

by

Table of contents

1 Getting Started: Coordinating Distributed Applications with ZooKeeper.....	2
1.1 Installing and Running ZooKeeper in Single Server Mode.....	2
1.2 Connecting to ZooKeeper.....	3
1.3 Programming to ZooKeeper.....	3
1.4 Running Replicated ZooKeeper.....	3
1.5 Other Optimizations.....	4

1. Getting Started: Coordinating Distributed Applications with ZooKeeper

This document contains information to get you started quickly with Zookeeper. It is aimed primarily at developers hoping to try it out, and contains simple installation instructions for a single ZooKeeper server, a few commands to verify that it is running, and a simple programming example. Finally, as a convenience, there are a few sections regarding more complicated installations, for example running replicated deployments, and optimizing the transaction log. However for the complete instructions for commercial deployments, please refer to the [Zookeeper Administrator's Guide](#).

1.1. Installing and Running ZooKeeper in Single Server Mode

Setting up a ZooKeeper server in standalone mode is straightforward. The server is contained in a single JAR file, so installation consists of copying a JAR file and creating a configuration.

Note:

Zookeeper requires Java 1.5 or more recent.

Once you have downloaded the ZooKeeper source, cd to the root of your ZooKeeper source, and run "ant jar". For example:

```
$ cd ~/dev/zookeeper
$ ~/dev/zookeeper/: ant jar
```

This should generate a JAR file called zookeeper.jar. To start Zookeeper, compile and run zookeeper.jar. [tbd, some more instruction here. Perhaps a command line? Are these two steps or one?]

To start ZooKeeper you need a configuration file. Here is a sample file:

```
tickTime=2000
dataDir=/var/zookeeper/
clientPort=2181
```

This file can be called anything, but for the sake of this discussion, call it **zoo.cfg**. Here are the meanings for each of the fields:

tickTime

the basic time unit in milliseconds used by ZooKeeper. It is used to do heartbeats and the minimum session timeout will be twice the tickTime.

dataDir

the location to store the in-memory database snapshots and, unless specified otherwise, the transaction log of updates to the database.

clientPort

the port to listen for client connections

Now that you created the configuration file, you can start ZooKeeper:

```
java -cp zookeeper-dev.jar:java/lib/log4j-1.2.15.jar:conf  
org.apache.zookeeper.server.quorum.QuorumPeerMain zoo.cfg
```

ZooKeeper logs messages using log4j -- more detail available in the [Logging](#) section of the Programmer's Guide.[tbd: real reference needed] You will see log messages coming to the console and/or a log file depending on the log4j configuration.

The steps outlined here run ZooKeeper in standalone mode. There is no replication, so if Zookeeper process fails, the service will go down. This is fine for most development situations, but to run Zookeeper in replicated mode, please see [Running Replicated Zookeeper](#).

1.2. Connecting to ZooKeeper

Once ZooKeeper is running, you have several option for connection to it:

- **Java:** Use `java -cp zookeeper.jar:java/lib/log4j-1.2.15.jar:conf org.apache.zookeeper.ZooKeeperMain 127.0.0.1:2181`

This lets you perform simple, file-like operations.

- **C:** compile `cli_mt` (multi-threaded) or `cli_st` (single-threaded) by running `_make cli_mt_` or `_make cli_st_` in the `c` subdirectory in the ZooKeeper sources.

You can run the program using `LD_LIBRARY_PATH=. cli_mt 127.0.0.1:2181` or `LD_LIBRARY_PATH=. cli_st 127.0.0.1:2181`. This will give you a simple shell to execute file system like operations on ZooKeeper.

1.3. Programming to ZooKeeper

ZooKeeper has a Java bindings and C bindings. They are functionally equivalent. The C bindings exist in two variants: single threaded and multi-threaded. These differ only in how the messaging loop is done. For more information, see the [Programming Examples in the Zookeeper Programmer's Guide](#) for sample code using of the different APIs.

1.4. Running Replicated ZooKeeper

Running ZooKeeper in standalone mode is convenient for evaluation, some development, and testing. But in production, you should run ZooKeeper in replicated mode. A replicated group of servers in the same application is called a *quorum*, and in replicated mode, all servers in the quorum have copies of the same configuration file. The file is similar to the one used in standalone mode, but with a few differences. Here is an example:

```
tickTime=2000
dataDir=/var/zookeeper/
clientPort=2181
initLimit=5
syncLimit=2
server.1=zoo1:2888 server.2=zoo2:2888
server.3=zoo3:2888
```

The new entry, **initLimit** is timeouts ZooKeeper uses to limit the length of time the Zookeeper servers in quorum have to connect to a leader. The entry **syncLimit** limits how far out of date a server can be from a leader.

With both of these timeouts, you specify the unit of time using **tickTime**. In this example, the timeout for initLimit is 5 ticks at 2000 milliseconds a tick, or 10 seconds.

The entries of the form *server.X* list the servers that make up the ZooKeeper service. When the server starts up, it knows which server it is by looking for the file **myid** in the data directory. That file contains the server number, in ASCII.

Finally, note the "2888" port numbers after each server name. These are the "electionPort" numbers of the servers (as opposed to clientPorts), that is ports for [tbd].

Note:

If you want to test multiple servers on a single machine, define the electionPort for each server in that server's config file, using the line `electionPort=xxxx` as means of avoiding clashes.

1.5. Other Optimizations

There are a couple of other configuration parameters that can greatly increase performance:

- To get low latencies on updates it is important to have a dedicated transaction log directory. By default transaction logs are put in the same directory as the data snapshots and **myid** file. The `dataLogDir` parameter indicates a different directory to use for the transaction logs.
- [tbd: what is the other config param?]