

ZooKeeper 3.0.1 Release Notes

by

Table of contents

| | |
|---|---|
| 1 Migration Instructions when Upgrading to version 3..... | 2 |
| 1.1 Migrating Client Code..... | 2 |
| 1.2 Migrating Server Data..... | 3 |
| 1.3 Migrating Server Configuration..... | 5 |
| 2 ZooKeeper 3.0.1 Changes..... | 5 |

These release notes include new developer and user facing incompatibilities, features, and major improvements.

- [Migration Instructions](#)
- [Changes](#)

1. Migration Instructions when Upgrading to version 3

You should only have to read this section if you are upgrading from a previous version of ZooKeeper, version 2 or earlier, to version 3, otw skip down to [changes](#)

A small number of changes in this release have resulted in non-backward compatible Zookeeper client user code and server instance data. The following instructions provide details on how to migrate code and data from version 2.2.1 to version 3.0.0.

Note: ZooKeeper increments the major version number (major.minor.fix) when backward incompatible changes are made to the source base. As part of the migration from SourceForge we changed the package structure (com.yahoo.zookeeper.* to org.apache.zookeeper.*) and felt it was a good time to incorporate some changes that we had been withholding. As a result the following will be required when migrating from 2.2.1 to 3.0.0 version of ZooKeeper.

- [Migrating Client Code](#)
- [Migrating Server Data](#)
- [Migrating Server Configuration](#)

1.1. Migrating Client Code

The underlying client-server protocol has changed in version 3.0.0 of ZooKeeper. As a result clients must be upgraded along with serving clusters to ensure proper operation of the system (old pre-3.0.0 clients are not guaranteed to operate against upgraded 3.0.0 servers and vice-versa).

1.1.1. Watch Management

In previous releases of ZooKeeper any watches registered by clients were lost if the client lost a connection to a ZooKeeper server. This meant that developers had to track watches they were interested in and reregister them if a session disconnect event was recieved. In this release the client library tracks watches that a client has registered and reregisters the watches when a connection is made to a new server. Applications that still manually reregister interest should continue working properly as long as they are able to handle unsolicited watches. For

example, an old application may register a watch for /foo and /goo, lose the connection, and reregister only /goo. As long as the application is able to receive a notification for /foo, (probably ignoring it) the application does not have to be changed. One caveat to the watch management: it is possible to miss an event for the creation and deletion of a znode if watching for creation and both the create and delete happens while the client is disconnected from ZooKeeper.

This release also allows clients to specify call specific watch functions. This gives the developer the ability to modularize logic in different watch functions rather than cramming everything in the watch function attached to the ZooKeeper handle. Call specific watch functions receive all session events for as long as they are active, but will only receive the watch callbacks for which they are registered.

1.1.2. Java API

1. The java package structure has changed from **com.yahoo.zookeeper*** to **org.apache.zookeeper***. This will probably effect all of your java code which makes use of ZooKeeper APIs (typically import statements)
2. A number of constants used in the client ZooKeeper API were re-specified using enums (rather than ints). See [ZOOKEEPER-7](#), [ZOOKEEPER-132](#) and [ZOOKEEPER-139](#) for full details
3. [ZOOKEEPER-18](#) removed KeeperStateChanged, use KeeperStateDisconnected instead

Also see [the current java API](#)

1.1.3. C API

1. A number of constants used in the client ZooKeeper API were renamed in order to reduce namespace collision, see [ZOOKEEPER-6](#) for full details

1.2. Migrating Server Data

The following issues resulted in changes to the on-disk data format (the snapshot and transaction log files contained within the ZK data directory) and require a migration utility to be run.

- [ZOOKEEPER-27 Unique DB identifiers for servers and clients](#)
- [ZOOKEEPER-32 CRCs for ZooKeeper data](#)
- [ZOOKEEPER-33 Better ACL management](#)
- [ZOOKEEPER-38 headers \(version+\) in log/snap files](#)

The following must be run once, and only once, when upgrading the ZooKeeper server instances to version 3.0.0.

Note:

The <dataLogDir> and <dataDir> directories referenced below are specified by the *dataLogDir* and *dataDir* specification in your ZooKeeper config file respectively. *dataLogDir* defaults to the value of *dataDir* if not specified explicitly in the ZooKeeper server config file (in which case provide the same directory for both parameters to the upgrade utility).

