STORAGE CONSOLE	24
3.3. RED HAT STORAGE AND RED HAT ENTERPRISE VIRTUALIZATION INTEGRATION	31
3.4. RED HAT STORAGE AND RED HAT OPENSTACK INTEGRATION	32
CHAPTER 4. TECHNOLOGY PREVIEWS	33
4.1. GSTATUS COMMAND	33
4.2. STRIPED VOLUMES	33
4.3. DISTRIBUTED-STRIPED VOLUMES	33
4.4. DISTRIBUTED-STRIPED-REPLICATED VOLUMES	34
4.5. STRIPED-REPLICATED VOLUMES	34
4.6. REPLICATED VOLUMES WITH REPLICA COUNT GREATER THAN 3	34
4.7. STOP REMOVE BRICK OPERATION	34
4.8. READ-ONLY VOLUME	34
4.9. NFS GANESHA	35
4.10. NON UNIFORM FILE ALLOCATION	35
APPENDIX A. REVISION HISTORY	36

CHAPTER 1. INTRODUCTION

Red Hat Storage is a software only, scale-out storage solution that provides flexible and agile unstructured data storage for the enterprise. Red Hat Storage provides new opportunities to unify data storage and infrastructure, increase performance, and improve availability and manageability to meet a broader set of the storage challenges and needs of an organization.

GlusterFS, a key building block of Red Hat Storage, is based on a stackable user space design and can deliver exceptional performance for diverse workloads. GlusterFS aggregates various storage servers over different network interfaces and connects them to form a single large parallel network file system. The POSIX compliant GlusterFS servers use XFS file system format to store data on disks. These servers be accessed using industry standard access protocols including Network File System (NFS) and Server Message Block SMB (also known as CIFS).

Red Hat Storage Servers for On-premise can be used in the deployment of private clouds or data centers. Red Hat Storage can be installed on commodity servers and storage hardware resulting in a powerful, massively scalable, and highly available NAS environment. Additionally, Red Hat Storage can be deployed in the public cloud using Red Hat Storage Server for Public Cloud, for example, within the Amazon Web Services (AWS) cloud. It delivers all the features and functionality possible in a private cloud or data center to the public cloud by providing massively scalable and high available NAS in the cloud.

Red Hat Storage Server for On-Premise

Red Hat Storage Server for On-Premise enables enterprises to treat physical storage as a virtualized, scalable, and centrally managed pool of storage by using commodity servers and storage hardware.

Red Hat Storage Server for Public Cloud

Red Hat Storage Server for Public Cloud packages GlusterFS as an Amazon Machine Image (AMI) for deploying scalable NAS in the AWS public cloud. This powerful storage server provides a highly available, scalable, virtualized, and centrally managed pool of storage for Amazon users.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

CHAPTER 2. WHAT'S NEW IN THIS RELEASE?

This chapter describes the key features and changes in the Red Hat Storage 3.0 update 4 release.

- **Support for Just a Bunch of Disks**

- Red Hat Storage supports Just a Bunch Of Disks (JBOD) on servers with up to 24 disks. The complexity of RAID is avoided by using 3-way replication with JBOD.

- **Three-way replication**

- Synchronous three-way replication is now fully supported in Red Hat Storage volumes. Three-way replicated volumes do not require hardware or software RAID. Three-way replication yields best results when used in conjunction with JBODs. Quorum functionality on three-way replicated volumes helps to prevent split-brain scenarios. 3-way replicated volumes provide additional resilience in public cloud deployments as well.

Three-way replication is also supported for Hadoop Deployments.

- **Performance Enhancements**

- **Event threads:** Two new performance enhancement options namely `event-thread.client` and `event-thread.server` are introduced. Configuring appropriate values minimizes the effect of slow throughput and response time in a Red Hat Storage trusted storage pool.
- **Virtual Memory parameters:** Two new virtual memory parameters can be configured to optimize performance of random, small file workloads and large file sequential workloads.

- **Change in the location of the Gluster command log file**

Gluster commands executed on a node in the Red Hat Storage Truster Storage Pool is logged in the file `/var/log/glusterfs/cmd_history.log`. It is no longer a hidden file.

- **Samba-3 is deprecated from the Red Hat Storage 3.0 Update 4 release**

The Samba version 3 is being deprecated from RHS 3.0 Update 4. Further updates will not be provided for Samba-3.x. It is recommended that you upgrade to Samba-4.x, which is provided in a separate channel or repository, for all updates including the security updates.

RHN channel: `rhel-x86_64-server-6-rh-gluster-3-samba`

CDN repo: `rh-gluster-3-samba-for-rhel-6-server-rpms`

- **The gluster CLI sub package is deprecated from the Red Hat Enterprise Linux optional channel**

The `glusterfs-cli` sub-package is removed from the product listing for `rhel-optional` channel or repository. No further versions of the sub-package will be shipped to the `rhel-optional` channel or repository.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

CHAPTER 3. KNOWN ISSUES

This chapter provides a list of known issues at the time of release.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

3.1. RED HAT STORAGE

Issues related to Snapshot

- BZ# [1202388](#)

After restoring a snapshot, attaching a new node to a Red Hat Storage Trusted Storage Pool fails due to checksum mismatch.

This failure occurs in following scenarios:

- If quota is enabled on the volume and a snapshot is taken, the system does not recompute the quota checksum when the snapshot is restored. This leads to a checksum mismatch and the new node goes into a rejected state.

Workaround: After snapshot restore, `glusterd` service must be restarted on existing nodes and the peer probe has to be executed on the new node.

- If there are different volume options configured in the origin volume and the snapshot volume, and a new node is added, subsequently the same snapshot is restored on the volume, the restored volume will have a different checksum in the new node. Restarting of any node in the Red Hat Storage Trusted Storage Pool results in the new node to move to peer-reject state.

Workaround: Copy the `/var/lib/glusterd/vols/VOLNAME/info` file from existing node to new node and restart the `glusterd` on new node.

Workaround: Ensure that you do not perform any file system operations on the snapshot that is about to be deleted.

- BZ# [1201820](#)

When a snapshot is deleted, the corresponding file system object in the User Serviceable Snapshot is also deleted. Any subsequent file system access results in the snapshot Daemon becoming unresponsive.

Workaround: Ensure that you do not perform any file system operations on the snapshot that is about to be deleted.

- BZ# [1191033](#)

When a large number of snapshots are activated and are accessed via User Serviceable Snapshots, the inode limit for the mounted filesystem may get exhausted, causing the deletion of inode entries from the inode table. Accessing the `.snaps` directory when this happens may result in `ENOENT` errors in FUSE, and `Invalid Argument` errors in NFS.

Workaround: Clear the kernel VFS cache by executing the following command:

```
# echo 3 >/proc/sys/vm/drop_caches
```

- [BZ# 1176835](#)

When the User Serviceable Snapshot feature is enabled, performing any operation or system call that uses the `statfs()` or `statvfs()` system calls on a `.snaps` repository fails with the following error: **Stale file handle**

- [BZ# 1160621](#)

If the current directory is not a part of the snapshot, for example, `snap1`, then the user cannot enter the `.snaps/snap1` directory.

- [BZ# 1169790](#)

When a volume is down and there is an attempt to access `.snaps` directory, a negative cache entry is created in the Kernel Virtual File System(VFS) cache for the `.snaps` directory. After the volume is brought back online, accessing the `.snaps` directory fails with an `ENOENT` error because of the negative cache entry.

Workaround: Clear the kernel VFS cache by executing the following command:

```
# echo 3 >/proc/sys/vm/drop_caches
```

- [BZ# 1170145](#)

After the restore operation is complete, if restore a volume while you are in the `.snaps` directory, the following error message is displayed from the mount point - **"No such file or directory"**.

Workaround:

1. Navigate to the parent directory of the `.snaps` directory.
2. Drop VFS cache by executing the following command:

```
# echo 3 >/proc/sys/vm/drop_caches
```

3. Change to the `.snaps` folder.

- [BZ# 1170365](#)

Virtual inode numbers are generated for all the files in the `.snaps` directory. If there are hard links, they are assigned different inode numbers instead of the same inode number.

- [BZ# 1170502](#)

On enabling the User Serviceable Snapshot feature, if a directory or a file by name `.snaps` exists on a volume, it appears in the output of the `ls -a` command.

