

Release Notes for Patches for the MapR 5.0.0 Release

Release Notes for the December 2016 Patch

Released 12/09/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-40889.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-40889.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-40889.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-40889.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-40889.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-40889.x86_64.deb
Win32	Client	mapr-client-5.0.0.40889GA-1.win32.zip
Win64	Client	mapr-client-5.0.0.40889GA-1.amd64.zip
Mac	Client	mapr-client-5.0.0.40889GA-1.x86_64.tar.gz

Fixes

Bug 13187

Description

The `maprcli volume create` command was not setting group ownership to user's primary group.

Resolution

With this fix, the `maprcli volume create` command will set group ownership to user's primary group.

Bug 20965

Description

When working with multiple clusters, synchronization issues was causing MapRFileSystem to return NullPointerException.

Resolution

With this fix, MapRFileSystem has been improved to better support working with multiple clusters and MapRFileSystem contains fixes for synchronization issues.

Bug 24139

Description

If limit spread was enabled and the nodes were more than 85% full, CLDB did not allocate containers for IOs on non-local volumes.

Resolution

With this fix, CLDB will now allocate new containers to ensure that the IO does not fail.

Bug 24249

Description

When running map/reduce jobs with older versions of the MapR classes, the system hung because the older classes linked to the native library installed on cluster nodes that have been updated to a newer MapR version.

Resolution

With this fix, the new `fs.mapr.bailout.on.library.mismatch` parameter detects mismatched libraries, fails the map/reduce job, and logs an error message. The parameter is enabled by default. You can disable the parameter on all the TaskTracker nodes and resubmit the job for the task to continue to run. To disable the parameter, you must set it to `false` in the `core-site.xml` file.

Bug 24618

Description

Remote mirror volumes could not be created on secure clusters using MCS even when the appropriate tickets were present.

Resolution

With this fix, remote mirror volumes can now be created on secure clusters using MCS.

Bug 24965

Description

On large clusters, sometimes the bind failed with the message indicating unavailability of port when running MR jobs, specifically reducer tasks.

Resolution

With this fix, the new `fs.mapr.bind.retries` configuration parameter in `core-site.xml` file, if set to `true`, will retry to bind during client initialization for 5 minutes before failing. By default, the `fs.mapr.bind.retries` configuration parameter is set to `false`.

Bug 24969

Description

The `maprcli volume create` command was not setting group ownership to user's primary group when the user's primary GID was not the first GID in the list of GIDs.

Resolution

With this fix, the primary GID of the user performing the operation will now be the first GID in the list of GIDs.

Bug 24971

Description

When the mirroring operation started after a CLDB failover, sometimes it was sending requests to slave CLDB where data was stale, resulting in the the mirroring operation hanging. If the CLDB failover happened again during this time, the new CLDB master was discarding data resynchronized by the old mirroring operation, but marking the mirroring operation as successful. This resulted in data mismatch between source and destination.

Resolution

With this fix, mirroring requests will be sent to master CLDB node only.

Bug 25003

Description

When a specific queue uses all of its resources, the UsedResources tab in the Resource Manager UI might show a greater value than shown in the MaxResources tab. This happens when another application is submitted and the application master container size is added in.

Resolution

With this fix, no more containers can be assigned to a queue when its UsedResource has reached the MaxResource limit.

Bug 25041

Description

Whenever a newly added node was made the master of the name container, MFS crashed while deleting files in the background.

Resolution

With this fix, MFS will not crash when a newly added node is made the master of the name container.

Release Notes for the October 2016 Patch

Released 10/24/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-40097.x86_64.rpm
---------	--------	--

Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-40097.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-40097.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-40097.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-40097.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-40097.x86_64.deb
Win32	Client	mapr-client-5.0.0.40097GA-1.win32.zip
Win64	Client	mapr-client-5.0.0.40097GA-1.amd64.zip
Mac	Client	mapr-client-5.0.0.40097GA-1.x86_64.tar.gz

Fixes

Bug 14105

Description

When nodes attempt to register with duplicate IDs, CLDB does not register the nodes and log meaningful error messages.

Resolution

With this fix, when nodes attempt to register with duplicate IDs, CLDB will log appropriate error messages.

Bug 24505

Description

A job failed when the JvmManager went into an inconsistent state.

