

Hortonworks Data Platform Win

Release Notes

(May 21, 2013)

Hortonworks Data Platform Win 1.1.0 (GA) : Release Notes

Copyright © 2013 Hortonworks, Inc. All rights reserved.

The Hortonworks Data Platform, powered by Apache Hadoop, is a massively scalable and 100% open source platform for storing, processing and analyzing large volumes of data. It is designed to deal with data from many sources and formats in a very quick, easy and cost-effective manner. The Hortonworks Data Platform consists of the essential set of Apache Hadoop projects including MapReduce, Hadoop Distributed File System (HDFS), HCatalog, Pig, Hive, HBase, Zookeeper and Ambari. Hortonworks is the major contributor of code and patches to many of these projects. These projects have been integrated and tested as part of the Hortonworks Data Platform release process and installation and configuration tools have also been included.

Unlike other providers of platforms built using Apache Hadoop, Hortonworks contributes 100% of our code back to the Apache Software Foundation. The Hortonworks Data Platform is Apache-licensed and completely open source. We sell only expert technical support, [training](#) and partner-enablement services. All of our technology is, and will remain free and open source. Please visit the [Hortonworks Data Platform](#) page for more information on Hortonworks technology. For more information on Hortonworks services, please visit either the [Support](#) or [Training](#) page. Feel free to [Contact Us](#) directly to discuss your specific needs.

Licensed under the Apache License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at

<http://www.apache.org/licenses/LICENSE-2.0>

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

Table of Contents

1. Release Notes HDP-Win-1.1.0 (GA)	1
1.1. Product Version: HDP-Win-1.1.0 (GA)	1
1.2. Patch Information	1
1.2.1. Patch Information for Hadoop	2
1.2.2. Patch Information for Pig	7
1.2.3. Patch Information for Hive	8
1.2.4. Patch Information for HCatalog	8
1.3. Improvements	9
1.4. Known Issues	9
2. Release Notes HDP-Win-1.1.0	11
2.1. Product Version: HDP-Win-1.1.0	11
2.2. Patch Information	11
2.2.1. Patch Information for Hadoop	11
2.2.2. Patch Information for Pig	17
2.2.3. Patch Information for Hive	18
2.2.4. Patch Information for HCatalog	18
2.3. Known Issues	18

1. Release Notes HDP-Win-1.1.0 (GA)

RELEASE NOTES: Hortonworks Data Platform 1.1.0 (GA) for Windows powered by Apache Hadoop In this document:

- [Product Version: HDP-Win-1.1.0 \(GA\)](#)
- [Patch Information](#)
- [Improvements](#)
- [Known Issues](#)

1.1. Product Version: HDP-Win-1.1.0 (GA)

All HDP 1.1.0 components listed here are official Apache releases of the most recent stable versions available. Hortonworks' philosophy is to provide patches only when absolutely necessary to assure the interoperability of the components. Unless you are explicitly directed by Hortonworks Support to take a patch update, each of the HDP 1.1.0 components needs to remain at the following package version levels to ensure a certified and supported copy of HDP 1.1.0.

This release of Hortonworks Data Platform (HDP) deploys the following Hadoop-related components:

- Apache Hadoop 1.0.3
- Apache Pig 0.9.3
- Apache HCatalog 0.4.1
- Apache Templeton 0.1.4
- Apache Hive 0.9.0
- Apache Sqoop 1.4.2
- Apache Oozie 3.2.0

1.2. Patch Information

In this section:

- [Patch information for Hadoop](#)
- [Patch information for Pig](#)
- [Patch information for Hive](#)
- [Patch information for HCatalog](#)

1.2.1. Patch Information for Hadoop

The following list provides information on the **Apache JIRAs** for this release:

- [HDFS-6527](#): Backport HADOOP-7389: Use of TestingGroups by tests causes subsequent tests to fail.
- [HADOOP-8836](#): UGI should throw exception in case `winutils.exe` cannot be loaded.
- [MAPREDUCE-4657](#): WindowsResourceCalculatorPlugin has Null Pointer Exception.
- [HADOOP-8903](#): Added support `HADOOP_USER_CLASSPATH_FIRST` in Windows Hadoop cmd.
- [HADOOP-8908](#): Refactor `winutil.exe` related code.
- [HADOOP-8911](#): CRLF characters in source and text files.
- [HADOOP-8912](#): Add `.gitattributes` file to prevent CRLF and LF mismatches for source and text files.
- [HADOOP-8868](#): `FileUtil#chmod` should normalize the path before calling into shell APIs.
- [HDFS-4065](#): `TestDFSShell.testGet` sporadically fails attempting to corrupt block files due to race condition.
- [HADOOP-8902](#): Enabled Gridmix v1 and v2 benchmarks on the Windows platform.
- [HADOOP-8564](#): Port and extend Hadoop native libraries for Windows to address DataNode concurrent reading and writing issue.
- [HADOOP-4093](#): Fix a bug that `AzureBlockPlacementPolicy#chooseTarget` only returns one DataNode when replication factor is greater than three.
- [HADOOP-8907](#): Provide means to look for `zlib1.dll` next to `hadoop.dll` on Windows.
- [HADOOP-8420](#): Hadoop Common creating `package-info.java` must not depend on `sh`.
- [HADOOP-8972](#): Move `winutils` tests from `bat` to Java.
- [HADOOP-9026](#): `Hadoop.cmd` fails to initialize if user's `%path%` variable has parenthesis.
- [HADOOP-9027](#): Build fails on Windows without `sh/sed/echo` in the path.
- [HADOOP-8847](#): Change `untar` to use Java API on Windows instead of spawning `tar` process.
- [HADOOP-9062](#): `hadoop-env.cmd` overwrites the value of `*_OPTS` set before install.
- [HADOOP-9074](#): Hadoop install scripts for Windows.
- [HADOOP-9110](#): `winutils` is off-by-one error indexing `MONTHS` array can cause access violation.