- [BZ# 1174618](#)

If the User Serviceable Snapshot feature is enabled, and a directory has a pre-existing `.snaps` folder, then accessing that folder can lead to unexpected behavior.

Workaround: Rename the pre-existing `.snaps` folder with another name.

- [BZ# 1167648](#)

Performing operations which involve client graph changes such as volume set operations, restoring snapshot etc eventually leads to out of memory scenarios for the client processes which mount the volume.

- [BZ# 1141433](#)

An incorrect output is displayed while setting the *snap-max-hard-limit* and *snap-max-soft-limit* options for volume or system/cluster.

Example:

- Command: `gluster snapshot config snap-max-hard-limit 100`

Output : *snapshot config : System for snap-max-hard-limit set successfully*

Expected Output : *snapshot config : snap-max-hard-limit for the cluster set successfully.*

- Command: `gluster snapshot config vol1 snap-max-hard-limit 100`

Output : *snapshot config : vol1 for snap-max-hard-limit set successfully.*

Expected output : *snapshot config : snap-max-hard-limit for vol1 set successfully.*

The same example is applicable for *snap-max-soft-limit* option also.

- [BZ# 1133861](#)

New snap bricks fails to start if the total snapshot brick count in a node goes beyond 1K.

Workaround: Deactivate unused snapshots.

- [BZ# 1126789](#)

If any node or `glusterd` service is down when snapshot is restored then any subsequent snapshot creation fails.

Workaround: Do not restore a snapshot, if node or `glusterd` service is down.

- [BZ# 1122064](#)

The snapshot volumes does not handshake based on versions. If a node or `glusterd` service is down when snapshot activate or deactivate command executed, the node on which the command was executed is not updated when the node or `glusterd` service is up and continues to be in the same state.

- [BZ# 1139624](#)

While taking snapshot of a gluster volume, it creates another volume which is similar to the original volume. Gluster volume consumes some amount of memory when it is in started state, so as Snapshot volume. Hence, the system goes to out of memory state.

Workaround: Deactivate unused snapshots to reduce the memory foot print.

- [BZ# 1129675](#)

If glusterd is down in one of the nodes in cluster or if the node itself is down, then performing a snapshot restore operation leads to the inconsistencies:

- Executing `gluster volume heal vol-name info` command displays the error message *Transport endpoint not connected*
- Error occurs when clients try to connect to glusterd service.

Workaround: Perform snapshot restore only if all the nodes and their corresponding glusterd services are running.

Restart glusterd service using `# service glusterd start` command.

- [BZ# 1105543](#)

When a node with old snap entry is attached to the cluster, the old entries are propagated throughout the cluster and old snapshots which are not present are displayed.

Workaround: Do not attach a peer with old snap entries.

- [BZ# 1104191](#)

The Snapshot command fails if snapshot command is run simultaneously from multiple nodes when high write or read operation is happening on the origin or parent volume.

Workaround: Avoid running multiple snapshot commands simultaneously from different nodes.

- [BZ# 1059158](#)

The `NFS mount` option is not supported for snapshot volumes.

- [BZ# 1114015](#)

A wrong warning message, *Changing snapshot-max-hard-limit will lead to deletion of snapshots if they exceed the new limit. Do you want to continue? (y/n)* is displayed while setting the configuration options (`snap-max-hard-limit`, `snap-max-soft-limit`) for system or volume. The snapshot will not be deleted even if configuration values are changed.

- [BZ# 1113510](#)

The output of `gluster volume info` information (`snap-max-hard-limit` and `snap-max-soft-limit`) even though the values that are not set explicitly and must not be displayed.

- [BZ# 1111479](#)

Attaching a new node to the cluster while snapshot delete was in progress, deleted snapshots successfully but gluster snapshot list shows some of the snaps are still present.

Workaround: Do not attach or detach new node to the trusted storage pool operation while snapshot is in progress.

- [BZ# 1092510](#)

If you create a snapshot when the rename of directory is in progress (here, its complete on hashed subvolume but not on all of the subvolumes), on snapshot restore, directory which was undergoing rename operation will have same GFID for both source and destination. Having

same GFID is an inconsistency in DHT and can lead to undefined behavior.

This is since in DHT, a rename (source, destination) of directories is done first on hashed-subvolume and if successful, then on rest of the subvolumes. At this point in time, if you have both source and destination directories present in the cluster with same GFID - destination on hashed-subvolume and source on rest of the subvolumes. A parallel lookup (on either source or destination) at this time can result in creation of directories on missing subvolumes - source directory entry on hashed and destination directory entry on rest of the subvolumes. Hence, there would be two directory entries - source and destination - having same GFID.

- [BZ# 1104635](#)

During snapshot delete, if a node goes down, the snapshot delete command fails. Stale entries would be present in the node which went down and when the node comes back online, the stale entry is propagated to other nodes and this results an invalid snapshot entry which may not be deletable using CLI.

Workaround: Manually delete the snapshot, including the back-end LVM, from all the nodes and restart `glusterd` service on all nodes.

Issues related to Nagios

- [BZ# 1136207](#)

Volume status service shows *All bricks are Up* message even when some of the bricks are in unknown state due to unavailability of `glusterd` service.

- [BZ# 1109683](#)

When a volume has a large number of files to heal, the `volume self heal info` command takes time to return results and the `nrpe` plug-in times out as the default timeout is 10 seconds.

Workaround:

In `/etc/nagios/gluster/gluster-commands.cfg` increase the timeout of `nrpe` plug-in to 10 minutes by using the `-t` option in the command.

Example: `$USER1$/gluster/check_vol_server.py $ARG1$ $ARG2$ -o self-heal -t 600`

- [BZ# 1094765](#)

When certain commands invoked by Nagios plug-ins fail, irrelevant outputs are displayed as part of performance data.

- [BZ# 1107605](#)

Executing `sadf` command used by the Nagios plug-ins returns invalid output.

Workaround: Delete the datafile located at `/var/log/sa/saDD` where DD is current date. This deletes the datafile for current day and a new datafile is automatically created and which is usable by Nagios plug-in.

- [BZ# 1107577](#)

The Volume self heal service returns a WARNING when there unsynchronized entries are present in the volume, even though these files may be synchronized during the next run of self-heal process if `self-heal` is turned on in the volume.

- [BZ# 1109744](#)

Notification message is sent only once regarding quorum loss and not for each volume.

- [BZ# 1121009](#)

In Nagios, CTDB service is created by default for all the gluster nodes regardless of whether CTDB is enabled on the Red Hat Storage node or not.

- [BZ# 1089636](#)

In the Nagios GUI, incorrect status information is displayed as *Cluster Status OK : None of the Volumes are in Critical State*, when volumes are utilized beyond critical level.

- [BZ# 1109739](#)

Cluster quorum monitoring Nagios plug-in displays only the latest messages received, so the user is unable to determine all the volumes for which quorum is lost.

Workaround:

Event log contains the list of the messages received. Also, if quorum is lost on a cluster, all volumes with `quorum-type` server will be affected.

- [BZ# 1079289](#)

When the memory utilization is very high, some or all the services may go to critical state and display the message: *CHECK_NRPE: Socket timeout after 10 seconds*.

- [BZ# 1106421](#)

When `cluster.quorum-type` is set to `none` for all volumes in the cluster, the Cluster quorum monitoring plug-in only receives passive checks based on `rsyslog` messages and hence remains in Pending state as there are no service notifications available.

- [BZ# 1085331](#)

Volume Utilization graph displays the error *perfddata directory for host_directory for host_name does not exist*, when volume utilization data is not available.

- [BZ# 1111828](#)

In Nagios GUI, Volume Utilization graph displays an error when volume is restored using its snapshot.

Issues related to Rebalancing Volumes

- [BZ# 1110282](#)

Executing `rebalance status` command, after stopping rebalance process, fails and displays a message that the rebalance process is not started.

- [BZ# 1140531](#)

Extended attributes set on a file while it is being migrated during a rebalance operation are lost.

Workaround: Reset the extended attributes on the file once the migration is complete.

- Rebalance does not proceed if any of the subvolumes of dht in the volume are down. This could be any brick in the case of a pure distributed volume. In a distributed replicated set, rebalance will proceed as long as at least one brick of each replica set is up.