Resolution

With this fix, jobs no longer fail as a result of the JvmManager entering an inconsistent state.

Bug 24562

Description

CLDB (container location database) performance suffered because Warden gave the CLDB service a lower CPU priority.

Resolution

With this fix, Warden uses a new algorithm to set the correct CPU priority for the CLDB service.

Bug 24651

Description

CLDB threw an exception and failed over when the snapshots list was iterated over while snapshots were being created.

Resolution

With this fix, CLDB will no longer fail over when snapshots list is iterated over while new snapshots are being created.

Bug 24656

Description

MFS was churning cpu while taking snapshot because of some debug code in the builds.

Resolution

With this fix, MFS will no longer churn CPU as the debug code has been disabled.

Bug 24700

Description

The Job Tracker user interface failed with a NullPointerException when a user submitted a Hive job with a null value in a method.

Resolution

With this fix, the Job Tracker interface does not fail when a Hive job is run with a null value in a method.

Release Notes for the September 2016 Patch

Released 9/23/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-39744.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-39744.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-39744.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-39744.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-39744.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-39744.x86_64.deb
Win32	Client	mapr-client-5.0.0.39744GA-1.win32.zip
Win64	Client	mapr-client-5.0.0.39744GA-1.amd64.zip
Mac	Client	mapr-client-5.0.0.39744GA-1.x86_64.tar.gz

Fixes

Bug 23652

Description

The POSIX loopbacknfs client did not automatically refresh renewed service tickets.

Resolution

With this fix, the POSIX loopbacknfs client will:

- Automatically use the renewed service ticket without requiring a restart if the ticket is replaced before expiration (ticket expiry time + grace period of 55 minutes). If the ticket is replaced after expiration (which is ticket expiry time + grace period of 55 minutes), the POSIX loopbacknfs client will not refresh the ticket as the mount will become stale.
- Allow impersonation if a service ticket is replaced before ticket expiration (which is ticket expiry time + grace period of 55 minutes) with a servicewithimpersonation ticket.
- Honor all changes to user/group IDs of the renewed ticket.

Bug 24053

Description

During client initialization, the client crashed if there was an error during initialization.

Resolution

With this fix, the client will not crash if there is an error during initialization.

Bug 24119

Details

Warden adjusts the FileServer (MFS) and Node Manager (NM) memory incorrectly when NM and TaskTracker (TT) are on the same node. This can result in too much memory being allocated to MFS.

Resolution

With this fix, Warden does not adjust MFS memory when NM and TT are on the same node. Memory adjustment is implemented only when TT and MapR-FS (but no NM) are on the same node.

Release Notes for the August 2016 Patch

Released 8/27/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-39340.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-39340.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-39340.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-39340.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-39340.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-39340.x86_64.deb
Win32	Client	mapr-client-5.0.0.39340GA-1.win32.zip
Win64	Client	mapr-client-5.0.0.39340GA-1.amd64.zip
Mac	Client	mapr-client-5.0.0.39340GA-1.x86_64.tar.gz

Fixes

Bug 23944

Details

In some cases, when a local write times out with ETIMEDOUT error, the NFS server re-uses shared pages, before mfs releases those pages, resulting in mfs crash.

Resolution

With this fix, on ETIMEDOUT error for local writes in NFS server, NFS server will not reuse those pages.

Bug 24022

Details

Mirroring of a volume on a container which does not have a master container caused the mirror thread to hang.

Resolution

With this fix, mirroring will not hang when the container associated with the volume has no master.

Bug 24063

Details

During mirroring, the volume property update messages were logged (in cldb.log) every 5 seconds because the log level was set to INFO.

Resolution

With this fix, the log level is now DEBUG and the log will not contain multiple volume property update messages from mirroring.

Bug 24140

Details

While trying to access secure cluster from Windows, if the MAPR_TICKETFILE_LOCATION pointed to an incorrect location, the MapRClient threw an assert and the application crashed.

Resolution

With this fix, the application will not crash if the MAPR_TICKETFILE_LOCATION points to an incorrect location. Instead, the application will exit gracefully.