- [HADOOP-9102](#): winutils task `isAlive` does not return a non-zero exit code if the requested task is not alive.
- [HADOOP-9061](#): Java6+Windows does not work well with symlinks.
- [HADOOP-8645](#): `HADOOP_HOME` and `-Dhadoop.home` (from `hadoop` wrapper script) are not uniformly handled.
- [HADOOP-9185](#): `TestFileCreation.testFsClose` should clean up on exit.
- [HADOOP-9191](#): `TestAccessControlList` and `TestJobHistoryConfig` fail with JDK7.
- [HADOOP-9177](#): Address issues that reported by static code analysis on winutils.
- [HADOOP-9250](#): Windows installer bugfixes
- [MAPREDUCE-4915](#): `TestShuffleExceptionCount` fails with open JDK7
- [MAPREDUCE-4914](#): `TestMiniMRDFSsort` fails with openJDK7.
- [HADOOP-9179](#): `TestFileSystem` fails with open JDK7.
- [HDFS-4358](#): `TestCheckpoint` failure with JDK7.
- [HDFS-4355](#): `TestNameNodeMetrics.testCorruptBlock` fails with open JDK7
- [MAPREDUCE-4909](#): `TestKeyValueTextInputFormat` fails with Open JDK 7 on Windows
- [HADOOP-9175](#): `TestWritableName` fails with Open JDK 7
- [HADOOP-9174](#): `TestSecurityUtil` fails with Open JDK 7
- [HDFS-4337](#): Backport HDFS-4240: For nodegroup-aware block placement, when a node is excluded, the nodes in the same nodegroup should also be excluded.
- [HDFS-4341](#): Set default data dir permission in `MiniDFSclusterWithNodeGroup`.
- [HDFS-4320](#): Add a separate configuration for NameNode RPC address instead of using `fs.default.name`.
- [HDFS-3942](#): Backport HDFS-3495 and HDFS-4234: Update Balancer to support new `NetworkTopology` with `NodeGroup` and use generic code for choosing `DataNode` in Balancer.
- [MAPREDUCE-782](#): Use `PureJavaCrc32` in MapReduce spills.
- [HDFS-496](#): Backport: Use `PureJavaCrc32` in HDFS.
- [HADOOP-7096](#): Allow setting of end-of-record delimiter for `TextInputFormat`.
- [HADOOP-8617](#): Backport HADOOP-6148, HADOOP-6166 and HADOOP-7333 for a pure Java CRC32 calculator implementation.

- [HADOOP-9111](#): Change some JUnit 3 tests to JUnit 4 so that `@Ignore` tests can be run with Ant v1.8.x.
- [HADOOP-9090](#): Support on-demand publish of metrics.
- [HADOOP-9099](#): `TestNetUtils` fails if "UnknownHost" is resolved as a valid hostname.
- [HADOOP-9095](#): Backport [HADOOP-8372](#): `NetUtils.normalizeHostName()` incorrectly handles hostname starting with a numeric character.
- [HADOOP-9036](#): Fix racy test case `TestSinkQueue`
- [HADOOP-8900](#): `BuiltInGzipDecompressor` throws `IOException` - stored gzip size doesn't match decompressed size.
- [HADOOP-6496](#): `HttpServer` sends wrong content-type for CSS files.
- [HADOOP-7868](#): Hadoop native fails to compile when default linker option is `-Wl,-as-needed`.
- [HDFS-3941](#): Backport [HDFS-3498](#) and [HDFS-3601](#): Support replica removal in `BlockPlacementPolicy` and make `BlockPlacementPolicyDefault` extensible for reusing code in subclasses, and add `BlockPlacementPolicyWithNodeGroup` to support block placement with 4-layer network topology.
- [HADOOP-8820](#): Backport [HADOOP-8469](#) and [HADOOP-8470](#): Make `NetworkTopology` class pluggable and add `NetworkTopologyWithNodeGroup`, a 4-layer implementation of `NetworkTopology`.
- [HADOOP-8645](#): `HADOOP_HOME` and `-Dhadoop.home` (from hadoop wrapper script) are not uniformly handled.
- [MAPREDUCE-1806 \(HADOOP-136\)](#): Fixed issues in `CombineFileInputFormat` that caused failure while using Sqoop to export files in ASV.
- [HADOOP-8874](#): Added an API to retrieve valid `HADOOP_HOME` and `bin` path. This patch adds a consistency layer for `HADOOP_HOME` lookups and provides abstractions to qualify bin paths of hadoop binary components.
- [HADOOP-8731](#): Added Public distributed cache support for Windows.
- [HADOOP-9040](#): Added fixes for `TaskController`.
- [HADOOP-9071](#): Improved Ivy log levels.
- [HADOOP-8872](#): Fixed issue caused while invoking `FileSystem.length()` method on a Windows machine using JDK 6.x.
- [MAPREDUCE-4564](#): Fixed shell timeout mechanism. This fix enables successful termination of those processes that are spawned by `Winutils`.
- [MAPREDUCE-4561](#): Added support for node health scripts on Windows
- [MAPREDUCE-4597](#): Fixed intermittent failures for `TestKillSubProcesses`.