While running rebalance on a volume, ensure that all the bricks of the volume are in the operating or connected state.

- BZ# [960910](#)

After executing `rebalance` on a volume, running the `rm -rf` command on the mount point to remove all of the content from the current working directory recursively without being prompted may return *Directory not Empty* error message.

- BZ# [862618](#)

After completion of the rebalance operation, there may be a mismatch in the failure counts reported by the `gluster volume rebalance status` output and the rebalance log files.

- BZ# [1039533](#)

While Rebalance is in progress, adding a brick to the cluster displays an error message, `failed to get index` in the gluster log file. This message can be safely ignored.

- BZ# [1064321](#)

When a node is brought online after rebalance, the status displays that the operation is completed, but the data is not rebalanced. The data on the node is not rebalanced in a remove-brick rebalance operation and running commit command can cause data loss.

Workaround: Run the `rebalance` command again if any node is brought down while rebalance is in progress, and also when the rebalance operation is performed after remove-brick operation.

Issues related to Geo-replication

- BZ# [1204044](#)

The script `stop-all-gluster-processes.sh` fails to stop the Geo-replication processes.

Workaround: If the script fails on a Red Hat Storage node, manually search for the `gsyncd` processes and kill them.

- BZ# [1179701](#)

If a master volume of a Red Hat Storage node goes into an irrecoverable state and it is replaced by a new node, the new node does not have changelogs because of which it performs a file system crawl operation and misses to sync a few files from the master volume.

- BZ# [1102524](#)

The Geo-replication worker goes to faulty state and restarts when resumed. It works as expected when it is restarted, but takes more time to synchronize compared to resume.

Issues related to Self-heal

- BZ# [877895](#)

When one of the bricks in a replicate volume is offline, the `ls -lR` command from the mount point reports *Transport end point not connected*

When one of the two bricks under replication goes down, the entries are created on the other brick. The Automatic File Replication translator remembers that the directory that is down contains stale data. If the brick that is online is killed before the self-heal happens on that directory, operations like `readdir()` fail.

- [BZ# 1063830](#)

Performing add-brick or remove-brick operations on a volume having replica pairs when there are pending self-heals can cause potential data loss.

Workaround: Ensure that all bricks of the volume are online and there are no pending self-heals. You can view the pending heal info using the command `gluster volume heal volname info`.

Issues related to replace-brick operation

- After the `gluster volume replace-brick VOLNAME Brick New-Brick commit force` command is executed, the file system operations on that particular volume, which are in transit, fail.
- After a replace-brick operation, the stat information is different on the NFS mount and the FUSE mount. This happens due to internal time stamp changes when the `replace-brick` operation is performed.

Issues related to Directory Quota

- [BZ# 1001453](#)

Truncating a file to a larger size and writing to it violates the quota hard limit. This is since the XFS pre-allocation logic applied on the truncated file does not extract the actual disk space it consumed.

- [BZ# 1003755](#)

Directory Quota feature does not work well with hard links. With a directory that has Quota limit set, the disk usage seen with the `du -hs directory` command and the disk usage seen with the `gluster volume quota VOLNAME list directory` command may differ. It is recommended that applications writing to a volume with directory quotas enabled, do not use hard links.

- [BZ# 1016419](#)

Quota does not account for the disk blocks consumed by a directory. Even if a directory grows in size because of the creation of new directory entries, the size as accounted by quota does not change. You can create any number of empty files but you will not be able to write to the files once you reach the quota hard limit. For example, if the quota hard limit of a directory is 100 bytes and the disk space consumption is exactly equal to 100 bytes, you can create any number of empty files without exceeding quota limit.

- [BZ# 1020275](#)

Creating files of different sizes leads to the violation of the quota hard limit.

- [BZ# 1021466](#)

After setting Quota limit on a directory, creating sub directories and populating them with files and renaming the files subsequently while the I/O operation is in progress causes a quota limit violation.

- BZ# [998893](#)

Zero byte sized files are created when a write operation exceeds the available quota space. Since Quota does not account for the disk blocks consumed by a directory(as per Bug 1016419), the write operation creates the directory entry but the subsequent write operation fails because of unavailable disk space.

- BZ# [1023430](#)

When a quota directory reaches its limit, renaming an existing file on that directory leads to Quota violation. This is because the renamed is treated as a new file.

- BZ# [998791](#)

During a file rename operation if the hashing logic moves the target file to a different brick, then the rename operation fails if it is initiated by a non-root user.

- BZ# [999458](#)

Quota hard limit is violated for small quota sizes in the range of 10 MB to 100 MB.

- BZ# [1020713](#)

In a distribute or distribute replicate volume, while setting quota limit on a directory, if one or more bricks or one or more replica sets respectively, experience downtime, quota is not enforced on those bricks or replica sets, when they are back online. As a result, the disk usage exceeds the quota limit.

Workaround: Set quota limit again after the brick is back online.

- BZ# [1032449](#)

In the case when two or more bricks experience a downtime and data is written to their replica bricks, invoking the quota list command on that multi-node cluster displays different outputs after the bricks are back online.

Issues related to NFS

- BZ# [1203901](#) and BZ# [1203850](#)

Multi-client sequential write operations on the file system leads to the Red Hat Storage node to become unresponsive. See the [Knowledge Centric Solution](#) for more information on the steps to proceed further.

- After you restart the NFS server, the unlock within the grace-period feature may fail and the locks help previously may not be reclaimed.
- `fcntl` locking (NFS Lock Manager) does not work over IPv6.
- You cannot perform NFS mount on a machine on which glusterfs-NFS process is already running unless you use the NFS mount `-o nolock` option. This is because glusterfs-nfs has already registered NLM port with portmapper.

- If the NFS client is behind a NAT (Network Address Translation) router or a firewall, the locking behavior is unpredictable. The current implementation of NLM assumes that Network Address Translation of the client's IP does not happen.
- `nfs.mount-udp` option is disabled by default. You must enable it to use posix-locks on Solaris when using NFS to mount on a Red Hat Storage volume.
- If you enable the `nfs.mount-udp` option, while mounting a subdirectory (exported using the `nfs.export-dir` option) on Linux, you must mount using the `-o proto=tcp` option. UDP is not supported for subdirectory mounts on the GlusterFS-NFS server.
- For NFS Lock Manager to function properly, you must ensure that all of the servers and clients have resolvable hostnames. That is, servers must be able to resolve client names and clients must be able to resolve server hostnames.
- BZ# [1040418](#)

The length of the argument to `nfs.export-dir` (or any other gluster set option) is limited to internal buffer size of the Linux shell. In a typical set up, the default size of this buffer is 131072 bytes.

Issues related to nfs-ganesha

- BZ# [1116374](#)

The `nfs-ganesha` daemon crashes if started in `NIV_FULL_DEBUG` level.

Workaround: Use other log levels supported by `nfs-ganesha` while starting the server.

- BZ# [1116336](#)

The `nfs-ganesha` process remains active after setting `nfs-ganesha.enable` to `off` as executing `kill -s TERM` command does not kill `nfs-ganesha`.

Workaround: Use `kill -9` on the process ID of `ganesha.nfsd` process and then use CLI options to export new entries.

- BZ# [1115901](#)

Multi-node `nfs-ganesha` is not supported in this release.

Workaround: In a multi-node volume setup, perform all CLI commands and steps on one of the nodes only.

- BZ# [1114574](#)

Executing `rpcinfo -p` command after stopping `nfs-ganesha` displays NFS related programs

Workaround: Run `rpcinfo -d` command each of the NFS related services listed in `rpcinfo -p` and start the Red Hat Storage volume forcefully using the following command:

```
# gluster vol start volume force
```

- BZ# [1091936](#)

When ACL support is enabled, `getattr` of ACL attribute on the files with no ACLs set return value as `NULL`. This leads to discrepancies while trying to read ACLs on the files present in the

system.

- [BZ# 1054124](#)

After files and directories are created on the mount point with root-squash enabled for `nfs-ganesha`, executing `ls` command displays `user : group` as `4294967294 : 4294967294` instead of `nfsnobody : nfsnobody`. This is because the client maps only 16 bit unsigned representation of -2 to `nfsnobody` whereas 4294967294 is 32 bit equivalent of -2. This is currently a limitation in upstream `nfs-ganesha 2.0` and will be fixed in the future release.

