

Release Notes for Patches for the MapR 5.2.0 Release

Release Notes, December 2016

Released 12/09/2016

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

Packages

RedHat	Server	mapr-patch-5.2.0.39122.GA-40967.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.2.0.39122.GA-40967.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-40967.x86_64.rpm
Red Hat	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-40967.x86_64.rpm
Red Hat	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-40967.x86_64.rpm
Ubuntu	Server	mapr-patch-5.2.0.39122.GA-40967.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.2.0.39122.GA-40967.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-40967.x86_64.deb
Ubuntu	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-40967.x86_64.deb
Ubuntu	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-40967.x86_64.deb
Win32	Client	mapr-client-5.2.0.40967GA-1.win32.zip
Win64	Client	windows/mapr-client-5.2.0.40967GA-1.amd64.zip
Mac	Client	mapr-client-5.2.0.40967GA-1.x86_64.tar.gz

Fixes:

Bug 12856

Description

When the `hadoop fs -rmr` command is run, it reads entire directory contents into memory before starting to delete anything resulting in Out Of Memory error.

Resolution

This fix includes a new `hadoop mfs -rmr <path>` command that:

- Will not build entire readdir file list in memory and once 1MB of readdir data is reached, the command will unlink and remove those directories.
- Will not fetch the attributes of the entries in readdir.

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 23257

Description

In MCS, new NFS VIPs were visible in the **NFS HA > VIP Assignments** tab, but not in the **NFS HA > NFS Setup** tab.

Resolution

With this fix, the NFS VIPs will be available in both the **NFS HA > VIP Assignments** tab and the **NFS HA > NFS Setup** tab.

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 24155

Description

Disk setup was timing out if running trim on flash drives took some time.

Resolution

With this fix, disk setup will complete successfully and the warning message (“Starting Trim of SSD drives, it may take a long time to complete”) is entered in the log file.

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 24352

Description

In this patch, the mirror synchronization has been optimized for changes in a small percentage of the inodes. During mirror resync operation, the destination will send the recent version number from the last mirror resync operation. While scanning inodes to identify the inodes that

have changed since the last resync operation, MFS will now compare the version number sent by the destination with the allocation group, which keeps track of all the inodes. If the allocation group version is:

- Higher than the last resync version, then MFS will check for the changed inodes in the allocation group.
- Less than or equal to the last resync version, MFS will not read all the inodes in the allocation group because the allocation group has not changed since the last resync operation.

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 24846

Description

If the topology of a node changed, after a CLDB failover, the list of nodes under a topology could not be determined as the new non-leaf topologies were not being updated.

Resolution

With this fix, the inner nodes of topology graph will be updated correctly and the list of nodes under an inner (non-leaf) topology will be determined correctly.

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 request 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 24915

Description

In version 5.1, running the `expandaudit` utility on volumes can result in very large (more than 1GB) audit log files due to incorrect GETATTR (get attributes) cache handling.

Resolution

With this fix, the `expandaudit` utility has been updated so that it will not perform subsequent GETATTR calls if the original call to the same file identifier failed.

Bug 24610

Description

In a secure cluster, when there are intermittent connection drops (between MFS-MFS or client-MFS), the client and/or server could crash during authentication.

Resolution

With this fix, the client and/or server will not crash during authentication if there are intermittent connection drops.

Bug 24585

Description

Excessive logging in CLDB audit caused `cldbaudit.log` file to grow to large sizes.

Resolution

With this fix, to reduce the size of `cldbaudit.log` file, the queries to CLDB for ZK string will no longer be logged for auditing.

Bug 25177

Description

When using FairScheduler with `maxAMShare` enabled, total `amResourceUsage` per queue is not calculated properly, which may cause applications to hang in ACCEPTED state.

Resolution:

AM resource usage is now calculated as expected and YARN jobs no longer get stuck in the ACCEPTED state.

Bug 25184

Description

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

Resolution

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

Bug 25290

Description

Sometimes, while the writes were in progress, the FUSE process crashed and the group IDs of the user changed.

Resolution

With this fix, the FUSE process will not crash while writes are in progress.

Bug 25426

Description

The server was rejecting encrypted writes as the expected length was not matching the RPC data length and this caused the server to crash.

Resolution

With this fix, the server will no longer crash as the expected length will always match the RPC data length for encrypted writes.

Release Notes, October 2016

Released 10/24/2016

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

Packages

RedHat	Server	mapr-patch-5.2.0.39122.GA-40164.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.2.0.39122.GA-40164.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-40164.x86_64.rpm
Red Hat	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-40164.x86_64.rpm
Red Hat	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-40164.x86_64.rpm
Ubuntu	Server	mapr-patch-5.2.0.39122.GA-40164.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.2.0.39122.GA-40164.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-40164.x86_64.deb
Ubuntu	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-40164.x86_64.deb

Ubuntu	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-40164.x86_64.deb
Win32	Client	mapr-client-5.2.0.40164GA-1.win32.zip
Win64	Client	windows/mapr-client-5.2.0.40164GA-1.amd64.zip
Mac	Client	mapr-client-5.2.0.40164GA-1.x86_64.tar.gz

Fixes:

Bug 11349

Description

Hoststats did not work on POSIX edge node.