- [HADOOP-8739](#): Fixed command line parsing on Windows.
- [HADOOP-8664](#): Fixed the Hadoop streaming job issue that required the user to provide full path to commands.
- [HDFS-3766](#): Fixed TestStorageRestore on Windows.
- [HADOOP-8634](#): Fixed the errors caused when FileSystem deleteonExit method is invoked.
- [HDFS-3763](#): Fixed the TestNameNodeMXBean failures on Windows.
- [MAPREDUCE-4510](#): Fixed redundant checks and logging of **getconf** on Windows.
- [HADOOP-8731](#): Added public distributed cache support for Windows. This fixes the failures for TestTrackerDistributedCacheManager.
- [HADOOP-8763](#): Fixed issues caused when setting group owner on Windows.
- [HADOOP-8694](#): Added symlink support to Windows platform.
- [HADOOP-8732](#): Fixed test failures caused due to incorrect process serialization on Windows.
- [HDFS-3564](#): Added enhancements to the block placement policy. This enhancement enables a pluggable placement policy and provides a new API that enables moving blocks for balancing. It also enables the placement policy to decide the number of racks and provides ability to extend the policy.
- [HDFS-3566](#): Add AzureBlockPlacementPolicy to handle fault and upgrade domains in Azure. This policy distributes replicas across both the fault and the upgrade domains to ensure zero data loss.
- [HADOOP-8457](#): Fixed the file ownership issue for users in the Administrators groups on Windows.
- [MAPREDUCE-4374](#): Added support for configurable environment for child map/reduce tasks on Windows.
- [HADOOP-8453](#): Added unit tests for Winutils program. Winutils is the Windows console program that emulates the Linux command line utilities used by Hadoop.
- [HDFS-385](#): Added new experimental API BlockPlacementPolicy allows investigating alternate rules for locating block replicas.
- [HADOOP-8899](#): Fixed issues caused because of Classpath exceeding the maximum operating system (OS) limit.
- [MAPREDUCE-1806](#): Fixed issues with CombineFileInputFormat.
- [HADOOP-8935](#): Improved Winutils to handle the failures caused for the **winutils ls** command.
- [HADOOP-6496](#): Fixed the HTTPServer issue that caused incorrect rendering of the web interface for HBase.

- [HADOOP-7827](#): Fixed issue with JSP pages for web interfaces.
- [HADOOP-8903](#): Added support for `HADOOP_USER_CLASSPATH_FIRST` environment variable in the `hadoop.cmd` file.
- [HADOOP-8880](#): Fixed Hive test failures caused because of missing Jersey JAR files in the POM template.
- [HADOOP-8733](#): Fixed the failures caused by the dependencies in the `test.sh` script file.
- [MAPREDUCE-4400](#): Fixed performance regression for small jobs and workflows.
- [HADOOP-8734](#): Fixed LocalJobRunner to support private distributed cache.
- [HDFS-3833](#): Fixed TestDFSShell failures on Windows caused due to concurrent file read/write.
- [MAPREDUCE-4598](#): Added support for node health scripts on Windows.
- [HADOOP-8657](#): Fixed TestCLI to remove hardcoded value for the file length.
- [HDFS-3163](#): Fixed failures for TestHDFSCLI.testAll which occurred when the user name was not provided in lowercase.
- [HADOOP-8618](#): Fixed build failures caused due to merging of the Hadoop v1.0.3 branch.
- [HADOOP-8544](#): Moved an assertion location in `winutils chmod` command.
- [HADOOP-7389](#): Fixed test failures caused when tests use the TestingGroups.
- [HADOOP-8414](#): Fixed issues caused by localhost resolving to incorrect address on Windows.
- [MAPREDUCE-4368](#): Fixed TaskRunner to handle the event when `java.library.path` contains a quoted path with embedded spaces on Windows platform.
- [MAPREDUCE-4369](#): Fixed streaming job failures with WindowsResourceCalculatorPlugin.
- [MAPREDUCE-4332](#): Fixed command length abort issues on Windows.
- [HADOOP-8487](#): Fixed the HDFS tests to use correct test paths.
- [HADOOP-8534](#): Fixed failures for those tests that left configuration files open.
- [HADOOP-8486](#): Fixed the resource leak caused because of open file handles for SequenceFile.
- [MAPREDUCE-4203](#): Added an implementation of the process tree for Windows.
- [HADOOP-8454](#): Fixed bugs for the `chmod` command in Winutils program.
- [HADOOP-8409](#): Fixed TestCommandLineJobSubmission and TestGenericOptionsParser to work for Windows.
- [MAPREDUCE-4260](#): Added support to use JobObject for spawning tasks on Windows platform.