- [BZ# 1054739](#)

As multiple sockets are used with `nfs-ganesha`, executing `showmount -e` command displays duplicate information.

Issues related to Object Store

- The GET and PUT commands fail on large files while using Unified File and Object Storage.

Workaround: You must set the `node_timeout=60` variable in the proxy, container, and the object server configuration files.

Issues related to distributed Geo-replication

- [BZ# 984591](#)

After stopping a Geo-replication session, if the files synced to the slave volume are renamed then when Geo-replication starts again, the renamed files are treated anew, (without considering the renaming) and synced on to the slave volumes again. For example, if 100 files were renamed, you would find 200 files on the slave side.

- [BZ# 987929](#)

While the `rebalance` process is in progress, starting or stopping a Geo-replication session results in some files not get synced to the slave volumes. When a Geo-replication sync process is in progress, running the `rebalance` command causes the Geo-replication sync process to stop. As a result, some files do not get synced to the slave volumes.

- [BZ# 1029799](#)

Starting a Geo-replication session when there are tens of millions of files on the master volume takes a long time to observe the updates on the slave mount point.

- [BZ# 1026072](#)

The Geo-replication feature keeps the status details including the `changeLog` entries in the `/var/run/gluster` directory. On Red Hat Storage Server, this directory is a `tmpfs` mountpoint, therefore there is a data loss after a reboot.

- [BZ# 1027727](#)

When there are hundreds of thousands of hard links on the master volume prior to starting the Geo-replication session, some hard links are not getting synchronized to the slave volume.

- [BZ# 1029899](#)

During a Geo-replication session, after you set the checkpoint, and subsequently when one of the active nodes goes down, the passive node replaces the active node. At this point the checkpoint for replaced node is displayed as invalid.

- [BZ# 1030256](#)

During a Geo-replication session, when create and write operations are in progress, if one of the active nodes goes down, there is a possibility for some files to undergo a synchronization failure to the slave volume.

- [BZ# 1063028](#)

When geo-replication session is running between master and slave, ACLs on the master volume are not reflected on the slave as ACLs (which are extended attributes) are not synced to the slave by Geo replication.

- [BZ# 1056226](#)

User-set xattrs are not synced to the slave as Geo-replication does not process **SETXATTR** fops in changelog (and in hybrid crawl).

- [BZ# 1063229](#)

After the upgrade, two Geo-replication monitor processes run for the same session. Both process try to use the same `xsync changelog` file to record the changes.

Workaround: Before running `geo-rep create force` command, kill the Geo-replication monitor process.

Issues related to Red Hat Storage Volumes:

- [BZ# 877988](#)

Entry operations on replicated bricks may have a few issues with `md-cache` module enabled on the volume graph.

For example: When one brick is down, while the other is up an application is performing a hard link call `link()` would experience `EEXIST` error.

Workaround: Execute this command to avoid this issue:

```
# gluster volume set VOLNAME stat-prefetch off
```

- [BZ# 986090](#)

Currently, the Red Hat Storage server has issues with mixed usage of hostnames, IPs and FQDNs to refer to a peer. If a peer has been probed using its hostname but IPs are used during `add-brick`, the operation may fail. It is recommended to use the same address for all the operations, that is, during peer probe, volume creation, and adding/removing bricks. It is preferable if the address is correctly resolvable to a FQDN.

- [BZ# 882769](#)

When a volume is started, by default the NFS and Samba server processes are also started automatically. The simultaneous use of Samba or NFS protocols to access the same volume is not supported.

Workaround: You must ensure that the volume is accessed either using Samba or NFS protocols.

- BZ# [852293](#)

The management daemon does not have a rollback mechanism to revert any action that may have succeeded on some nodes and failed on the those that do not have the brick's parent directory. For example, setting the `volume-id` extended attribute may fail on some nodes and succeed on others. Because of this, the subsequent attempts to recreate the volume using the same bricks may fail with the error *brickname or a prefix of it is already part of a volume*

Workaround:

1. You can either remove the brick directories or remove the glusterfs-related extended attributes.
2. Try creating the volume again.

- BZ# [994950](#)

An input-output error is seen instead of the Disk quota exceeded error when the quota limit exceeds. This issue is fixed in the Red Hat Enterprise Linux 6.5 Kernel.

- BZ# [913364](#)

An NFS server reboot does not reclaim the file LOCK held by a Red Hat Enterprise Linux 5.9 client.

- BZ# [896314](#)

GlusterFS Native mount in Red Hat Enterprise Linux 5.x shows lower performance than the RHEL 6.x versions for high burst I/O applications. The FUSE kernel module on Red Hat Enterprise Linux 6.x has many enhancements for dynamic write page handling and special optimization for large burst of I/O.

Workaround: It is recommended that you use Red Hat Enterprise Linux 6.x clients if you observe a performance degradation on the Red Hat Enterprise Linux 5.x clients.

- BZ# [1017728](#)

On setting the quota limit as a decimal digit and setting the `deem-statfs` on, a difference is noticed in the values displayed by the `df -h` command and `gluster volume quota VOLNAME list` command. In case of the `gluster volume quota VOLNAME list` command, the values do not get rounded off to the next integer.

- BZ# [1030438](#)

On a volume, when read and write operations are in progress and simultaneously a rebalance operation is performed followed by a remove-brick operation on that volume, then the `rm -rf` command fails on a few files.

- BZ# [1100590](#)

The `cp -a` operation from the NFS mount point hangs if barrier is already enabled.

Issues related to POSIX ACLs:

- Mounting a volume with `-o acl` can negatively impact the directory read performance. Commands like recursive directory listing can be slower than normal.
- When POSIX ACLs are set and multiple NFS clients are used, there could be inconsistency in the way ACLs are applied due to attribute caching in NFS. For a consistent view of POSIX ACLs in a multiple client setup, use the `-o noac` option on the NFS mount to disable attribute caching. Note that disabling the attribute caching option could lead to a performance impact on the operations involving the attributes.

Issues related to Samba

- BZ# [1013151](#)

Accessing a Samba share may fail if GlusterFS is updated while Samba is running.

Workaround: On each node where GlusterFS is updated, restart Samba services after GlusterFS is updated.

- BZ# [994990](#)

When the same file is accessed concurrently by multiple users for reading and writing. The users trying to write to the same file will not be able to complete the write operation because of the lock not being available.

Workaround: To avoid the issue, execute the command:

```
# gluster volume set VOLNAME storage.batch-fsync-delay-usec 0
```

- BZ# [1031783](#)

If Red Hat Storage volumes are exported by Samba, NT ACLs set on the folders by Microsoft Windows clients do not behave as expected.

Issues related to multi-threaded epoll

BZ# [1200679](#)

Setting the value of `client.event-threads` or `server.event-threads` values to a lower value reduces the number of epoll threads only when subsequent IO operations are performed on the mount point of the volume.

Issues related to sos

BZ# [1190997](#)

The `sosreport` command does not collect the gluster command logs without flags or options in the `/var/log/glusterfs/cmd_history.log` file.

Workaround: Use the `sosreport -a` command to collect logs in the `/var/log/glusterfs/cmd_history.log` file.

General issues

- If files and directories have different GFIDs on different back-ends, the glusterFS client may hang or display errors.

Contact Red Hat Support for more information on this issue.

- BZ# [920002](#)

The POSIX compliance tests fail in certain cases on Red Hat Enterprise Linux 5.9 due to mismatched timestamps on FUSE mounts. These tests pass on all the other Red Hat Enterprise Linux 5.x and Red Hat Enterprise Linux 6.x clients.

- BZ# [916834](#)

The `quick-read` translator returns stale file handles for certain patterns of file access. When running the `dbench` application on the mount point, a `dbench: read fails on handle 10030` message is displayed.

