

CLI MiniCluster

Table of contents

| | |
|--------------------------------|---|
| 1 Purpose..... | 2 |
| 2 Hadoop Tarball..... | 2 |
| 3 Running the MiniCluster..... | 2 |

1 Purpose

Using the CLI MiniCluster, users can simply start and stop a single-node Hadoop cluster with a single command, and without the need to set any environment variables or manage configuration files. The CLI MiniCluster starts both a MapReduce and HDFS clusters. This is useful for cases where users want to quickly experiment with a real Hadoop cluster or test non-Java programs that rely on significant Hadoop functionality.

2 Hadoop Tarball

To get a Hadoop distribution, download a recent [stable release](#) from one of the Apache Download Mirrors. Unpack the downloaded Hadoop distribution. In the distribution, edit the file `conf/hadoop-env.sh` to define at least `JAVA_HOME` to be the root of your Java installation.

3 Running the MiniCluster

From inside the root directory of the extracted tarball, you can start the CLI MiniCluster using the following command:

```
$ bin/hadoop jar hadoop-test-*.jar minicluster -jtport JT_PORT
-nnport NN_PORT
```

In the example command above, `JT_PORT` and `NN_PORT` should be replaced by the user's choice of these port numbers. If not specified, random free ports will be used.

There are a number of command line arguments that the users can use to control which services to start, and to pass other configuration properties. The available command line arguments:

| Argument | Description |
|--|--|
| <code>-D <property=value></code> | Options to pass into configuration object |
| <code>-datanodes <arg></code> | How many datanodes to start (default 1) |
| <code>-format</code> | Format the DFS (default false) |
| <code>-help</code> | Prints option help. |
| <code>-jhsport <arg></code> | JobHistoryServer port (default 0--we choose) |
| <code>-namenode <arg></code> | URL of the namenode (default is either the DFS cluster or a temporary dir) |
| <code>-nnport <arg></code> | NameNode port (default 0--we choose) |
| <code>-nodemanagers <arg></code> | How many nodemanagers to start (default 1) |

| Argument | Description |
|----------------------|---|
| -nodfs | Don't start a mini DFS cluster |
| -nomr | Don't start a mini MR cluster |
| -rmport <arg> | ResourceManager port (default 0--we choose) |
| -writeConfig <path> | Save configuration to this XML file. |
| -writeDetails <path> | Write basic information to this JSON file. |

To display this full list of available arguments, the user can pass the `-help` argument to the above command.