- [MAPREDUCE-4321](#): Fixed failures for DefaultTaskController on Windows.
- [HADOOP-8424](#): Fixed Classpath issues that caused intermittent failures for web user interface on Windows.
- [HDFS-3424](#): Fixed TestDatanodeBlockScanner and TestReplication failures on Windows.
- [HADOOP-8374](#): Improved support for hard link manipulation on Windows.
- [HADOOP-8440](#): Fixed failures for HarFileSystem.decodeHarURI.
- [HADOOP-8411](#): Fixed TestStorageDirectorFailure, TestTaskLogsTruncater, TestWebHdfsUrl and TestSecurityUtil failures on Windows.
- [HADOOP-8235](#): Added support file permissions and ownership on Windows for RawLocalFileSystem.
- [MAPREDUCE-4204](#): Improved ProcfsBasedProcessTree to enable the resource collection object to be pluggable.
- [MAPREDUCE-4201](#): Fixed issues related to obtaining PIDs on Windows.
- [HADOOP-8234](#): Enabled user group mappings on Windows platform.
- [HADOOP-8223](#): Applied initial patch for branch-1-win.
- [HDFS-4413](#): Secondary namenode won't start if HDFS isn't the default file system

1.2.2. Patch Information for Pig

The following list provides information on the **Apache JIRAs** for this release:

- [PIG-2958](#): Fixed failure issues caused when Pig tests have no associated logger.
- [PIG-2957](#): Fixed failures for TetsScriptUDF test.
- [PIG-2960](#): Increased the timeout for unit tests on Windows.
- [PIG-2943](#): Improved DevTests and Windows checks to include the **Util.Windows** method.
- [PIG-2942](#): Fixed failures for DevTests and TestLoad tests.
- [PIG-2941](#): Added consistency in chaining the Ivy resolvers. This fix also adds the fallback mechanism.
- [PIG-2953](#): Added support for **which** utility on Windows
- [PIG-2956](#): Fixed issues with invalid cache specification for some streaming statement.
- [PIG-2954](#): Fixed test failures caused due to the dependency on **bash**.
- [PIG-2959](#): Fixes for `pig.cmd` to run on Windows.
- [PIG-2801](#): Fixes for grunt **sh** command.

- [PIG-2800](#): Fixed issues with `pig.additional.jars` path separator.
- [PIG-2799](#): Updated Pig streaming interface to run correctly on Windows without Cygwin support.
- [PIG-2798](#): Fixed issues with Pig streaming tests on Windows.
- [PIG-2797](#): Fixed the Pig tests to use `Util.generateURI` method.
- [PIG-2796](#): Fixed issues with invalid path names on HDFS caused because the Pig tests use local temporary paths.
- [PIG-2795](#): Added support to handle paths generated on Windows.
- [PIG-2794](#): Added utilities to facilitate testing on Windows platform.
- [PIG-2793](#): Improved Pig to work on Windows platform without Cygwin support.

1.2.3. Patch Information for Hive

The following list provides information on the **Apache JIRAs** for this release:

- [HIVE-3025](#): Fixed Hive archive command for Hadoop v 0.22 and 0.23.
- [HIVE-3448](#): Fixed failures for the `skewjoin.q` testcase on Windows.
- [HIVE-3441](#): Fixed failures caused due to the partition column strings in Windows file names.
- [HIVE-3436](#): Fixed the `script_pipe.q` failures on Windows.
- [HIVE-3483](#): Fixed issues with joins that use partitioned table on Windows.
- [HIVE-3317](#): Fixed `TestDocToUnix` unit tests on Windows.
- [HIVE-3320](#): Fixed test case failures caused by incorrect handling of CRLF line endings on Windows.
- [HIVE-3319](#): Fixed path related issues that caused the unit test failures for Windows.
- [HIVE-3327](#): Fixed failures caused while execution of the `/bin/cat` script files on Windows.
- [HIVE-3479](#): Fixed issues with negative unit tests.
- [HIVE-3494](#): Fixed JDBC test case failures on Windows.
- [HIVE-3480](#): Fixed file handle leaks in Symbolic and symlink related input formats.

1.2.4. Patch Information for HCatalog

The following list provides information on the **Apache JIRAs** for this release:

- [HCATALOG-512](#): Fixed HCatalog unit tests on Windows.

- [HCATALOG-514](#): Fixed HCatalog python scripts in the package build for Windows.

1.3. Improvements

- Added ability to perform push install from a central computer.
- Added ability to uninstall and reinstall over existing data.

Use the following instructions to upgrade from HDP-Win-1.1:

1. Uninstall HDP using `DESTROY_DATA=no`. On each cluster host, execute the following command from the command shell:

```
msiexec /x "$MSI_PATH" /lv "$PATH_to_Installer_Log_File" DESTROY_DATA="no"
```

2. Reinstall HDP using `DESTROY_DATA=no`. On each cluster host, execute the following command from the command shell:

```
msiexec /x "$MSI_PATH" /lv "$PATH_to_Installer_Log_File" DESTROY_DATA="no"
```

- Improved error feedback during deployments.