Work Around: Use the command below to avoid the issue:

```
# gluster volume set VOLNAME quick-read off
```

- BZ# [1030962](#)

On installing the Red Hat Storage Server from an ISO or PXE, the `kexec-tools` package for the `kdump` service gets installed by default. However, the `crashkernel=auto` kernel parameter required for reserving memory for the `kdump` kernel, is not set for the current kernel entry in the bootloader configuration file, `/boot/grub/grub.conf`. Therefore the `kdump` service fails to start up with the following message available in the logs.

```
kdump: No crashkernel parameter specified for running kernel
```

Workaround: After installing the Red Hat Storage Server, the `crashkernel=auto`, or an appropriate `crashkernel=sizeM` kernel parameter can be set manually for the current kernel in the bootloader configuration file. After that, the Red Hat Storage Server system must be rebooted, upon which the memory for the `kdump` kernel is reserved and the `kdump` service starts successfully. Refer to the following link for more information on [Configuring kdump on the Command Line](#)

Additional information: On installing a new kernel after installing the Red Hat Storage Server, the `crashkernel=auto` kernel parameter is successfully set in the bootloader configuration file for the newly added kernel.

- BZ# [866859](#)

The `sosreport` behavior change (to `glusterfs` and `sosreport`) is altered in the `statedump` behavior configuration file (`glusterdump.optionsfile`) and it is placed in `/tmp`. This file has information on the path and options you can set on the behavior of the `statedump` file. The `glusterfs` daemon searches for this file and subsequently places the `statedump` information in the specified location. **Workaround:** Change the configurations in `glusterfs` daemon to make it look at `/usr/local/var/run/gluster` for `glusterdump.optionsfile` by default. No changes to be performed to make `sosreport` write its configuration file in `/usr/local/var/run/gluster` instead of `/tmp`.

- BZ# [1054759](#)

A `vdsmd` crash report is detected by Automatic Bug Reporting Tool (ABRT) in Red Hat Storage Node as the `/etc/vdsm/vdsm.id` file was not found during the first time.

Workaround: Execute the command `/usr/sbin/dmidecode -s system-uuid > /etc/vdsm/vdsm.id` before adding the node to avoid the `vdsmd` crash report.

- [BZ# 1058032](#)

While migrating VMs, libvirt changes the ownership of the guest image, unless it detects that the image is on a shared filesystem and the VMs can not access the disk images as the required ownership is not available.

Workaround: Perform the steps:

1. Power-off the VMs before migration.
2. After migration is complete, restore the ownership of the VM Disk Image (107:107)
3. Start the VMs after migration.

- [BZ# 990108](#)

Volume options that begin with *user.** are considered user options and these options cannot be reset as there is no way of knowing the default value.

- [BZ# 1065070](#)

The **python-urllib3** package fails to downgrade and this in turn results in Red Hat Storage downgrade process failure.

Workaround: Move the `/usr/lib/python2.6/site-packages/urllib3*` to `/tmp` and perform a fresh installation of the **python-urllib3** package.

- [BZ# 1086159](#)

The glusterd service crashes when volume management commands are executed concurrently with peer commands.

- [BZ# 1130270](#)

If a 32 bit Samba package is installed before installing Red Hat Storage Samba package, the installation fails as Samba packages built for Red Hat Storage do not have 32 bit variants

Workaround: Uninstall 32 bit variants of Samba packages.

- [BZ# 1139183](#)

The Red Hat Storage 3.0 version does not prevent clients with versions older Red Hat Storage 3.0 from mounting a volume on which rebalance is performed. Users with versions older than Red Hat Storage 3.0 mounting a volume on which rebalance is performed can lead to data loss.

You must install latest client version to avoid this issue.

- [BZ# 1127178](#)

If a replica brick goes down and comes up when `rm -rf` command is executed, the operation may fail with the message *Directory not empty*.

Workaround: Retry the operation when there are no pending self-heals.

- [BZ# 969020](#)

Renaming a file during remove-brick operation may cause the file not to get migrated from the removed brick.

Workaround: Check the removed brick for any files that might not have been migrated and copy those to the gluster volume before decommissioning the brick.

- BZ# [1007773](#)

When `remove-brick start` command is executed, even though the graph change is propagated to the NFS server, the directory inodes in memory are not refreshed to exclude the removed brick. Hence, new files that are created may end up on the removed-brick.

Workaround: If files are found on the removed-brick path after `remove-brick commit`, copy them via a gluster mount point before re-purposing the removed brick.

- BZ# [1116115](#)

When I/O operation is being performed at a high rate on the volume, faster than the rate at which the quota's accounting updates disk usage, the disk usage crosses the soft-limit mark without being failed with `EDQUOT`.

Workaround: Set `features.soft-timeout` and `features.hard-timeout` to a lower value, depending on the workload. Setting `features.soft-timeout` or `features.hard-timeout` to zero ensures that the disk usage accounting happens in line with the I/O operations performed on the volume, depending on the applicable timeout, but it could result in an increase in I/O latencies.



NOTE

The `features.soft-timeout` applies when the disk usage is lesser than the quota soft-limit set on a directory. `features.hard-timeout` applies when the disk usage is greater than the quota soft-limit but lesser than the hard-limit set on a directory.

- BZ# [1116121](#)

When I/O is being performed at a high rate on the volume, faster than the rate at which the quota's accounting updates disk usage, the disk usage crosses the soft-limit mark without an alert being logged.

Workaround: Set `features.soft-timeout` to a lower value, depending on the workload. Setting `features.soft-timeout` to zero would ensure that the disk usage accounting happens in line with the I/O performed on the volume, but it could result in an increase in I/O latencies.

- BZ# [1120437](#)

Executing `peer-status` command on probed host displays the IP address of the node on which the peer probe was performed.

Example When probing a peer, node B with hostname from node A, executing `peer status` command on node B, displays IP address of node A instead of its hostname.

Workaround: Probe node A from node B with hostname of node A. For example, execute the command: `# gluster peer probe HostnameA` from node B.

- BZ# [1097309](#)

If a NFS-client gets disconnected from the Red Hat Storage server without releasing the locks it obtained, these locks can prevent other NFS clients or Native (FUSE) clients to access the locked files. The NFSv3 protocol does not allow releasing the locks server side without a restart of the NFS servers. A restart triggers a grace period where existing locks need to get re-obtained. Locks are expired and made available to other NFS clients when the previous NFS clients do not re-request the previously obtained locks. More details related to NFS clients and lock recovery can be found in <https://access.redhat.com/site/solutions/895263>.

- BZ# [1122371](#)

The NFS server process and gluster self-heal daemon process restarts when gluster daemon process is restarted.

- BZ# [1110692](#)

Executing `remove-brick status` command, after stopping remove-brick process, fails and displays a message that the remove-brick process is not started.

- BZ# [1123733](#)

Executing a command which involves glusterd-glusterd communication `Examplegluster volume status` immediately after one of the nodes is down hangs and fails after 2 minutes with cli-timeout message. The subsequent command fails with the error message *Another transaction in progress* for 10 mins (frame timeout).

Workaround: Set a non-zero value for `ping-timeout` in `/etc/glusterfs/glusterd.vol` file.

- BZ# [1115915](#)

When a GlusterFS-native (FUSE) client loses its connection to the storage server without properly closing it, the brick process does not release resources (like locks) within an acceptable time. Other GlusterFS-native clients that require these resources can get blocked until the TCP-connection gets garbage collected by networking layers in the kernel and the brick processes get informed about it. This can introduce delays of 15-20 minutes before locks are released.

Workaround: Reduce the value of the system-wide `net.ipv4.tcp_retries2 sysctl`. Due to this change, the network layer of the Linux kernel times-out TCP-connections sooner.

- BZ# [1136718](#)

The afr self-heal can leave behind a partially healed file if the brick containing afr self-heal source file goes down in the middle of heal operation. If this partially healed file is migrated before the brick that was down comes online again, the migrated file would have incorrect data and the original file would be deleted.

- BZ# [1139193](#)

After `add-brick` operation, any application (like git) which attempts `opendir` on a previously present directory fails with `ESTALE/ENOENT` errors.

- BZ# [1141172](#)

If you rename a file from multiple mount points, there are chances of losing the file. This issue is witnessed since `mv` command sends unlinks instead of renames when source and destination happens to be hard links to each other. Hence, the issue is in `mv`, distributed as part of `coreutils` in various Linux distributions.

For example, if there are parallel renames of the form (mv a b) and (mv b a) where a and b are hard links to the same file, because of the above mentioned behavior of mv, unlink (a) and unlink (b) would be issued from both instances of mv. This results in losing both the links a and b and hence the file.

