Typically one machine in the cluster is designated as the `NameNode` and another machine the as `JobTracker`, exclusively. These are the *masters*. The rest of the machines in the cluster act as both `DataNode` *and* `TaskTracker`. These are the *slaves*.

The root of the distribution is referred to as `HADOOP_HOME`. All machines in the cluster usually have the same `HADOOP_HOME` path.

Configuration

The following sections describe how to configure a Hadoop cluster.

Configuration Files

Hadoop configuration is driven by two types of important configuration files:

1. Read-only default configuration - src/core/core-default.xml, src/hdfs/hdfs-default.xml and src/mapred/mapred-default.xml.
2. Site-specific configuration - *conf/core-site.xml*, *conf/hdfs-site.xml* and *conf/mapred-site.xml*.

To learn more about how the Hadoop framework is controlled by these configuration files, look here.

Additionally, you can control the Hadoop scripts found in the `bin/` directory of the distribution, by setting site-specific values via the `conf/hadoop-env.sh`.

Site Configuration

To configure the Hadoop cluster you will need to configure the *environment* in which the Hadoop daemons execute as well as the *configuration parameters* for the Hadoop daemons.

The Hadoop daemons are `NameNode/DataNode` and `JobTracker/TaskTracker`.

Configuring the Environment of the Hadoop Daemons

Administrators should use the `conf/hadoop-env.sh` script to do site-specific customization of the Hadoop daemons' process environment.

At the very least you should specify the `JAVA_HOME` so that it is correctly defined on each remote node.

Administrators can configure individual daemons using the configuration options `HADOOP_*_OPTS`. Various options available are shown below in the table.

Daemon	Configure Options
NameNode	HADOOP_NAMENODE_OPTS
DataNode	HADOOP_DATANODE_OPTS
SecondaryNamenode	HADOOP_SECONDARYNAMENODE_OPTS
JobTracker	HADOOP_JOBTRACKER_OPTS
TaskTracker	HADOOP_TASKTRACKER_OPTS

For example, To configure Namenode to use parallelGC, the following statement should be added in `hadoop-env.sh`:

```
export HADOOP_NAMENODE_OPTS="-XX:+UseParallelGC
${HADOOP_NAMENODE_OPTS}"
```

Other useful configuration parameters that you can customize include:

- `HADOOP_LOG_DIR` - The directory where the daemons' log files are stored. They are automatically created if they don't exist.

- `HADOOP_HEAPSIZE` - The maximum amount of heapsize to use, in MB e.g. `1000MB`. This is used to configure the heap size for the hadoop daemon. By default, the value is `1000MB`.

Configuring the Hadoop Daemons

This section deals with important parameters to be specified in the following:
`conf/core-site.xml`:

Parameter	Value	Notes
fs.default.name	URI of `NameNode`.	*hdfs://hostname/*

`conf/hdfs-site.xml`:

Parameter	Value	Notes
dfs.name.dir	Path on the local filesystem where the `NameNode` stores the namespace and transactions logs persistently.	If this is a comma-delimited list of directories then the name table is replicated in all of the directories, for redundancy.
dfs.data.dir	Comma separated list of paths on the local filesystem of a `DataNode` where it should store its blocks.	If this is a comma-delimited list of directories, then data will be stored in all named directories, typically on different devices.

`conf/mapred-site.xml`:

Parameter	Value	Notes
mapred.job.tracker	Host or IP and port of `JobTracker`.	*host:port* pair.
mapred.system.dir	Path on the HDFS where where the MapReduce framework stores system files e.g. `/hadoop/mapred/system/`.	This is in the default filesystem (HDFS) and must be accessible from both the server and client machines.
mapred.local.dir	Comma-separated list of paths on the local filesystem where temporary MapReduce data is written.	Multiple paths help spread disk i/o.
mapred.tasktracker. {map|reduce}.tasks.maximum	The maximum number of MapReduce tasks, which are run simultaneously on a given `TaskTracker`, individually.	Defaults to 2 (2 maps and 2 reduces), but vary it depending on your hardware.
dfs.hosts/dfs.hosts.exclude	List of permitted/excluded DataNodes.	If necessary, use these files to control the list of allowable datanodes.
mapred.hosts/mapred.hosts.exclude	List of permitted/excluded TaskTrackers.	If necessary, use these files to control the list of allowable TaskTrackers.
mapred.queue.names	Comma separated list of queues to which jobs can be submitted.	The MapReduce system always supports atleast one queue with the name as *default*. Hence, this parameter's value should always contain the string *default*. Some job schedulers supported in Hadoop, like the Capacity Scheduler, support multiple queues. If such a scheduler is being used, the list of configured queue names must be specified here. Once queues are defined, users can submit jobs to a queue using the property name

		mapred.job.queue.name in the job configuration. There could be a separate configuration file for configuring properties of these queues that is managed by the scheduler. Refer to the documentation of the scheduler for information on the same.
mapred.acls.enabled	Boolean, specifying whether checks for queue ACLs and job ACLs are to be done for authorizing users for doing queue operations and job operations.	If *true*, queue ACLs are checked while submitting and administering jobs and job ACLs are checked for authorizing view and modification of jobs. Queue ACLs are specified using the configuration parameters of the form *mapred.queue.queue-name.acl-name*, defined below under mapred-queue-acls.xml. Job ACLs are described at Job Authorization

`conf/mapred-queue-acls.xml`

Parameter	Value	Notes
mapred.queue.*queue-name*.acl-submit-job	List of users and groups that can submit jobs to the specified *queue-name*.	The list of users and groups are both comma separated list of names. The two lists are separated by a blank. Example: *user1,user2 group1,group2*. If you wish to define only a list of groups, provide a blank at the beginning of the value.
mapred.queue.*queue-name*.acl-administer-jobs	List of users and groups that can view job details, change the priority or kill jobs that have been submitted to the specified *queue-name*.	The list of users and groups are both comma separated list of names. The two lists are separated by a blank. Example: *user1,user2 group1,group2*. If you wish to define only a list of groups, provide a blank at the beginning of the value. Note that the owner of a job can always change the priority or kill his/her own job, irrespective of the ACLs.

Typically all the above parameters are marked as final to ensure that they cannot be overriden by user-applications.

Real-World Cluster Configurations

This section lists some non-default configuration parameters which have been used to run the *sort* benchmark on very large clusters.

- Some non-default configuration values used to run sort900, that is 9TB of data sorted on a cluster with 900 nodes:

Configuration File	Parameter	Value	Notes
conf/hdfs-site.xml	dfs.block.size	134217728	HDFS blocksize of 128MB for large file-systems.
conf/hdfs-site.xml	dfs.namenode.handler.count	40	More NameNode server threads to handle RPCs from large number of DataNodes.
conf/mapred-site.xml	mapred.reduce.parallel.copies	20	Higher number of parallel copies run by reduces to fetch outputs from very large number of maps.
conf/mapred-site.xml	mapred.map.child.java.opts	-Xmx512M	Larger heap-size for child jvms of maps.