1.4. Known Issues

- One of the regression test for Oozie fails with MS SQL because `java.io.IOException` does not finish within the `600000timeout`.
- One of the regression test for Oozie fails with MS SQL because `java.io.UnsupportedEncodingException` does not finish within the `600000timeout`.
- Running MapReduce jobs using pipes is currently not supported on Windows.
- Non-default compression codecs that are based on zlib or snappy are currently not supported on Windows
- It is possible to encounter the following exception while starting the Hive command line interface (CLI).

```
FAILED: Error in metadata: javax.jdo.JDOFatalDataStoreException: DERBY SQL
error: SQLCODE: -1, SQLSTATE: XJ041, SQLERRMC:
    Failed to create database 'metastore_db', see the next exception for
    details.
    ::SQLSTATE: XBM0JDirectory C:\Hadoop\hive-0.9.0\bin\metastore_db already
    exists.
```

This typically happens when the user attempts to install and uninstall HDP repeatedly. In such cases, the directories for Hive might not get deleted properly on uninstall. You can use any of the following workaround:

- Manually delete the `metastore_db` directory (`C:\Hadoop\hive-0.9.0\bin\metastore_db`).
- Uninstall HDP, delete all the files in the Hadoop directory (`C:\Hadoop`), and install HDP again

- HDP creates some files in the HDFS directory (`C:\hdfs`) that are not deleted on uninstall. This issue is observed when `hadoop.tmp.dir` is not defined to point to the `C:\hadoop` location. There is no impact on the deployment of your cluster. However, it is recommended that you manually delete the files in the HDFS directory (`C:\hdfs`) after you uninstall HDP.

2. Release Notes HDP-Win-1.1.0

RELEASE NOTES: Hortonworks Data Platform 1.1 for Windows (Beta) powered by Apache Hadoop

2.1. Product Version: HDP-Win-1.1.0

This release of Hortonworks Data Platform (HDP) deploys the following Hadoop-related components:

- Apache Hadoop 1.0.3
- Apache Pig 0.9.3
- Apache HCatalog 0.4.1
- Apache Templeton 0.1.4
- Apache Hive 0.9.0
- Apache Sqoop 1.4.2
- Apache Oozie 3.2.0

2.2. Patch Information

In this section:

- [Patch information for Hadoop](#)
- [Patch information for Pig](#)
- [Patch information for Hive](#)
- [Patch information for HCatalog](#)

2.2.1. Patch Information for Hadoop

The following list provides information on the **Apache JIRAs** for this release:

- [HDFS-6527](#): Backport HADOOP-7389: Use of TestingGroups by tests causes subsequent tests to fail.
- [HADOOP-8836](#): UGI should throw exception in case `winutils.exe` cannot be loaded.
- [MAPREDUCE-4657](#): WindowsResourceCalculatorPlugin has Null Pointer Exception.
- [HADOOP-8903](#): Added support `HADOOP_USER_CLASSPATH_FIRST` in Windows Hadoop cmd.
- [HADOOP-8908](#): Refactor `winutils.exe` related code.

- [HADOOP-8911](#): CRLF characters in source and text files.
- [HADOOP-8912](#): Add .gitattributes file to prevent CRLF and LF mismatches for source and text files.
- [HADOOP-8868](#): FileUtil#chmod should normalize the path before calling into shell APIs.
- [HDFS-4065](#): TestDFSShell.testGet sporadically fails attempting to corrupt block files due to race condition.
- [HADOOP-8902](#): Enabled Gridmix v1 and v2 benchmarks on the Windows platform.
- [HADOOP-8564](#): Port and extend Hadoop native libraries for Windows to address DataNode concurrent reading and writing issue.
- [HADOOP-4093](#): Fix a bug that AzureBlockPlacementPolicy#chooseTarget only returns one DataNode when replication factor is greater than three.
- [HADOOP-8907](#): Provide means to look for zlib1.dll next to hadoop.dll on Windows.
- [HADOOP-8420](#): Hadoop Common creating package-info.java must not depend on sh.
- [HADOOP-8972](#): Move winutils tests from bat to Java.
- [HADOOP-9026](#): Hadoop.cmd fails to initialize if user's %path% variable has parenthesis.
- [HADOOP-9027](#): Build fails on Windows without sh/sed/echo in the path.
- [HADOOP-8847](#): Change untar to use Java API on Windows instead of spawning tar process.
- [HADOOP-9062](#): hadoop-env.cmd overwrites the value of *_OPTS set before install.
- [HADOOP-9074](#): Hadoop install scripts for Windows.
- [HADOOP-9110](#): winutils ls off-by-one error indexing MONTHS array can cause access violation.
- [HADOOP-9102](#): winutils task isAlive does not return a non-zero exit code if the requested task is not alive.
- [HADOOP-9061](#): Java6+Windows does not work well with symlinks.
- [HADOOP-8645](#): HADOOP_HOME and -Dhadoop.home (from hadoop wrapper script) are not uniformly handled.
- [HADOOP-9185](#): TestFileCreation.testFsClose should clean up on exit.
- [HADOOP-9191](#): TestAccessControlList and TestJobHistoryConfig fail with JDK7.
- [HADOOP-9177](#): Address issues that reported by static code analysis on winutils.
- [HADOOP-9250](#): Windows installer bugfixes