Release Notes for the July 2016 Patch

Released 7/29/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	maprpatch5.0.0.32987.GA39045.x86_64.rpm
Red Hat	Client	maprpatchclient5.0.0.32987.GA39045.x86_64.rpm
Red Hat	Loopbacknfs	maprpatchloopbacknfs5.0.0.32987.GA39045.x86_64.rpm
Ubuntu	Server	maprpatchloopbacknfs5.0.0.32987.GA39045.x86_64.rpm
Ubuntu	Client	maprpatchclient5.0.0.32987.GA39045.x86_64.deb
Ubuntu	Loopbacknfs	maprpatchloopbacknfs5.0.0.32987.GA39045.x86_64.deb
Win32	Client	maprclient5.0.0.39045GA1.win32.zip
Win64	Client	maprclient5.0.0.39045GA1.amd64.zip
Mac	Client	maprclient5.0.0.39045GA1.x86_64.tar.gz

Bug 22368

Details

Sometimes the `mrconfig info dumpcontainers` command was looping its output indefinitely.

Resolution

With this fix, the command will not loop its output indefinitely.

Bug 23488

Details

When a volume with a replication factor below 3 and no minimum number of replication was created, the replication manager set the minimum replication value to 1 and containers with only one copy were not re-replicated.

Resolution

With this fix, by default, the minimum number of copies is 1 if the replication factor is less than or equal to 2.

Bug 23544

Details

Incorrect values are returned for Name Container Replication Factor. Namespace and Dataspace replication can be independently specified. However, if Namespace replica factors are not set explicitly, they assume the same values as that of data space for both 'desired' and 'min'.

If the 'min' and 'desired' repl factors of Dataspace are 2, and 3 respectively, when the replication manager checks for the desired replication of the namespace container, it would see '2' instead of '3'. Hence, it might fail to create an additional copy of the namespace containers (since it thinks the desired is 2, when in fact it should have been 3).

Volume info might show the desired and min correctly, but the available copies might fall short of the desired replica factor.

Resolution

This issue is fixed by invoking the correct function to return the namespace replica factor when is is not different from the data replica factor.

Bug 23629

Details

While allocating large number of inodes during resynchronization of containers, the source container would timeout if destination container did not respond within 5 minutes.

Resolution

With this fix, instead of sending large number of inodes during resynchronization, multiple commands with a fixed number of inodes per command will be sent to allocate the required number of inodes.

Bug 23715

Details

The MFS C and Java APIs did not return the requested number of bytes.

Resolution

With this fix, both C and Java APIs will return the requested number of bytes if present.

Bug 23795

Details

Some storage pools were going offline frequently with CRC errors because:

- NFS server was encrypting files even when network-encryption bit was not set on files. This happens when `dfs.feature.filecipherbit.support` feature is not enabled.
Note: This happens only on clusters that were upgraded from 3.0, or prior to 3.1 or later and configured as secure.
- During replication, if replica reported a CRC error for replicated writes, the source was retransmitting without verifying CRC at source. This caused the replica to go stale unnecessarily.

Resolution

With this fix:

- NFS server will no longer encrypt data unless "network-encryption" bit is set on the files.

- The source for the replica will check CRC and if it doesn't match, the source will crash. This will allow one of the replicas to become master.

Bug 23799

Details

When there is an error, the container resync work area was freed, but the inode resync work area was still referring to the container resync work area.

Resolution

With this fix, the container work area will wait till the completion of all node resync operations before releasing the work area.

Bug 23876

Details

Sometimes, the same node was getting added to the replica chain twice and this was blocking the next resync request.

Resolution

With this fix, the same the node will not get added twice as a check has been included to verify if a node is already in the chain before adding a node.

Release Notes for the June 2016 Patch

Released 6/24/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-38652.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-38652.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-38652.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-38652.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-38652.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-38652.x86_64.deb

Windows 32-bit	Client	mapr-client-5.0.0.38652GA-1.win32.zip
Windows 64-bit	Client	mapr-client-5.0.0.38652GA-1.amd64.zip
Mac OS X	Client	mapr-client-5.0.0.38652GA-1.x86_64.tar.gz

Fixes

Bug 20186

Details

The MapR dbclient logged ESTALE (Stale file handle) errors for successful operations on binary tables when MapR-DB client applications followed these steps:

1. User X initiated the truncation of a binary table.
2. User X issued an operation (get/put/scan, etc) against the table while MapR-DB was in the process of deleting the table, recreating it, and then refreshing inode information.