- BZ# [979926](#)

When any process establishes a TCP connection with `glusterfs` servers of a volume using port > 1023, the server rejects the requests and the corresponding file or management operations fail. By default, `glusterfs` servers treat ports > 1023 as unprivileged.

Workaround: To disable this behavior, enable `rpc-auth-allow-insecure` option on the volume using the steps given below:

1. To allow `insecure` connections to a volume, run the following command:

```
#gluster volume set VOLNAME rpc-auth-allow-insecure on
```

2. To allow `insecure` connections to `glusterd` process, add the following line in `/etc/glusterfs/glusterd.vol` file:

```
option rpc-auth-allow-insecure on
```

3. Restart `glusterd` process using the following command:

```
# service glusterd restart
```

4. Restrict connections to trusted clients using the following command:

```
#gluster volume set VOLNAME auth.allow IP address
```

- BZ# [1139676](#)

Renaming a directory may cause both source and target directories to exist on the volume with the same GFID and make some files in these directories not visible from the mount point. The files will still be present on the bricks.

Workaround: The steps to fix this issue are documented in:

<https://access.redhat.com/solutions/1211133>

- BZ# [1030309](#)

During directory creations attempted by geo-replication, though an `mkdir` fails with `EEXIST`, the directory might not have a complete layout for sometime and the directory creation fails with `Directory exists` message. This can happen if there is a parallel `mkdir` attempt on the same name. Till the other `mkdir` completes, layout is not set on the directory. Without a layout, entry creations within that directory fails.

Workaround: Set the layout on those subvolumes where the directory is already created by the parallel `mkdir` before failing the current `mkdir` with `EEXIST`.

**NOTE**

This is not a complete fix as the other `mkdir` might not have created directories on all subvolumes. The layout is set on the subvolumes where directory is already created. Any file or directory names which hash to these subvolumes on which layout is set, can be created successfully.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

3.2. RED HAT STORAGE CONSOLE

Issues related to Red Hat Storage Console

- [BZ#1164662](#)

The **Trends** tab in the Red Hat Storage Console appears to be empty after the ovirt engine restarts. This is due to the Red Hat Storage Console UI-plugin failing to load on the first instance of restarting the ovirt engine.

Workaround: Refresh(F5) the browser page to load the **Trends** tab.

- [BZ#1165269](#)

When you add a Red Hat Storage node in the Red Hat Storage Console using its IP address and remove it from the Red Hat Storage trusted Storage Pool, and consequently use the FQDN of the node to add it again to the Trusted Storage pool, the operation fails. This results in the node becoming unresponsive.

Workaround: Remove the unresponsive node from the Red Hat Storage Trusted Storage Pool and try adding it again.

- [BZ#1167305](#)

The **Trends** tab on the Red Hat Storage Console does not display the thin-pool utilization graphs in addition to the brick utilization graphs. Currently, there is no mechanism for the UI plugin to detect if the volume is provisioned using the thin provisioning feature.

- [BZ#1167572](#)

On editing the cluster version in the Edit Cluster dialog box on the Red Hat Storage Console, the compatible version field gets loaded with the highest available compatibility version by default, instead of the current version of the cluster.

Workaround: Select the correct version of the cluster in the **Edit Cluster** dialog box before clicking on the **OK** button.

- [BZ#916095](#)

If Red Hat Storage node is added to the cluster using IP address and the same Red Hat Storage node is later added using the FQDN (Fully Qualified Domain Name), the installation fails.

- [BZ# 990108](#)

Resetting the `user.cifs` option using the **Create Volume** operation on the **Volume Options** tab on the Red Hat Storage Console reports a failure.

- BZ# [978927](#)

Log messages that Red Hat Storage Console is trying to update VM/Template information are displayed.

- BZ# [880509](#)

When run on versions higher than Firefox 17, the Red Hat Storage Console login page displays a browser incompatibility warning. Red Hat Storage Console can be best viewed in Firefox 17 and higher versions.

- BZ# [1049759](#)

When `rhsc-log-collector` command is run, after collecting logs from different servers, the Terminal becomes garbled and unusable.

Workaround: Run the `reset` command.

- BZ# [1054366](#)

In Internet Explorer 10, while creating a new cluster with Compatibility version 3.3, the **Host** drop down list does not open correctly. Also, if there is only one item, the drop down list gets hidden when the user clicks on it.

- BZ# [1053395](#)

In Internet Explorer, while performing a task, an error message Unable to evaluate payload is displayed.

- BZ# [1056372](#)

When no migration is occurring, incorrect error message is displayed for the **stop migrate** operation.

- BZ# [1049890](#)

When gluster daemon service is restarted, failed Rebalance is started automatically and the status is displayed as *Started* in the Red Hat Storage Console.

- BZ# [1048426](#)

When there are more entries in Rebalance Status and remove-brick Status window, the column names scrolls up along with the entries while scrolling the window.

Workaround: Scroll up the Rebalance Status and remove-brick Status window to view the column names.

- BZ# [1053112](#)

When large sized files are migrated, the stop migrate task does not stop the migration immediately but only after the migration is complete.

- BZ# [1040310](#)

If the Rebalance Status dialog box is open in the Red Hat Storage Console while Rebalance is being stopped from the Command Line Interface, the status is currently updated as *Stopped*. But if the Rebalance Status dialog box is not open, the task status is displayed as *Unknown* because the status update relies on the gluster Command Line Interface.

- [BZ# 1051696](#)

When a cluster with compatibility version 3.2 contains Red Hat Storage 2.1 U2 nodes, creating Volume with bricks in root partition fails and the force option to allow bricks in root partition is not displayed.

Workaround: Do not create bricks in root partition or move the Cluster Compatibility Version to 3.3.

- [BZ# 838329](#)

When incorrect create request is sent through REST api, an error message is displayed which contains the internal package structure.

- [BZ# 1049863](#)

When Rebalance is running on multiple volumes, viewing the brick advanced details fails and the error message could not fetch brick details, please try again later is displayed in the **Brick Advanced Details** dialog box.

- [BZ# 1022955](#)

Rebalance or remove-brick cannot be started immediately after stopping Rebalance or remove-brick, when a large file migration is in progress, as part of the previous operation (rebalance or remove-brick), even though it says it has stopped.

- [BZ# 1015455](#)

The information on successfully completed Rebalance volume task is cleared from the Red Hat Storage Console after 5 minutes. The information on failed tasks is cleared after 1 hour.

- [BZ# 1038691](#)

The RESTful Service Description Language (RSDL) file displays only the response type and not the detailed view of the response elements.

Workaround: Refer the URL/API schema for detailed view of the elements of response type for the actions.

- [BZ# 1024184](#)

If there is an error while adding bricks, all the "." characters of FQDN / IP address in the error message will be replaced with "_" characters.

- [BZ# 982625](#)

Red Hat Storage Console allows adding Red Hat Storage 2.0+ and Red Hat Storage 2.1 servers into a 3.0 Cluster which is not supported in Red Hat Storage.

- [BZ# 975399](#)

When Gluster daemon service is restarted, the host status does not change to UP from Non-Operational immediately in the Red Hat Storage Console. There would be a 5 minute interval for auto-recovery operations which detect changes in Non-Operational hosts.

- [BZ# 971676](#)

While enabling or disabling Gluster hooks, the error message displayed if all the servers are not in UP state is incorrect.

- BZ# [1054759](#)

A `vdsm-tool` crash report is detected by Automatic Bug Reporting Tool (ABRT) in Red Hat Storage Node as the `/etc/vdsm/vdsm.id` file was not found during the first time.

Workaround: Execute the command `/usr/sbin/dmidecode -s system-uuid > /etc/vdsm/vdsm.id` before adding the node to avoid the `vdsm-tool` crash report.

- BZ# [1057122](#)

While configuring the Red Hat Storage Console to use a remote database server, on providing either `yes` or `no` as input for `Database host name validation` parameter, it is considered as `No`.

- BZ# [1042808](#)

When `remove-brick` operation fails on a volume, the Red Hat Storage node does not allow any other operation on that volume.

Workaround: Perform `commit` or `stop` for the failed `remove-brick` task, before another task can be started on the volume.