- [MAPREDUCE-4915](#): `TestShuffleExceptionCount` fails with open JDK7
- [MAPREDUCE-4914](#): `TestMiniMRDFSsort` fails with openJDK7.
- [HADOOP-9179](#): `TestFileSystem` fails with open JDK7.
- [HDFS-4358](#): `TestCheckpoint` failure with JDK7.
- [HDFS-4355](#): `TestNameNodeMetrics.testCorruptBlock` fails with open JDK7
- [MAPREDUCE-4909](#): `TestKeyValueTextInputFormat` fails with Open JDK 7 on Windows
- [HADOOP-9175](#): `TestWritableName` fails with Open JDK 7
- [HADOOP-9174](#): `TestSecurityUtil` fails with Open JDK 7
- [HDFS-4337](#): Backport HDFS-4240: For nodegroup-aware block placement, when a node is excluded, the nodes in the same nodegroup should also be excluded.
- [HDFS-4341](#): Set default data dir permission in `MiniDFSclusterWithNodeGroup`.
- [HDFS-4320](#): Add a separate configuration for NameNode RPC address instead of using `fs.default.name`.
- [HDFS-3942](#): Backport HDFS-3495 and HDFS-4234: Update Balancer to support new `NetworkTopology` with `NodeGroup` and use generic code for choosing `DataNode` in Balancer.
- [MAPREDUCE-782](#): Use `PureJavaCrc32` in MapReduce spills.
- [HDFS-496](#): Backport: Use `PureJavaCrc32` in HDFS.
- [HADOOP-7096](#): Allow setting of end-of-record delimiter for `TextInputFormat`.
- [HADOOP-8617](#): Backport HADOOP-6148, HADOOP-6166 and HADOOP-7333 for a pure Java CRC32 calculator implementation.
- [HADOOP-9111](#): Change some JUnit 3 tests to JUnit 4 so that `@Ignore` tests can be run with Ant v1.8.x.
- [HADOOP-9090](#): Support on-demand publish of metrics.
- [HADOOP-9099](#): `TestNetUtils` fails if "UnknownHost" is resolved as a valid hostname.
- [HADOOP-9095](#): Backport HADOOP-8372: `NetUtils.normalizeHostName()` incorrectly handles hostname starting with a numeric character.
- [HADOOP-9036](#): Fix racy test case `TestSinkQueue`
- [HADOOP-8900](#): `BuiltInGzipDecompressor` throws `IOException` - stored gzip size doesn't match decompressed size.
- [HADOOP-6496](#): `HttpServer` sends wrong content-type for CSS files.

- [HADOOP-7868](#): Hadoop native fails to compile when default linker option is `-Wl,-as-needed`.
- [HDFS-3941](#): Backport HDFS-3498 and HDFS-3601: Support replica removal in `BlockPlacementPolicy` and make `BlockPlacementPolicyDefault` extensible for reusing code in subclasses, and add `BlockPlacementPolicyWithNodeGroup` to support block placement with 4-layer network topology.
- [HADOOP-8820](#): Backport HADOOP-8469 and HADOOP-8470: Make `NetworkTopology` class pluggable and add `NetworkTopologyWithNodeGroup`, a 4-layer implementation of `NetworkTopology`.
- [HADOOP-8645](#): `HADOOP_HOME` and `-Dhadoop.home` (from hadoop wrapper script) are not uniformly handled.
- [MAPREDUCE-1806 \(HADOOP-136\)](#): Fixed issues in `CombineFileInputFormat` that caused failure while using Sqoop to export files in ASV.
- [HADOOP-8874](#): Added an API to retrieve valid `HADOOP_HOME` and `bin` path. This patch adds a consistency layer for `HADOOP_HOME` lookups and provides abstractions to qualify `bin` paths of hadoop binary components.
- [HADOOP-8731](#): Added Public distributed cache support for Windows.
- [HADOOP-9040](#): Added fixes for `TaskController`.
- [HADOOP-9071](#): Improved Ivy log levels.
- [HADOOP-8872](#): Fixed issue caused while invoking `FileSystem.length()` method on a Windows machine using JDK 6.x.
- [MAPREDUCE-4564](#): Fixed shell timeout mechanism. This fix enables successful termination of those processes that are spawned by **Winutils**.
- [MAPREDUCE-4561](#): Added support for node health scripts on Windows
- [MAPREDUCE-4597](#): Fixed intermittent failures for `TestKillSubProcesses`.
- [HADOOP-8739](#): Fixed command line parsing on Windows.
- [HADOOP-8664](#): Fixed the Hadoop streaming job issue that required the user to provide full path to commands.
- [HDFS-3766](#): Fixed `TestStorageRestore` on Windows.
- [HADOOP-8634](#): Fixed the errors caused when `FileSystem.deleteOnExit` method is invoked.
- [HDFS-3763](#): Fixed the `TestNameNodeMXBean` failures on Windows.
- [MAPREDUCE-4510](#): Fixed redundant checks and logging of **getconf** on Windows.
- [HADOOP-8731](#): Added public distributed cache support for Windows. This fixes the failures for `TestTrackerDistributedCacheManager`.
- [HADOOP-8763](#): Fixed issues caused when setting group owner on Windows.