The attempt to modify the table was unsuccessful and the dbclient logged an ESTALE error. However, the dbclient retried the operation after the truncation process was over and successfully modified the table.

Resolution

With this fix, the dbclient no longer logs the ESTALE error before retrying an action on a table after the truncation process is finished.

Bug 23382

Details

CLDB fails over with an exception when a node with stale containers is removed.

Resolution

With this fix, a node with stale containers can be removed successfully from the cluster and CLDB exceptions are not thrown.

Bug 23473

Details

In this type of situation in MapR-DB, the first of a series of puts for a row would succeed, while the remaining puts in the series would fail without errors:

1. A tablet T is split into T1 and T. The dbclient still has tablet T cached with the original key range.
2. The dbclient issues a series of puts against a rowkey that used to be in T, but which is now in T1.
3. The server returns an ERANGE error for the first put, but not for the remaining puts in the series.
4. The dbclient retries the first put and succeeds, but does not retry the remaining puts because the dbclient never received the ERANGE error for those puts.

This problem could occur for different types of errors that applied to all of the puts issued together for a single row.

Resolution

With this fix, the server returns the relevant error message for all of the puts in a series for a single row.

Bug 23541

Details

A `ddlopen` of `libmapr_pam.so` using immediate symbol resolution throws an undefined symbol error.

Resolution

An updated `libmapr_pam.so` has been provided that links to `libpam.so`.

Bug 23544

Details

Incorrect values are returned for Name Container Replication Factor. Namespace and Dataspace replication can be independently specified. However, if Namespace replica factors are not set explicitly, they assume the same values as that of data space for both 'desired' and 'min'.

If the 'min' and 'desired' repl factors of Dataspace are 2, and 3 respectively, when the replication

manager checks for the desired replication of the namespace container, it would see '2' instead of '3'. Hence, it might fail to create an additional copy of the namespace containers (since it thinks the desired is 2, when in fact it should have been 3).

Volume info might show the desired and min correctly, but the available copies might fall short of the desired replica factor.

Resolution

This issue is fixed by invoking the correct function to return the namespace replica factor when it is not different from the data replica factor.

Release Notes for May 2016 Patch

Released 5/20/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-38215.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-38215.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-38215.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-38215.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-38215.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-38215.x86_64.deb
Windows 32-bit	Client	mapr-client-5.0.0.38215GA-1.win32.zip
Windows 64-bit	Client	mapr-client-5.0.0.38215GA-1.amd64.zip
Mac OS X	Client	mapr-client-5.0.0.38215GA-1.x86_64.tar.gz

Fixes

Bug 21613

Details

Resynchronization hung after an RPC failure because replicas did not handle the RPC failure scenario.

Resolution

After the fix, resynchronization fails gracefully after an RPC failure.

Bug 22491

Details

A typographical error in log messages prevented MapR support from accurately confirming that container resync errors were causing mirroring failures.

Resolution

With this fix, the typographical errors are corrected.

Bug 23311

Details

If a table was a source in table replication, an incremental bulk load of one or more non-replicated column families would cause the mfs service to core.

For example, suppose a source table contained the column families cf1, cf2, and cf3. Only cf3 was being replicated. If an incremental bulk load was started for cf1 and cf2, the mfs service cored.

Resolution

With this fix, the mfs service no longer cores in this type of situation.

Release Notes for the April 2016 Patch

Released 4/22/2016

These release notes describe the fixes that are included in this patch.

Packages

Red Hat	Server	mapr-patch-5.0.0.32987.GA-38013.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.0.0.32987.GA-38013.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-38013.x86_64.rpm
Ubuntu	Server	mapr-patch-5.0.0.32987.GA-37930.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.0.0.32987.GA-37930.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.0.0.32987.GA-37930.x86_64.deb
Windows 64-bit	Client	mapr-client-5.0.0.37930GA-1.amd64.zip
Mac OS X	Client	mapr-client-5.0.0.37930GA-1.x86_64.tar.gz

Fixes

Bug 21163

Details

On a MapR fileserver, resync errors could occur when data on the primary and replica containers were resynchronized after a network or node failure to ensure that all of the containers had the same data.

Resolution

With this fix, MapR fileserver now report resync errors to the CLDB.