- BZ# [1060991](#)

In Red Hat Storage Console, Technology Preview warning is not displayed for `stop remove-brick` operation.

- BZ# [1057450](#)

Brick operations like adding and removing a brick from Red Hat Storage Console fails when Red Hat Storage nodes in the cluster have multiple FQDNs (Fully Qualified Domain Names).

Workaround: Host with multiple interfaces should map to the same **FQDN** for both Red Hat Storage Console and `gluster peer probe`.

- BZ# [958803](#)

When a brick process goes down, the brick status is not updated and displayed immediately in the Red Hat Storage Console as the Red Hat Storage Console synchronizes with the `gluster Command Line Interface` every 5 minutes for brick status.

- BZ# [1038663](#)

Framework restricts displaying delete actions for collections in RSDL display.

- BZ# [1061677](#)

When Red Hat Storage Console detects a `remove-brick` operation which is started from `gluster Command Line Interface`, engine does not acquire lock on the volume and `Rebalance` task is allowed.

Workaround: Perform `commit` or `stop` on `remove-brick` operation before starting `Rebalance`.

- BZ# [1061813](#)

After stopping, committing, or retaining bricks from the Red Hat Storage Console UI, the details of files scanned, moved, and failed are not displayed in the Tasks pane.

Workaround: Use *Status* option in *Activities* column to view the details of the remove-brick operation.

- BZ# [924826](#)

In Red Hat Storage Console, parameters related to Red Hat Enterprise Virtualization are displayed while searching for Hosts using the Search bar.

- BZ# [1062612](#)

When Red Hat Storage 2.1 Update 2 nodes are added to 3.2 cluster, users are allowed to perform Rebalance and remove-brick tasks which are not supported for 3.2 cluster.

- BZ# [977355](#)

When resolving a missing hook conflict, if one of the servers in the cluster is not online, an error message is displayed without the server name. Hence, the server which was down can not be identified.

Workaround: Identify the information on the server which was down from the Hosts tab.

- BZ# [1046055](#)

While creating volume, if the bricks are added in root partition, the error message displayed does not contain the information that **Allow bricks in root partition and re-use the bricks by clearing xattrs** option needs to be selected to add bricks in root partition.

Workaround: Select **Allow bricks in root partition and re-use the bricks by clearing xattrs** option to add bricks in root partition.

- BZ# [1060991](#)

In Red Hat Storage Console UI, Technology Preview warning is not displayed for stop remove-brick operation.

- BZ# [1066130](#)

Simultaneous start of Rebalance on volumes that span same set of hosts fails as gluster daemon lock is acquired on participating hosts.

Workaround: Start Rebalance again on the other volume after the process starts on first volume.

- BZ# [1086718](#)

The Red Hat Access Plug-in related answers are not written to the answer file correctly when the `rhsc-setup --generate-answer=answer-file` command is executed and hence the next offline execution of `rhsc-setup (rhsc-setup --offline --config-append=answer-file)` asks Red Hat Access Plug-in related questions again.

Workaround:

1. Add the entry given below in the answer-file
`OVESETUP_SYSTEM/configureRedhatAccessPlugin=bool:False`
2. Execute the Red Hat Storage Console setup script in the offline mode `rhsc-setup --offline --config-append=answer-file`

- BZ# [1086723](#)

Auto installation of Red Hat Storage Console using the answer file in Red Hat Storage Console-2.1.2 fails with the following warning message:

[WARNING] Current installation does not support upgrade. It is recommended to clean your installation using rhsc-cleanup and re-install.

Workaround:

1. Execute the command: `yum update rhsc-setup`
2. Perform the Red Hat Storage Console auto installation using the `answer-file` with the command: `rhsc-setup --config-append=answer-file`

- BZ# [1201740](#)

The Red Hat Storage Console overrides the Red Hat Enterprise Linux values for the `vm.dirty_ratio` and `dirty_background_ratio` to 5 and 2 respectively. This occurs when activating the tuned profile `rhs-virtualization` while adding Red Hat Storage nodes to the Red Hat Storage Console. This could decrease the performance of the Red Hat Storage Trusted Storage Pool.

Workaround: Restore the values of `vm.dirty_ratio` and `dirty_background_ratio` to the default values by activating the default tuned profile by executing the following command:

```
# tuned-adm profile default
```

- BZ# [1202731](#)

When you upgrade the Red Hat Storage Console from version 2.1.6 to 3.0.4, the `rhsc-log-collector` package does not get upgraded automatically.

Workaround: Execute the following command to manually upgrade the package:

```
# yum upgrade rhsc-log-collector
```

- BZ# [1200248](#)

The **Trends** tab on the Red Hat Storage Console does not display all the network interfaces available on a host. This limitation is because the Red Hat Storage Console `ui-plugin` does not have this information.

Workaround: The graphs associated with the hosts are available in the Nagios UI on the Red Hat Storage Console. You can view the graphs by clicking the **Nagios home** link

- BZ# [1194150](#)

The volumes of RDMA type are not supported by the Red Hat Storage Console. Such volumes

can be created using the gluster CLI and the volumes are displayed on the UI. If the RDMA volume does not support TCP then TCP port is displayed as N/A and the corresponding brick appears offline. The bricks that belong to pure RDMA volumes which do not support TCP, appear offline on the console but work properly in the Red Hat Storage server.

- [BZ# 1196144](#)

A few changes in the nrpe.cfg file requires restarting the nrpe service when the gluster-nagios-addons package is updated. As a result, the brick and disk monitoring does not work properly unless the nrpe service is restarted.

Workaround: Restart the nrpe service using the following command:

```
# service nrpe reload
```

Issues related to the Red Hat Storage Console Command Line Interface:

- [BZ# 928926](#)

When you create a cluster through API, enabling both *gluster_service* and *virt_service* is allowed though this is not supported.

- [BZ# 1059806](#)

In Red Hat Storage Console Command Line Interface, removing a cluster using its name fails with an error message and the same cluster gets deleted if UUID is used in remove command.

- [BZ# 1108688](#)

An image in the Nagios home page is not transferred via SSL and the Security details displays the following message:

Connection Partially Encrypted.

- [BZ# 1111549](#)

In the Red Hat Storage Console, if the name provided in the **Name** field of the host in the **New Host**" pop-up is different from the **Hostname** provided in Nagios, the utilization details for the hosts are not displayed in Trends tab.

- [BZ# 1100960](#)

Nagios does not support SELinux fully and this impacts rhsc-setup and normal usage of Red Hat Console with Nagios.

Workaround: Run SELinux in permissive mode.

- [BZ# 1113103](#)

The Network Utilization graph shows an error that RRD data does not exist.

Workaround: After midnight of the next day, the sadf output file gets corrected automatically and graph works fine.

- [BZ# 1121612](#)

The mount points for a host can not be detected from Trends tab in Red Hat Storage Console GUI.

View mount point Disk Utilization using Nagios GUI.

- [BZ# 1101181](#)

After stopping the rebalance/remove-brick operation from Red Hat Storage Console GUI, clicking on rebalance/remove-brick status button throws *Unable to fetch data* error.

- [BZ# 1125960](#)

Both Red Hat Storage Console and Hadoop Ambari are using Nagios to monitor Red Hat Storage Server nodes, but there is version conflict for the package nagios-plugins, which blocks the installation of nagios required by Hadoop Ambari. Red Hat Storage server 3.0 nodes are pre-bundled with latest version of nagios-plugin. If you want to use Red Hat Storage Console, then you will not be able to install Nagios via Ambari and utilize the pre-build hadoop monitors/alerts. HDP 2.0.6 is not supported for Red Hat Storage 3.0 release.

[23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33Report a bug](#)

3.3. RED HAT STORAGE AND RED HAT ENTERPRISE VIRTUALIZATION INTEGRATION

Issues related to Red Hat Enterprise Virtualization and Red Hat Storage Integration

- In the case that the Red Hat Storage server nodes and the Red Hat Enterprise Virtualization Hypervisors are present in the same data center, the servers of both types are listed for selection when you create a virtual machine or add a storage domain. Red Hat recommends that you create a separate data center for the Red Hat Storage server nodes.

- [BZ# 867236](#)

While deleting a virtual machine using the Red Hat Enterprise Virtualization Manager, the virtual machine is deleted but remains in the actual storage. This consumes unnecessary storage.