- [HADOOP-8694](#): Added symlink support to Windows platform.
- [HADOOP-8732](#): Fixed test failures caused due to incorrect process serialization on Windows.
- [HDFS-3564](#): Added enhancements to the block placement policy. This enhancement enables a pluggable placement policy and provides a new API that enables moving blocks for balancing. It also enables the placement policy to decide the number of racks and provides ability to extend the policy.
- [HDFS-3566](#): Add `AzureBlockPlacementPolicy` to handle fault and upgrade domains in Azure. This policy distributes replicas across both the fault and the upgrade domains to ensure zero data loss.
- [HADOOP-8457](#): Fixed the file ownership issue for users in the Administrators groups on Windows.
- [MAPREDUCE-4374](#): Added support for configurable environment for child map/reduce tasks on Windows.
- [HADOOP-8453](#): Added unit tests for Winutils program. Winutils is the Windows console program that emulates the Linux command line utilities used by Hadoop.
- [HDFS-385](#): Added new experimental API `BlockPlacementPolicy` allows investigating alternate rules for locating block replicas.
- [HADOOP-8899](#): Fixed issues caused because of Classpath exceeding the maximum operating system (OS) limit.
- [MAPREDUCE-1806](#): Fixed issues with `CombineFileInputFormat`.
- [HADOOP-8935](#): Improved Winutils to handle the failures caused for the `winutils ls` command.
- [HADOOP-6496](#): Fixed the `HTTPServer` issue that caused incorrect rendering of the web interface for HBase.
- [HADOOP-7827](#): Fixed issue with JSP pages for web interfaces.
- [HADOOP-8903](#): Added support for `HADOOP_USER_CLASSPATH_FIRST` environment variable in the `hadoop.cmd` file.
- [HADOOP-8880](#): Fixed Hive test failures caused because of missing Jersey JAR files in the POM template.
- [HADOOP-8733](#): Fixed the failures caused by the dependencies in the `test.sh` script file.
- [MAPREDUCE-4400](#): Fixed performance regression for small jobs and workflows.
- [HADOOP-8734](#): Fixed `LocalJobRunner` to support private distributed cache.
- [HDFS-3833](#): Fixed `TestDFSShell` failures on Windows caused due to concurrent file read/write.
- [MAPREDUCE-4598](#): Added support for node health scripts on Windows.

- [HADOOP-8657](#): Fixed TestCLI to remove hardcoded value for the file length.
- [HDFS-3163](#): Fixed failures for TestHDFSCLI.testAll which occurred when the user name was not provided in lowercase.
- [HADOOP-8618](#): Fixed build failures caused due to merging of the Hadoop v1.0.3 branch.
- [HADOOP-8544](#): Moved an assertion location in **winutils chmod** command.
- [HADOOP-7389](#): Fixed test failures caused when tests use the TestingGroups.
- [HADOOP-8414](#): Fixed issues caused by localhost resolving to incorrect address on Windows.
- [MAPREDUCE-4368](#): Fixed TaskRunner to handle the event when `java.library.path` contains a quoted path with embedded spaces on Windows platform.
- [MAPREDUCE-4369](#): Fixed streaming job failures with WindowsResourceCalculatorPlugin.
- [MAPREDUCE-4332](#): Fixed command length abort issues on Windows.
- [HADOOP-8487](#): Fixed the HDFS tests to use correct test paths.
- [HADOOP-8534](#): Fixed failures for those tests that left configuration files open.
- [HADOOP-8486](#): Fixed the resource leak caused because of open file handles for SequenceFile.
- [MAPREDUCE-4203](#): Added an implementation of the process tree for Windows.
- [HADOOP-8454](#): Fixed bugs for the **chmod** command in Winutils program.
- [HADOOP-8409](#): Fixed TestCommandLineJobSubmission and TestGenericOptionsParser to work for Windows.
- [MAPREDUCE-4260](#): Added support to use JobObject for spawning tasks on Windows platform.
- [MAPREDUCE-4321](#): Fixed failures for DefaultTaskController on Windows.
- [HADOOP-8424](#): Fixed Classpath issues that caused intermittent failures for web user interface on Windows.
- [HDFS-3424](#): Fixed TestDatanodeBlockScanner and TestReplication failures on Windows.
- [HADOOP-8374](#): Improved support for hard link manipulation on Windows.
- [HADOOP-8440](#): Fixed failures for HarFileSystem.decodeHarURI.
- [HADOOP-8411](#): Fixed TestStorageDirectorFailure, TestTaskLogsTruncater, TestWebHdfsUrl and TestSecurityUtil failures on Windows.
- [HADOOP-8235](#): Added support file permissions and ownership on Windows for RawLocalFileSystem.