1. Shutdown the ZooKeeper server cluster.
2. Backup your <dataLogDir> and <dataDir> directories
3. Run upgrade using
 - `bin/zkServer.sh upgrade <dataLogDir> <dataDir>`
 or
 - `java -classpath pathnolog4j:pathtozookeeper.jar UpgradeMain <dataLogDir> <dataDir>`

where <dataLogDir> is the directory where all transaction logs (log.*) are stored.
<dataDir> is the directory where all the snapshots (snapshot.*) are stored.

4. Restart the cluster.

If you have any failure during the upgrade procedure keep reading to sanitize your database.

This is how upgrade works in ZooKeeper. This will help you troubleshoot in case you have problems while upgrading

1. Upgrade moves files from <dataLogDir> and <dataDir> to <dataLogDir>/version-1/ and <dataDir>/version-1 respectively (version-1 sub-directory is created by the upgrade utility).
2. Upgrade creates a new version sub-directory <dataDir>/version-2 and <dataLogDir>/version-2
3. Upgrade reads the old database from <dataDir>/version-1 and <dataLogDir>/version-1 into the memory and creates a new upgraded snapshot.
4. Upgrade writes the new database in <dataDir>/version-2.

Troubleshooting.

1. In case you start ZooKeeper 3.0 without upgrading from 2.0 on a 2.0 database - the servers will start up with an empty database. This is because the servers assume that <dataDir>/version-2 and <dataLogDir>/version-2 will have the database to start with.

Since this will be empty in case of no upgrade, the servers will start with an empty database. In such a case, shutdown the ZooKeeper servers, remove the version-2 directory (remember this will lead to loss of updates after you started 3.0.) and then start the upgrade procedure.

2. If the upgrade fails while trying to rename files into the version-1 directory, you should try and move all the files under <dataDir>/version-1 and <dataLogDir>/version-1 to <dataDir> and <dataLogDir> respectively. Then try upgrade again.
3. If you do not wish to run with ZooKeeper 3.0 and prefer to run with ZooKeeper 2.0 and have already upgraded - you can run ZooKeeper 2 with the <dataDir> and <dataLogDir> directories changed to <dataDir>/version-1 and <dataLogDir>/version-1. Remember that you will lose all the updates that you made after the upgrade.

1.3. Migrating Server Configuration

There is a significant change to the ZooKeeper server configuration file.

The default election algorithm, specified by the *electionAlg* configuration attribute, has changed from a default of *0* to a default of *3*. See [Cluster Options](#) section of the administrators guide, specifically the *electionAlg* and *server.X* properties.

You will either need to explicitly set *electionAlg* to it's previous default value of *0* or change your *server.X* options to include the leader election port.

2. ZooKeeper 3.0.1 Changes

| Issue | Notes |
|-------------------------------|---|
| ZOOKEEPER-211 | Not all Mock tests are working |
| ZOOKEEPER-223 | change default level in root logger to INFO. |
| ZOOKEEPER-212 | fix the snapshot to be asynchronous. |
| ZOOKEEPER-213 | fix programmer guide C api docs to be in sync with latest zookeeper.h |
| ZOOKEEPER-219 | fix events.poll timeout in watcher test to be longer. |
| ZOOKEEPER-217 | Fix errors in config to be thrown as Exceptions. |
| ZOOKEEPER-228 | fix apache header missing in DBTest. |
| ZOOKEEPER-218 | fix the error in the barrier example code. |
| ZOOKEEPER-206 | documentation tab should contain the version |

| | |
|-------------------------------|---|
| | number and other small site changes. |
| ZOOKEEPER-226 | fix exists calls that fail on server if node has null data. |
| ZOOKEEPER-204 | SetWatches needs to be the first message after auth messages to the server |
| ZOOKEEPER-208 | Zookeeper C client uses API that are not thread safe, causing crashes when multiple instances are active. |
| ZOOKEEPER-227 | gcc warning from recordio.h |
| ZOOKEEPER-232 | fix apache licence header in TestableZookeeper |

