

# Map Reductions in Hadoop

This tutorial covers the basic concepts and code of Hadoop.

The Hadoop system makes it easy to parallelize a particular form of computation across a large data set, just like Google's MapReduce. It automates the process of splitting up and coordinating a computation and its data to run on thousands of machines at once. Although the implementation of Hadoop is complex, writing code that uses Hadoop is actually not very hard (hence, a successful abstraction).

This tutorial walks through a few code examples to illustrate the key features. Both examples make use of literature text sources - Shakespeare's complete works and Mark Twain's Huckleberry Finn. These and many others works are available at [Project Gutenberg](#). However, we also encourage you to find and use alternate text sources that are interesting to you, even if they're not available at Project Gutenberg.

## Hadoop Concepts

To use Hadoop, you write two classes -- a Mapper and a Reducer. The Mapper class contains a map function, which is called once for each input and outputs any number of intermediate <key, value> pairs. What code you put in the map function depends on the problem you are trying to solve. Let's start with a short example.

Suppose the goal is to create an "index" of a body of text -- we are given a text file, and we want to output a list of words annotated with the line-number at which each word appears.

For that problem, an appropriate Map strategy is: for each word in the input, output the pair <word, line-number> For example, suppose we have this five-line High school football coach quote as our input data set:

```
We are not what  
we want to be,  
but at least  
we are not what  
we used to be.
```

Running the Map code that for each word, outputs a pair <word, line-number>, yielding the set of pairs...

```
<we, 1>  
<are, 1>  
<not, 1>  
<what, 1>  
<we, 2>  
<want, 2>  
<to, 2>
```

<be, 2>  
<but, 3>  
<at, 3>  
<least, 3>  
<we, 4>  
<are, 4>  
<not, 4>  
<what, 4>  
<we, 5>  
<used, 5>  
<to, 5>  
<be, 5>

For now we can think of the <key, value> pairs as a nice linear list, but in reality, the Hadoop process runs in parallel on many machines. Each process has a little part of the overall Map input (called a map shard), and maintains its own local cache of the Map output. (For a description of how it really works, see Hadoop MapReduce or the Google White Paper.)

After the Map phase produces the intermediate <key, value> pairs they are efficiently and automatically grouped by key by the Hadoop system in preparation for the Reduce phase (this grouping is known as the Shuffle phase of a map-reduce). For the above example, that means all the "we" pairs are grouped together, all the "are" pairs are grouped together like this, showing each group as a line...

<we, 1> <we, 2> <we, 4> <we, 5>  
<are, 1> <are, 4>  
<not, 1> <not, 4>  
<what, 1> <what, 4>  
<want, 2>  
<to, 2> <to, 5>  
<be, 2> <be 5>  
<but, 3>  
<at, 3>  
<least, 3>  
<used, 5>

The Reducer class contains a reduce function, which is then called once for each key -- one reduce call for "we", one for "are", and so on. Each reduce looks at all the values for that key and outputs a "summary" value for that key in the final output. So in the above example, the reduce is called once for the "we" key, and passed the values the mapper output, 1, 4, 2, and 5 (the values going into reduce are not in any particular order). Suppose reduce computes a summary value string made of the line numbers sorted into increasing order, then the output of the Reduce phase on the above pairs will produce the pairs shown below. The Reduce phase also sorts the output <key,value> pairs into increasing order by key:

<are, 1 4>  
<at, 3>  
<be, 2 5>  
<but, 3>  
<least, 3>  
<not, 1 4>  
<to, 2 5>  
<we, 1 2 4 5>  
<what, 1 4>  
<want, 2>  
<used, 5>

Like Map, Reduce is also run in parallel on a group of machines. Each machine is assigned a subset of the keys to work on (known as a reduce shard), and outputs its results into a separate file. By default these are named "part-#####".