Workaround: Delete the virtual machine manually using the command line interface. Deleting the virtual image file frees the space.

- [BZ# 918032](#)

In this release, the `direct-io-mode=enable` mount option does not work on the Hypervisor.

- [BZ# 979901](#)

Virtual machines may experience very slow performance when a rebalance operation is initiated on the storage volume. This scenario is observed when the load on storage servers are extremely high. Hence, it is recommended to run the rebalance operation when the load is low.

- [BZ# 856121](#)

When a volume starts, a `.glusterfs` directory is created in the back-end export directory. When a `remove-brick` command is performed, it only changes the volume configuration to remove the brick and stale data is present in back-end export directory.

Workaround: Run this command on the Red Hat Storage Server node to delete the stale data.

```
# rm -rf /export-dir
```

- BZ# [866908](#)

The `gluster volume heal VOLNAME info` command gives stale entries in its output in a few scenarios.

Workaround: Execute the command after 10 minutes. This removes the entries from internal data structures and the command does not display stale entries.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

3.4. RED HAT STORAGE AND RED HAT OPENSTACK INTEGRATION

Issues related to Red Hat OpenStack and Red Hat Storage integration

- BZ# [1004745](#)

If a replica pair is down while taking a snapshot of a Nova instance on top of a Cinder volume hosted on a Red Hat Storage volume, the snapshot process may not complete as expected.

- BZ# [991490](#)

Mount options specified in `glusterfs_shares_config` file are not honored when it is specified as part of `enabled_backends` group.

- BZ# [980977](#) and BZ# [1017340](#)

If storage becomes unavailable, the volume actions fail with `error_deleting` message.

Workaround: Run `gluster volume delete VOLNAME force` to forcefully delete the volume.

- BZ# [1042801](#)

Cinder volume migration fails to migrate from one glusterFS back-end cluster to another. The migration fails even though the target volume is created.

- BZ# [1062848](#)

When a nova instance is rebooted while rebalance is in progress on the Red Hat Storage volume, the root file system will be mounted as read-only after the instance comes back up. Corruption messages are also seen on the instance.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

CHAPTER 4. TECHNOLOGY PREVIEWS

This chapter provides a list of all available Technology Preview features in Red Hat Storage 3.0 release.

Technology Preview features are currently not supported under Red Hat Storage subscription services, may not be functionally complete, and are generally not suitable for production environments. However, these features are included for customer convenience and to provide wider exposure to the feature.

Customers may find these features useful in a non-production environment. Customers are also free to provide feedback and functionality suggestions for a Technology Preview feature before it becomes fully supported. Errata will be provided for high-severity security issues.

During the development of a Technology Preview feature, additional components may become available to the public for testing. Red Hat intends to fully support Technology Preview features in the future releases.



NOTE

All Technology Preview features in Red Hat Enterprise Linux 6.5 release, will be considered as technology preview features in Red Hat Storage 3.0 release. For more information on technology preview features of Red Hat Enterprise Linux 6.5 release, see *Technology Previews* chapter of *Red Hat Enterprise Linux 6.5 Technical Notes*

[23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33Report a bug](#)

4.1. GSTATUS COMMAND

The `gstatus` command provides an easy-to-use, high-level view of the health of a trusted storage pool with a single command. It gathers information by executing the GlusterFS commands, to gather information about the statuses of the Red Hat Storage nodes, volumes, and bricks.

For more information, see section *gstatus command* in the *Red Hat Storage Administration Guide*

[23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33Report a bug](#)

4.2. STRIPED VOLUMES

Data is striped across bricks in a Striped volume. It is recommended that you use striped volumes only in high concurrency environments where accessing very large files is critical.

For more information, refer to section *Creating Striped Volumes* in the *Red Hat Storage 3 Administration Guide*.

[23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33Report a bug](#)

4.3. DISTRIBUTED-STRIPED VOLUMES

The distributed striped volumes stripe data across two or more nodes in the trusted storage pool. It is recommended that you use distributed striped volumes to scale storage and to access very large files during critical operations in high concurrency environments.

For more information, refer to section *Creating Distributed Striped Volumes* in the *Red Hat Storage 3 Administration Guide*.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.4. DISTRIBUTED-STRIPED-REPLICATED VOLUMES

Distributed striped replicated volumes distribute striped data across replicated bricks in a trusted storage pool. It is recommended that you use distributed striped replicated volumes in highly concurrent environments where there is parallel access of very large files and where performance is critical. Configuration of this volume type is supported only for Map Reduce workloads.

For more information, refer the section *Creating Distributed Striped Replicated Volumes* in the *Red Hat Storage 3 Administration Guide*.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.5. STRIPED-REPLICATED VOLUMES

The striped replicated volumes stripe data across replicated bricks in a trusted storage pool. It is recommended that you use striped replicated volumes in highly concurrent environments where there is parallel access of very large files and where performance is critical. In this release, configuration of this volume type is supported only for Map Reduce workloads.

For more information, refer the section *Creating Striped Replicated Volumes* in the *Red Hat Storage 3 Administration Guide*.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.6. REPLICATED VOLUMES WITH REPLICA COUNT GREATER THAN 3

The replicated volumes create copies of files across multiple bricks in the volume. It is recommended that you use replicated volumes in environments where high-availability and high reliability are critical. Creating replicated volumes with replica count more than 3 is under technology preview.

For more information, refer the section *Creating Replicated Volumes* in the *Red Hat Storage 3 Administration Guide*.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.7. STOP REMOVE BRICK OPERATION

You can stop a remove brick operation after you have opted to remove a brick through the Command Line Interface and Red Hat Storage Console. After executing a remove-brick operation, you can choose to stop the remove-brick operation by executing the `remove-brick stop` command. The files that are already migrated during remove-brick operation, will not be reverse migrated to the original brick.

For more information, refer the section *Stopping Remove Brick Operation* in the *Red Hat Storage 3 Administration Guide* and section *Stopping a Remove Brick Operation* in the *Red Hat Storage 3 Console Administration Guide*.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.8. READ-ONLY VOLUME

Red Hat Storage enables you to mount volumes with read-only permission. While mounting the client, you can mount a volume as read-only and also make the entire volume as read-only, which applies for all the clients using the `volume set` command.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.9. NFS GANESHA

With the 2.0 release in the open source community, *nfs-ganesha* supports Red Hat Storage volumes. *nfs-ganesha* 2.0 includes improved protocol support and stability. With Red Hat Storage 3 release, the *nfs-ganesha* feature in technology preview is being refreshed to bring in a number of significant enhancements and capabilities including dynamic configuration (addition and deletion of exports at run time), NFSv4 ACL's for glusterFS, glusterFS multi volume support, NFSv4 PseudoFS support and NFSv4/v3 kerberos authentication. With this feature, Red Hat Storage volumes can be exported using *nfs-ganesha* server for consumption by both NFSv3 and NFSv4 clients.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

4.10. NON UNIFORM FILE ALLOCATION

When a client on a server creates files, the files are allocated to a brick in the volume based on the file name. This allocation may not be ideal, as there is higher latency and unnecessary network traffic for read/write operations to a non-local brick or export directory. NUFA ensures that the files are created in the local export directory of the server, and as a result, reduces latency and conserves bandwidth for that server accessing that file. This can also be useful for applications running on mount points on the storage server.

If the local brick runs out of space or reaches the minimum disk free limit, instead of allocating files to the local brick, NUFA distributes files to other bricks in the same volume if there is space available on those bricks. You must enable NUFA before creating any data in the volume.

23050%2C+3.0+Update+4+Release+Notes-308-06-2015+18%3A49%3A33[Report a bug](#)

APPENDIX A. REVISION HISTORY

Revision 3-10 Deleted the bug BZ# 1136714 in the Known Issues chapter.	Mon Jun 08 2015	Divya Muntimadugu
Revision 3-9 Added the bug BZ# 1203901 in the Known Issues chapter.	Wed Apr 08 2015	Pavithra Srinivasan
Revision 3-8 Added three-way replication support for Hadoop deployments in the What's New in this Release? chapter.	Mon Mar 30 2015	Divya Muntimadugu
Revision 3-7 Version for 3.0.4 release.	Wed Mar 25 2015	Pavithra Srinivasan