Resolution

With this fix, hoststats can work on POSIX client edge nodes as well to display the statistics on MCS.

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 24408

Description

When running multiple producers as separate threads within a process, with a very small value for buffer.memory (say 1KB), some producers can stall. This is due to a lack of buffer memory.

Resolution

With this fix, the default value for minimum buffer memory is increased to 10kB.

Bug 24477

Description

Jobs failed if a local volume was not available and directories for mapreduce could not be initialized.

Resolution

With this fix, jobs no longer fail, and local volume recovery is enhanced.

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 24566

Description

An older version of the aws-sdk jar was built with MapR.

Resolution

With this fix, MapR upgraded the aws-sdk jar from version 1.7.4 to 1.7.15.

Bug 24630

Description

Under some conditions, using the 'ls' command with --full-time option produced incorrect results that showed as a negative number.

Resolution

With this fix, the correct timestamp is supplied.

Bug 24647

Description

On a node with multiple host IDs, CLDB crashed and failed over to new a new CLDB when a stale host ID was removed.

Resolution

With this fix, CLDB will not crash and fail over when a stale host ID is removed.

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 24658

Description

CLDB returned “no master” and an empty list for container lookup, which NFS server could not handle, because when multiple servers are down, there can be no master for a container.

Resolution

With this fix, NFS server will handle empty node list for container lookup.

Bug 24660

Description

MFS crashed because the maximum number of slots for backgrounded delete operations was not adequate. The incoming client operations reserving these slots were hanging and causing MFS to crash.

Resolution

With this fix, MFS will not crash as the number of slots for background operations has been increased.

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.

Bug 24712

Description

During container resynchronization, the same scratch space was being reused by internal parallel operations resulting in corruption.

Resolution

With this fix, internal parallel operations will use separate scratch spaces.

Bug 24992

Description

Installing a MapR patch caused jar files to be removed from under the drill/drill-1.4.0/jars/ directory

Resolution

Jar files are no longer incorrectly removed

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, September 2016

Released 9/23/2016

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

Packages

RedHat	Server	mapr-patch-5.2.0.39122.GA-39745.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.2.0.39122.GA-39745.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-39745.x86_64.rpm
Red Hat	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-39745.x86_64.rpm
Red Hat	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-39745.x86_64.rpm
Ubuntu	Server	mapr-patch-5.2.0.39122.GA-39745.x86_64.deb
Ubuntu	Client	mapr-patch-client-5.2.0.39122.GA-39745.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-39745.x86_64.deb
Ubuntu	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-39745.x86_64.deb
Ubuntu	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-39745.x86_64.deb
Win32	Client	mapr-client-5.2.0.39745GA-1.win32.zip
Win64	Client	windows/mapr-client-5.2.0.39745GA-1.amd64.zip
Mac	Client	mapr-client-5.2.0.39745GA-1.x86_64.tar.gz

Fixes:

Bug 11349

Description

Hoststats did not work on POSIX edge node.

Resolution

With this fix, hoststats can work on POSIX client edge nodes as well to display the statistics on MCS.

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 23975

Description

In version 5.1, MFS was failing to start on some docker containers as it was trying to figure out number of numa nodes from `/sys/devices/system/node`.

Resolution

With this fix, MFS will work on docker containers.

Bug 23981

Description

During remote writes, MFS was doing CRC checksum in the MFS thread and this caused MFS to churn a lot of CPU.

Resolution

With this fix, MFS will perform CRC checksum in other threads.

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 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.

Bug 24159

Description

The mtime was updated whenever a hard link was created. Also, when a hard link was created from the FUSE mount point, although the ctime was updated, the update timestamp only showed the minutes and seconds and not the nanoseconds.

Resolution

With this fix, mtime will not change on the hard link and when a hard link is created from the FUSE mount point, the timestamp for ctime will include nanoseconds.

Bug 24280

Description

Running the `maprcli dashboard info` command occasionally throws a `TimeoutException` error.

Resolution

With this fix, the internal timeout command was increased to provide more allowance for command processing.

Bug 24232

Description

In certain cases, files were created with stale chunk IDs, which prevented users from accessing files in the parent directory.

Resolution

With this fix, files will not be created with stale chunk IDs.

24315

Description

If you use the NFS client and you used the `dd` command with `iflag=direct`, an incorrect amount of data may have been read.

Resolution

With this fix, the dd command will read exactly the expected amount of data when iflag=direct is set.

Bug 24324

Description

The “disk not found” error was thrown because the script used to list disks was looking up disks in every instance of the fileserver process.

Resolution

With this fix, the script will look for disks only in the specific instance of the fileserver process.