Table 1: Changes Since ZooKeeper 3.0.0

Version 2.2.1 code, documentation, binaries, etc... are still accessible on [SourceForge](#)

| Issue | Notes |
|-------------------------------|--|
| ZOOKEEPER-43 | Server side of auto reset watches. |
| ZOOKEEPER-132 | Create Enum to replace CreateFlag in ZooKeeper.create method |
| ZOOKEEPER-139 | Create Enums for WatcherEvent's KeeperState and EventType |
| ZOOKEEPER-18 | keeper state inconsistency |
| ZOOKEEPER-38 | headers in log/snap files |
| ZOOKEEPER-8 | Stat enchaned to include num of children and size |
| ZOOKEEPER-6 | List of problem identifiers in zookeeper.h |
| ZOOKEEPER-7 | Use enums rather than ints for types and state |
| ZOOKEEPER-27 | Unique DB identifiers for servers and clients |
| ZOOKEEPER-32 | CRCs for ZooKeeper data |
| ZOOKEEPER-33 | Better ACL management |
| ZOOKEEPER-203 | fix datadir typo in releasenotes |
| ZOOKEEPER-145 | write detailed release notes for users migrating from 2.x to 3.0 |

| | |
|-------------------------------|---|
| ZOOKEEPER-23 | Auto reset of watches on reconnect |
| ZOOKEEPER-191 | forrest docs for upgrade. |
| ZOOKEEPER-201 | validate magic number when reading snapshot and transaction logs |
| ZOOKEEPER-200 | the magic number for snapshot and log must be different |
| ZOOKEEPER-199 | fix log messages in persistence code |
| ZOOKEEPER-197 | create checksums for snapshots |
| ZOOKEEPER-198 | apache license header missing from FollowerSyncRequest.java |
| ZOOKEEPER-5 | Upgrade Feature in Zookeeper server. |
| ZOOKEEPER-194 | Fix terminology in zookeeperAdmin.xml |
| ZOOKEEPER-151 | Document change to server configuration |
| ZOOKEEPER-193 | update java example doc to compile with latest zookeeper |
| ZOOKEEPER-187 | CreateMode api docs missing |
| ZOOKEEPER-186 | add new "releasenotes.xml" to forrest documentation |
| ZOOKEEPER-190 | Reorg links to docs and navs to docs into related sections |
| ZOOKEEPER-189 | forrest build not validated xml of input documents |
| ZOOKEEPER-188 | Check that election port is present for all servers |
| ZOOKEEPER-185 | Improved version of FLETest |
| ZOOKEEPER-184 | tests: An explicit include directive is needed for the usage of memcpy functions |
| ZOOKEEPER-183 | Array subscript is above array bounds in od_completion, src/cli.c. |
| ZOOKEEPER-182 | zookeeper_init accepts empty host-port string and returns valid pointer to zhandle_t. |
| ZOOKEEPER-17 | zookeeper_init doc needs clarification |

| | |
|-------------------------------|--|
| ZOOKEEPER-181 | Some Source Forge Documents did not get moved over: javaExample, zookeeperTutorial, zookeeperInternals |
| ZOOKEEPER-180 | Placeholder sections needed in document for new topics that the umbrella jira discusses |
| ZOOKEEPER-179 | Programmer's Guide "Basic Operations" section is missing content |
| ZOOKEEPER-178 | FLE test. |
| ZOOKEEPER-159 | Cover two corner cases of leader election |
| ZOOKEEPER-156 | update programmer guide with acl details from old wiki page |
| ZOOKEEPER-154 | reliability graph diagram in overview doc needs context |
| ZOOKEEPER-157 | Peer can't find existing leader |
| ZOOKEEPER-155 | improve "the zookeeper project" section of overview doc |
| ZOOKEEPER-140 | Deadlock in QuorumCnxManager |
| ZOOKEEPER-147 | This is version of the documents with most of the [tbd...] scrubbed out |
| ZOOKEEPER-150 | zookeeper build broken |
| ZOOKEEPER-136 | sync causes hang in all followers of quorum. |
| ZOOKEEPER-134 | findbugs cleanup |
| ZOOKEEPER-133 | hudson tests failing intermittently |
| ZOOKEEPER-144 | add toString support for watcher event, and enums for event type/state |
| ZOOKEEPER-21 | Improve zk ctor/watcher |
| ZOOKEEPER-142 | Provide Javadoc as to the maximum size of the data byte array that may be stored within a znode |
| ZOOKEEPER-93 | Create Documentation for Zookeeper |
| ZOOKEEPER-117 | threading issues in Leader election |