## Line Indexer Example

As a first Hadoop code example we will look at a simple "line indexer". Given an input text, offset indexer uses Hadoop to produce an index of all the words in the text. For each word, the index has a list of all the locations where the word appears and a text excerpt of each line where the word appears. Running the line indexer on the complete works of Shakespeare yields the following entry for the word "cipher".

```
38624    To cipher what is writ in learned books,  
12046    To cipher me how fondly I did dote;  
34739    Mine were the very cipher of a function,  
16844    MOTH   To prove you a cipher.  
66001    ORLANDO Which I take to be either a fool or a cipher.
```

The Hadoop code below for the line indexer is actually pretty short. The Map code extracts one word at a time from the input, and the Reduce code combines all the data for one word.

## Line Indexer Map

A Java Mapper class is defined in terms of its input and intermediate <key, value> pairs. To declare one, simply subclass from MapReduceBase and implement the Mapper interface. The Mapper interface provides a single method: *public void map(WritableComparable key, Writable value, OutputCollector output, Reporter reporter)*. Note: these inner classes probably need to be declared "static". If you get an error saying *ClassName.<init>()* is not defined, try declaring your class static. The map function takes four parameters which in this example correspond to:

- WritableComparable key - the byte-offset
- Writable value - the line from the file
- OutputCollector - output - this has the .collect method to output a <key, value> pair

- Reporter reporter - you can ignore this for now

The Hadoop system divides the (large) input data set into logical "records" and then calls `map()` once for each record. How much data constitutes a record depends on the input data type; For text files, a record is a single line of text. The main method is responsible for setting output key and value types.

Since in this example we want to output `<word, offset>` pairs, the types will both be `Text` (a basic string wrapper, with UTF8 support). It is necessary to wrap the more basic types because all input and output types for Hadoop must implement `WritableComparable`, which handles the writing and reading from disk.

## Line Indexer Map

For the line indexer problem, the map code takes in a line of text and for each word in the line outputs a string key/value pair `<word, offset:line>`.

The Map code below accomplishes that by...

- Parsing each word out of *value*. For the parsing, the code delegates to a utility *StringTokenizer* object that implements *hasMoreTokens()* and *nextToken()* to iterate through the tokens.
- For each word, getting the location from *key*.
- Calling *output.collect(word, value)* to output a `<key, value>` pair for each word.

```
public static class LineIndexerMapper extends MapReduceBase implements
Mapper {
    private final static Text word = new Text();
    private final static Text summary = new Text();
    public void map(WritableComparable key, Writable val,
        OutputCollector output, Reporter reporter)
        throws IOException {
        String line = val.toString();
        summary.set(key.toString() + ":" + line);
        StringTokenizer itr = new StringTokenizer(line.toLowerCase());
        while(itr.hasMoreTokens()) {
            word.set(itr.nextToken());
            output.collect(word, summary);
        }
    }
}
```

When run on many machines, each mapper gets part of the input -- so for example with 100 Gigabytes of data on 200 mappers, each mapper would get roughly its own 500 Megabytes of data to go through. On a single mapper, `map()` is called going through the data in its natural order, from start to finish.

The Map phase outputs <key, value> pairs, but what data makes up the key and value is totally up to the Mapper code. In this case, the Mapper uses each word as a key, so the reduction below ends up with pairs grouped by word. We could instead have chosen to use the line-length as the key, in which case the data in the reduce phase would have been grouped by line length. In fact, the map() code is not required to call output.collect() at all. It may have its own logic to prune out data simply by omitting collect. Pruning things in the Mapper is efficient, since it is highly parallel, and already has the data in memory. By shrinking its output, we shrink the expense of organizing and moving the data in preparation for the Reduce phase.

## Line Indexer Reduce

Defining a Reducer is just as easy. Simply subclass MapReduceBase and implement the Reducer interface: *public void reduce(WritableComparable key, Iterator values, OutputCollector output, Reporter reporter)*. The *reduce()* method is called once for each key; the values parameter contains all of the values for that key. The Reduce code looks at all the values and then outputs a single "summary" value.

Given all the values for the key, the Reduce code typically iterates over all the values and either concatenates the values together in some way to make a large summary object, or combines and reduces the values in some way to yield a short summary value.