24392

Details

Hadoop 2.7.1 uses a new state store data format. When you upgrade from Hadoop 2.5.1 to Hadoop 2.7.1, the ResourceManager fails to recover the state because it cannot handle the old state store data format.

Resolution:

With this fix, when the Resource Manager reads the data format used by Hadoop 2.5.1, it will use the Hadoop 2.5.1 method to read the state store data and recovery succeeds.

24413

Description

CLDB was crashing when volume replication was greater than 3.

Resolution

With this fix, CLDB will not crash when volume replication factor is greater than 3.

Bug 24446

Description

Due to incorrect attribute cache handling in NFS server, the getattr call sometimes returned stale mtime because the attribute cache was not getting updated properly at the time of setattr.

Resolution

With this fix, the attributes are now properly cached.

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.

Release Notes for the August 2016 Patch

Released 8/27/2016

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

Packages

RedHat	Server	mapr-patch-5.2.0.39122.GA-39350.x86_64.rpm
Red Hat	Client	mapr-patch-client-5.2.0.39122.GA-39350.x86_64.rpm
Red Hat	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-39350.x86_64.rpm
Red Hat	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-39350.x86_64.rpm
Red Hat	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-39350.x86_64.rpm
Ubuntu	Server	mapr-patch-5.2.0.39122.GA-39350.x86_64.deb

Ubuntu	Client	mapr-patch-client-5.2.0.39122.GA-39350.x86_64.deb
Ubuntu	Loopbacknfs	mapr-patch-loopbacknfs-5.2.0.39122.GA-39350.x86_64.deb
Ubuntu	Posix-client-basic	mapr-patch-posix-client-basic-5.2.0.39122.GA-39350.x86_64.deb
Ubuntu	Posix-client-platinum	mapr-patch-posix-client-platinum-5.2.0.39122.GA-39350.x86_64.deb
Win32	Client	mapr-client-5.2.0.39350GA-1.win32.zip
Win64	Client	mapr-client-5.2.0.39350GA-1.amd64.zip
Mac	Client	mapr-client-5.2.0.39350GA-1.x86_64.tar.gz

Fixes

Bug 20498 and 24143

Details

JobTracker attempts to restart TaskTrackers resulted in the loading of the job configuration object multiple times. Loading the job configuration object multiple times caused a JobTracker lock contention and it also caused the JobTracker service to become unresponsive.

Resolution

With this fix, the JobTracker caches the job configuration object for each job. Then, it uses the cached configuration object associated with each job for all of the job's task completion events.

Bug 23530 and 23260

Details

In certain cases, containers were stuck in incomplete state when either the request for role or response from the replicate state command was lost.

Resolution

With this fix, containers in incomplete state as a result of lost request (for role) or response (from the replicate state command) are automatically fixed.

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.

Bug 23931

Details

Nested user queues did not inherit labels from their parent queue. As a result, no labels were configured for nested user queues.

Resolution

With this fix, each nested user queue inherits the label and label policy from its parent queue.

Bug 24054

Details

When MapR-DB is used for time series data, a large number of tablets are generated with empty partitions. This can be merged into single tablet but, in this case, the single tablet will have thousands of empty partitions. Having a single tablet with thousands of empty partitions may result in a degradation of MapR-DB performance.

Resolution

With this fix, bucket flush is optimized for the tablet when a tablet has thousands of partitions.

Bug 24056

Details

When user resolution on the server failed, the client crashed while printing the error message because not enough number of arguments were passed to the gtrace function.

Resolution

With this fix, the `username` parameter will now be passed to the gtrace function.

Bug 24059

Details

There was no way to retrieve the list of containers that were unaware of the rack.

Resolution

With this fix, the `-queue` option in `maprcli dump replicationmanagerqueueinfo` command takes 5 as value to return the list of containers which are not rack aware.

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 24097

Details

Retrieval of zookeeper address for every hadoop command required a scan of the complete cluster information, which made `cldb` threads busy and unresponsive for new requests causing hadoop commands to take more time.

Resolution

With this fix, the complete cluster will not be scanned to retrieve zookeeper information.

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.

Bug 24170

Details

If the new ($\geq 5.1.0$) FileClient talks to older clusters, applications like YARN, which requires the ZooKeeper string, might fail because of the new way of retrieving the ZooKeeper address (which is available on clusters $\geq 5.1.0$).

Resolution

With this fix, the FileClient will fallback to the older way of retrieving the ZooKeeper information from CLDB if the new way of retrieving the information does not work.

Bug 24190

Details

During table replication on 10,000 tables, the destination table takes a long time (30 minutes or so) to catch up with the primary table. This is due to a bug in idleWorker which resulted in open buckets being flushed only during bucketflush instead of by the idleWorker.

Resolution

With this fix, idleWorker flushes idle open buckets for replication every 2 seconds.

Bug 24238

Details

MFS crashed with asserts in Unlink and create because of an issue with rename when the target file had multiple hard links.

Resolution

With this fix, MFS will no longer crash with asserts.