- [MAPREDUCE-4204](#): Improved ProcfsBasedProcessTree to enable the resource collection object to be pluggable.
- [MAPREDUCE-4201](#): Fixed issues related to obtaining PIDs on Windows.
- [HADOOP-8234](#): Enabled user group mappings on Windows platform.
- [HADOOP-8223](#): Applied initial patch for branch-1-win.
- [HDFS-4413](#): Secondary namenode won't start if HDFS isn't the default file system

2.2.2. Patch Information for Pig

The following list provides information on the **Apache JIRAs** for this release:

- [PIG-2958](#): Fixed failure issues caused when Pig tests have no associated logger.
- [PIG-2957](#): Fixed failures for TetsScriptUDF test.
- [PIG-2960](#): Increased the timeout for unit tests on Windows.
- [PIG-2943](#): Improved DevTests and Windows checks to include the **Util.Windows** method.
- [PIG-2942](#): Fixed failures for DevTests and TestLoad tests.
- [PIG-2941](#): Added consistency in chaining the Ivy resolvers. This fix also adds the fallback mechanism.
- [PIG-2953](#): Added support for **which** utility on Windows
- [PIG-2956](#): Fixed issues with invalid cache specification for some streaming statement.
- [PIG-2954](#): Fixed test failures caused due to the dependency on **bash**.
- [PIG-2959](#): Fixes for `pig.cmd` to run on Windows.
- [PIG-2801](#): Fixes for grunt **sh** command.
- [PIG-2800](#): Fixed issues with `pig.additional.jars` path separator.
- [PIG-2799](#): Updated Pig streaming interface to run correctly on Windows without Cygwin support.
- [PIG-2798](#): Fixed issues with Pig streaming tests on Windows.
- [PIG-2797](#): Fixed the Pig tests to use **Util.generateURI** method.
- [PIG-2796](#): Fixed issues with invalid path names on HDFS caused because the Pig tests use local temporary paths.
- [PIG-2795](#): Added support to handle paths generated on Windows.
- [PIG-2794](#): Added utilities to facilitate testing on Windows platform.
- [PIG-2793](#): Improved Pig to work on Windows platform without Cygwin support.

2.2.3. Patch Information for Hive

The following list provides information on the **Apache JIRAs** for this release:

- [HIVE-3025](#): Fixed Hive archive command for Hadoop v 0.22 and 0.23.
- [HIVE-3448](#): Fixed failures for the skewjoin.q testcase on Windows.
- [HIVE-3441](#): Fixed failures caused due to the partition column strings in Windows file names.
- [HIVE-3436](#): Fixed the script_pipe.q failures on Windows.
- [HIVE-3483](#): Fixed issues with joins that use partitioned table on Windows.
- [HIVE-3317](#): Fixed TestDocToUnix unit tests on Windows.
- [HIVE-3320](#): Fixed test case failures caused by incorrect handling of CRLF line endings on Windows.
- [HIVE-3319](#): Fixed path related issues that caused the unit test failures for Windows.
- [HIVE-3327](#): Fixed failures caused while execution of the `/bin/cat` script files on Windows.
- [HIVE-3479](#): Fixed issues with negative unit tests.
- [HIVE-3494](#): Fixed JDBC test case failures on Windows.
- [HIVE-3480](#): Fixed file handle leaks in Symbolic and symlink related input formats.

2.2.4. Patch Information for HCatalog

The following list provides information on the **Apache JIRAs** for this release:

- [HCATALOG-512](#): Fixed HCatalog unit tests on Windows.
- [HCATALOG-514](#): Fixed HCatalog python scripts in the package build for Windows.

2.3. Known Issues

- One of the regression test for Oozie fails with MS SQL because `java.io.IOException` does not finish within the `600000timeout`.
- One of the regression test for Oozie fails with MS SQL because `java.io.UnsupportedEncodingException` does not finish within the `600000timeout`.
- Running MapReduce jobs using pipes is currently not supported on Windows.
- Non-default compression codecs that are based on zlib or snappy are currently not supported on Windows

- It is possible to encounter the following exception while starting the Hive command line interface (CLI).

```
FAILED: Error in metadata: javax.jdo.JDOFatalDataStoreException: DERBY SQL
error: SQLCODE: -1, SQLSTATE: XJ041, SQLERRMC:
    Failed to create database 'metastore_db', see the next exception for
    details.
    ::SQLSTATE: XBM0JDirectory C:\Hadoop\hive-0.9.0\bin\metastore_db already
    exists.
```

This typically happens when the user attempts to install and uninstall HDP repeatedly. In such cases, the directories for Hive might not get deleted properly on uninstall. You can use any of the following workaround:

- Manually delete the `metastore_db` directory (`C:\Hadoop\hive-0.9.0\bin\metastore_db`).
- Uninstall HDP, delete all the files in the Hadoop directory (`C:\Hadoop`), and install HDP again
- HDP creates some files in the HDFS directory (`C:\hdfs`) that are not deleted on uninstall. This issue is observed when `hadoop.tmp.dir` is not defined to point to the `C:\hadoop` location. There is no impact on the deployment of your cluster. However, it is recommended that you manually delete the files in the HDFS directory (`C:\hdfs`) after you uninstall HDP.