Bug 22195

Details

If a replica clone container of a mirror volume tried to resync with an upstream container that had a clone container whose mirror resync had finished, the upstream container returned a RETRY error. The resync of the replica would be triggered only after the version roll forward happened on the upstream container.

Resolution

With this fix, the upstream container will now verify version numbers on both primary and clone containers, mirror IDs on both, and then allow the replication chain to establish.

Bug 22303

Details

If a replica container was ahead of the primary container during container resync operation, a cleanup of the replica is performed. If the source node failed during this cleanup, resync messages failed. Also, if resync was resumed before the cleanup was completed, resync resumed incorrectly.

Resolution

With this fix, when the source node is restarted after a failure, resync will be restarted if resync had been cleaning up a replica container at the time of failure.

Bug 22316

Details

After a node restart, the replica performs a fast resync to match versions with the primary. During a fast resync (that is, when versions are matching on primary and replica containers), there should not be any resync data. But if a new snapshot is created on the primary container and if the replica attempts to resync with primary, because the replica does not expect any resync data from the primary, the replica hits an assert.

Resolution

With this fix, if there is a need for a snapshot resync when a replica is performing a fast resync, the primary will return EAGAIN so that the replica retries with a slow resync, which allows for data transfer when there is data mismatch between the containers.

Bug 22318

Details

When volume mirroring was in progress, intra-volume resynchronization of data was not waiting until the mirroring process was complete.

When mirroring volumes, temporary snapshots are created to resynchronize the data. If replicas and primary replicas have snapshots, then resynchronization is allowed. If the primary replica has a snapshot but the replica does not, intra-volume resynchronization should wait until the mirroring process is complete.

Resolution

With this fix, if the primary replica has a snapshot but the replica does not, intra-volume resynchronization waits until the mirroring process is complete, meaning until each replica has a temporary snapshot.

Bug 22389

Details

When volume mirroring was in progress and some nodes in the cluster (where the mirror volume resides) failed and were restarted, some containers of the mirror volume would have old data and some would have new data.

Because of multiple node failures, a temporary snapshot of the container is deleted even though mirroring completes. The snapshot is used to resynchronize the data. As a result, containers, associated with deleted snapshots, have old data. In addition, the replica sends snapshot versions, which, if different, triggers the primary replica to try to resynchronize the data. Because snapshots are deleted, due to node failure and restart, the snapshot versions sent by the replica are different.

Resolution

With this fix, the replica sends the correct snapshot version, the primary replica verifies that both the replica and primary replica are at the same state, and MFS establishes the replica chain.

Bug 22534

Details

In situations where a client application looped between creating and deleting the same MapR-DB table, either of the following two circumstances could lead to a fileserver deadlock, preventing any other MapR filesystem operations in the volume hosting the table:

- The creation of a snapshot of the volume was triggered.
- A node hosting one of the containers of the table data failed.

Resolution

With this fix, fileserver deadlocks are no longer possible in these situations.

Bug 22808

Details

The calculation of the preemption utilization threshold of the Fair Scheduler's Dominant Resource Fairness (drf) scheduling policy did not consider disk usage as a resource. Instead, the preemption utilization threshold was calculated based on memory and CPU alone.

Resolution

With this fix, the drf scheduling policy considers memory, CPU, and disk usage when allocating resources to applications. For example, because MapReduce jobs require disk resources, preemption will now occur when the disk resources are at capacity.

Bug 22860

Details

Client applications holding two or more connections to the server could experience RPC timeouts in the following type of situation: After one connection establishes a session key with the server, all of the connections remain idle long enough to trigger a session key renewal on the server. Two or more requests are then sent in parallel on different connections. The first request processed on the server triggers a change of the previous session key to the new session key. The remaining requests subsequently reaching the server on the other connections have the old session key, rather than the new session key.

Resolution

With this fix, the requests with the old session key are now discarded by the server and the client retransmits the requests with the new session key after a timeout that generally lasts from one to two minutes.

Bug 22881

Details

When mirroring was started for a volume, a new container, if not present, was created for each container in the source volume and the new containers were deleted if the mirroring was stopped. While deleting the new containers, the volume mirror module missed the last container in each iteration because the volume mirror module was incrementing the start key container ID (CID) during each iteration.

Resolution

With this fix, the volume mirror module will query the list of containers without missing a container and delete them.