The *reduce()* method produces its final value in the same manner as *map()* did, by calling *output.collect(key, summary)*. In this way, the Reduce specifies the final output value for the (possibly new) key. It is important to note that when running over text files, the input key is the byte-offset within the file. If the key is propagated to the output, even for an identity map/reduce, the file will be filled with the offset values. Not only does this use up a lot of space, but successive operations on this file will have to eliminate them. For text files, make sure you don't output the key unless you need it (be careful with the IdentityMapper and IdentityReducer).

## Line Indexer Reduce Code

The line indexer Reducer takes in all the <word, offset> key/value pairs output by the Mapper for a single word. For example, for the word "cipher", the pairs look like: <cipher, 38624:To cipher what is writ in learned books>, <cipher, 12046:To cipher me how fondly I did dote;>, <cipher, ... >.

Given all those <key, value> pairs, the reduce outputs a single value string. For the line indexer problem, the strategy is simply to concatenate all the values together to make a single large string, using "^" to separate the values. The choice of "^" is arbitrary -- later code can split on the "^" to recover the separate values. So for the key "cipher" the output value string will look like "38624:To cipher what is writ in learned book^12046:To cipher me how fondly I did dote;^34739:Mine were the very cipher of a function,^ ...". ]

To do this, the Reducer code simply iterates over values to get all the value strings, and concatenates them together into our output String.

```
public static class LineIndexerReducer extends MapReduceBase implements
Reducer {
    public void reduce(WritableComparable key, Iterator values,
        OutputCollector output, Reporter reporter)
        throws IOException {
        boolean first = true;
        StringBuilder toReturn = new StringBuilder();
        while (values.hasNext()) {
            if (!first)
                toReturn.append('^');
            first = false;
            toReturn.append(values.next().toString());
        }
        output.collect(key, new Text(toReturn.toString()));
    }
}
```

## Line Indexer Main Program

Given the Mapper and Reducer code, the short *main()* below starts the Map-Reduction running. The Hadoop system picks up a bunch of values from the command line on its own, and then the *main()* also specifies a few key parameters of the problem in the JobConf object, such as what Map and Reduce classes to use and the format of the input and output files. Other parameters, i.e. the number of machines to use, are optional and the system will determine good values for them if not specified.

```
public static void main(String[] args) throws IOException {
    JobConf conf = new JobConf(LineIndexer.class);
    conf.setJobName("LineIndexer");
    // The keys are words (strings):
    conf.setOutputKeyClass(Text.class);
    // The values are offsets+line (strings):
    conf.setOutputValueClass(Text.class);

    conf.setMapperClass(LineIndexer.LineIndexerMapper.class);
    conf.setReducerClass(LineIndexer.LineIndexerReducer.class);
    if (args.length < 2) {
        System.out.println("Usage: LineIndexer <input path> <output
path>");
        System.exit(0);
    }
    conf.setInputPath(new Path(args[0]));
    conf.setOutputPath(new Path(args[1]));
    JobClient.runJob(conf);
}
```

## Running A Map-Reduction

To run a Hadoop job, simply SSH into any of the JobTracker nodes on the cluster. To run the job, it is first necessary to copy the input data files onto the distributed file system. If the data files are in the localInput/ directory, this is accomplished by executing:

```
./bin/hadoop dfs -put localInput dfsInput
```

The files will then be copied onto the dfs into the directory dfsInput. It is important to copy files into a well named directory that is unique. These files can be viewed with

```
./bin/hadoop dfs -ls dir
```

where dir is the name of the directory to be viewed.

You can also use

```
./bin/hadoop dfs -lsr dir
```

to recursively view the directories. Note that all "relative" paths given will be put in the /users/\$USER/[dir] directory.

Make sure that the dfsOutput directory does not already exist, as you will be presented with an error, and your job will not run (This prevents the accidental overwriting of data, but can be overridden).