| | |
|-------------------------------|--|
| ZOOKEEPER-137 | client watcher objects can lose events |
| ZOOKEEPER-131 | Old leader election can elect a dead leader over and over again |
| ZOOKEEPER-130 | update build.xml to support apache release process |
| ZOOKEEPER-118 | findbugs flagged switch statement in followerrequestprocessor.run |
| ZOOKEEPER-115 | Potential NPE in QuorumCnxManager |
| ZOOKEEPER-114 | cleanup ugly event messages in zookeeper client |
| ZOOKEEPER-112 | src/java/main ZooKeeper.java has test code embedded into it. |
| ZOOKEEPER-39 | Use Watcher objects rather than boolean on read operations. |
| ZOOKEEPER-97 | supports optional output directory in code generator. |
| ZOOKEEPER-101 | Integrate ZooKeeper with "violations" feature on hudson |
| ZOOKEEPER-105 | Catch Zookeeper exceptions and print on the stderr. |
| ZOOKEEPER-42 | Change Leader Election to fast tcp. |
| ZOOKEEPER-48 | auth_id now handled correctly when no auth ids present |
| ZOOKEEPER-44 | Create sequence flag children with prefixes of 0's so that they can be lexicographically sorted. |
| ZOOKEEPER-108 | Fix sync operation reordering on a Quorum. |
| ZOOKEEPER-25 | Fuse module for Zookeeper. |
| ZOOKEEPER-58 | Race condition on ClientCnxn.java |
| ZOOKEEPER-56 | Add clover support to build.xml. |
| ZOOKEEPER-75 | register the ZooKeeper mailing lists with nabble.com |
| ZOOKEEPER-54 | remove sleeps in the tests. |

| | |
|-------------------------------|---|
| ZOOKEEPER-55 | build.xml failes to retrieve a release number from SVN and the ant target "dist" fails |
| ZOOKEEPER-89 | invoke WhenOwnerListener.whenNotOwner when the ZK connection fails |
| ZOOKEEPER-90 | invoke WhenOwnerListener.whenNotOwner when the ZK session expires and the znode is the leader |
| ZOOKEEPER-82 | Make the ZooKeeperServer more DI friendly. |
| ZOOKEEPER-110 | Build script relies on svnant, which is not compatible with subversion 1.5 working copies |
| ZOOKEEPER-111 | Significant cleanup of existing tests. |
| ZOOKEEPER-122 | Fix NPE in jute's Utils.toCSVString. |
| ZOOKEEPER-123 | Fix the wrong class is specified for the logger. |
| ZOOKEEPER-2 | Fix synchronization issues in QuorumPeer and FastLeader election. |
| ZOOKEEPER-125 | Remove unwanted class declaration in FastLeaderElection. |
| ZOOKEEPER-61 | Address in client/server test cases. |
| ZOOKEEPER-75 | cleanup the library directory |
| ZOOKEEPER-109 | cleanup of NPE and Resource issue nits found by static analysis |
| ZOOKEEPER-76 | Commit 677109 removed the cobertura library, but not the build targets. |
| ZOOKEEPER-63 | Race condition in client close |
| ZOOKEEPER-70 | Add skeleton forrest doc structure for ZooKeeper |
| ZOOKEEPER-79 | Document jacob's leader election on the wiki recipes page |
| ZOOKEEPER-73 | Move ZK wiki from SourceForge to Apache |
| ZOOKEEPER-72 | Initial creation/setup of ZooKeeper ASF site. |
| ZOOKEEPER-71 | Determine what to do re ZooKeeper Changelog |

| | |
|-------------------------------|---|
| ZOOKEEPER-68 | parseACLs in ZooKeeper.java fails to parse elements of ACL, should be lastIndexOf rather than IndexOf |
| ZOOKEEPER-130 | update build.xml to support apache release process. |
| ZOOKEEPER-131 | Fix Old leader election can elect a dead leader over and over again. |
| ZOOKEEPER-137 | client watcher objects can lose events |
| ZOOKEEPER-117 | threading issues in Leader election |
| ZOOKEEPER-128 | test coverage on async client operations needs to be improved |
| ZOOKEEPER-127 | Use of non-standard election ports in config breaks services |
| ZOOKEEPER-53 | tests failing on solaris. |
| ZOOKEEPER-172 | FLE Test |
| ZOOKEEPER-41 | Sample startup script |
| ZOOKEEPER-33 | Better ACL management |
| ZOOKEEPER-49 | SetACL does not work |
| ZOOKEEPER-20 | Child watches are not triggered when the node is deleted |
| ZOOKEEPER-15 | handle failure better in build.xml:test |
| ZOOKEEPER-11 | ArrayList is used instead of List |
| ZOOKEEPER-45 | Restructure the SVN repository after initial import |
| ZOOKEEPER-1 | Initial ZooKeeper code contribution from Yahoo! |

Table 2: Changes Since ZooKeeper 2.2.1