Now that the data is available to all of the worker machines, the job can be executed from a local jar file:

```
./bin/hadoop jar OffsetIndexer.jar dfsInput dfsOutput
```

The job should be run across the worker machines, copying input and intermediate data as needed. The output of the reduce stage will be left in the dfsOutput directory. To copy these files to your local machine in the directory localOutput, execute:

```
./bin/hadoop dfs -get dfsOutput localOutput
```

## Running A Map-Reduction Locally

During testing, you may want to run your Map-Reduces locally so as not to adversely affect the compute clusters. This is easily accomplished by adding a line to the main method:

```
conf.set("mapred.job.tracker", "local");
```

## Seeing Job Progress

When you submit your job to run a line will be printed saying:

```
Running job: job_12345
```

where 'job\_12345' will correspond to whatever name your job has been given. Further status information will be printed in that terminal as the job progresses. However, it is also possible to monitor a job given its name from any node in the cluster. This is done by the command:

```
./bin/hadoop/ job -status job_12345
```

A small amount of status information will be displayed, along with a link to a tracking URL (eg, <http://jobtrackerhostname:50050/>). This page will be a job-specific status page, and provide links to main status pages for other jobs and the Hadoop cluster itself.

## Line Indexer Client

After the Hadoop job finishes and the output is copied to a local machine, it can be analyzed. The line indexer client reads in the index files and presents a simple command line interface - type a word to pull up that entry from the index, or leave the line blank to quit.

To run the client, execute the program `line_indexer.par`:

```
./line_indexer.par ./output/part-00000
The index file 'part-00000' will then be loaded and the following prompt displayed:
*****
* Line Indexer Client *
*****
Enter a word to query the index for.
Or 'exit' to quit.
>
```

Enter a word and the index listings generated by the map reduce will be printed:

```
*****
* Line Indexer Client *
*****
Enter a word to query the index for.
Or 'exit' to quit.
> ship
Index for: ship
119489 But we will ship him hence: and this vile deed
89045 Have lost my way for ever: I have a ship
89830 I will possess you of that ship and treasure.
34636 That their designment halts: a noble ship of Venice
34791 Third Gentleman The ship is here put in,
37636 The riches of the ship is come on shore!
23092 Ere he take ship for France, and in Southampton.
74453 For there I'll ship them all for Ireland.
131177 Like to a ship that, having 'scaped a tempest,
108668 And ship from thence to Flanders.
132817 Whiles, in his moan, the ship splits on the rock,
61495 ANTIGONUS Thou art perfect then, our ship hath touch'd upon
65730 ship boring the moon with her main-mast, and anon
66517 Clown I would you had been by the ship side, to have
114335 new ship to purge melancholy and air himself: for,
17361 Now am I like that proud insulting ship
112791 FERDINAND The ship is under sail, and here she comes amain.
```



48720 And in their ship I am sure Lorenzo is not.  
48868 SALARINO He came too late, the ship was under sail:  
55861 a ship of rich lading wrecked on the narrow seas;  
56018 a tall ship lie buried, as they say, if my gossip  
3094 Assure yourself, after our ship did split,  
25148 Make such unquiet, that the ship  
51644 So up and down the poor ship drives:  
54363 the wind is loud, and will not lie till the ship be  
92888 Lysimachus our Tyrian ship espies,  
7115 PROTEUS Go, go, be gone, to save your ship from wreck,  
616 SCENE A ship at Sea: an island.  
676 SCENE I On a ship at sea: a tempestuous noise  
2924 GONZALO I'll warrant him for drowning; though the ship were  
4713 It should the good ship so have swallow'd and  
14860 PROSPERO Of the king's ship  
15389 Supposing that they saw the king's ship wreck'd  
98063 I'll bring you to your ship and so to Naples,  
5225 Our helpful ship was splitted in the midst;  
5606 At length, another ship had seized on us;  
45794 If any ship put out, then straight away.  
50454 The ship is in her trim; the merry wind  
50657 What ship of Epidamnum stays for me?  
50696 DROMIO OF SYRACUSE A ship you sent me to, to hire waftage.

Found 42 occurrences.

After playing around with the client, enter 'exit' to shutdown the program.