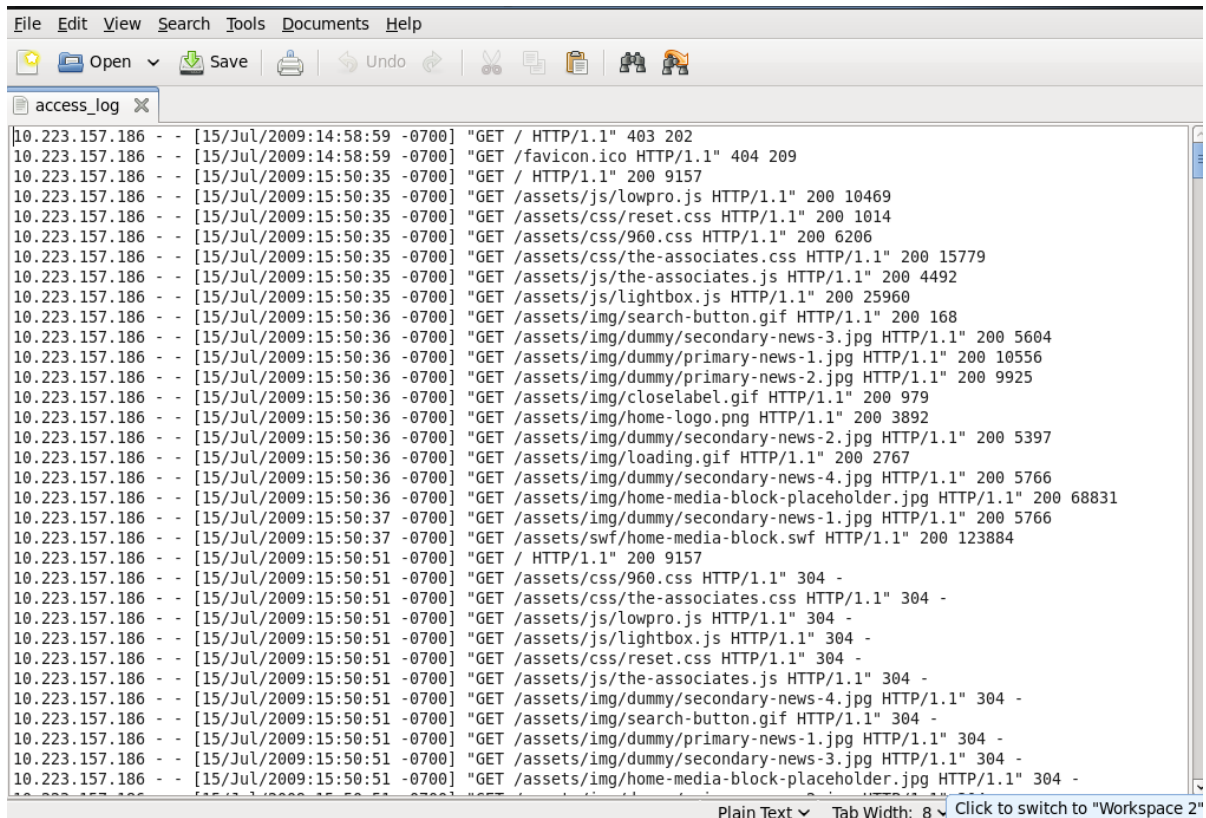


## Text file



## Code:

```
import java.io.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.*;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.lib.output.*;
import org.apache.hadoop.util.*;
public class webstatus extends Configured implements Tool
{
    //Map class
    public static class PartMonth extends Mapper<LongWritable,Text,Text,Text>
    {
        public void map(LongWritable key, Text value, Context context)
        {
            try{
                String [] str2 = value.toString().split(" ");
                String status=str2[str2.length-2];
                String ip = str2[0];
                context.write(new Text(ip), new Text(status));
            }
            catch(Exception e)
            {
                System.out.println(e.getMessage());
            }
        }
    }
}
```

```

}
}
}
//Reducer class
public static class ReduceClass extends Reducer<Text,Text,Text,IntWritable>
{
    public int max = -1;
    public void reduce(Text key, Iterable <Text> values, Context context) throws
    IOException, InterruptedException
    {
        int ipcount = 0;
        for (Text val : values)
        {
            ipcount += 1;
        }
        context.write(new Text(key), new IntWritable(ipcount));
    }
}
//Partitioner class
public static class CaderPartitioner extends
Partitioner < Text, Text >
{

    public int getPartition(Text key, Text value, int numReduceTasks)
    {
        {
            String status = value.toString();

            if(numReduceTasks == 0)
            {
                return 0;
            }
            else if(status.startsWith("2"))
            {
                return 1 % numReduceTasks;
            }
            else if(status.startsWith("3"))
            {
                return 2 % numReduceTasks;
            }
            else if(status.startsWith("4"))
            {
                return 3 % numReduceTasks;
            }
            else if(status.startsWith("5"))
            {
                return 4 % numReduceTasks;
            }
            else
            {
                return 5 % numReduceTasks;
            }
        }
    }
}
}

```

```

public int run(String[] arg) throws Exception
{
    Configuration conf = getConf();
    Job job = new Job(conf, "status count");
    job.setJarByClass(webstatus.class);
    FileInputFormat.setInputPaths(job, new Path(arg[0]));
    FileOutputFormat.setOutputPath(job, new Path(arg[1]));
    job.setMapperClass(PartMonth.class);
    job.setMapOutputKeyClass(Text.class);
    job.setMapOutputValueClass(Text.class);
    //set partitioner statement
    job.setPartitionerClass(CaderPartitioner.class);
    job.setReducerClass(ReduceClass.class);
    job.setNumReduceTasks(6);
    job.setInputFormatClass(TextInputFormat.class);
    job.setOutputFormatClass(TextOutputFormat.class);
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    System.exit(job.waitForCompletion(true)? 0 : 1);
    return 0;
}

public static void main(String ar[]) throws Exception
{
    int res = ToolRunner.run(new Configuration(), new webstatus(), ar);
    System.exit(0);
}
}

```

#### Code execution:

```

[training@localhost ~]$ hadoop jar /home/training/webstatus.jar /user/training/a
ccess_log /user/training/statuslogs
21/09/03 19:00:51 WARN mapred.JobClient: Use GenericOptionsParser for parsing th
e arguments. Applications should implement Tool for the same.
21/09/03 19:00:51 INFO input.FileInputFormat: Total input paths to process : 1
21/09/03 19:00:51 WARN snappy.LoadSnappy: Snappy native library is available
21/09/03 19:00:51 INFO snappy.LoadSnappy: Snappy native library loaded
21/09/03 19:00:51 INFO mapred.JobClient: Running job: job_202109031852_0001
21/09/03 19:00:52 INFO mapred.JobClient: map 0% reduce 0%
21/09/03 19:01:08 INFO mapred.JobClient: map 10% reduce 0%
21/09/03 19:01:10 INFO mapred.JobClient: map 18% reduce 0%
21/09/03 19:01:11 INFO mapred.JobClient: map 20% reduce 0%
21/09/03 19:01:13 INFO mapred.JobClient: map 25% reduce 0%
21/09/03 19:01:24 INFO mapred.JobClient: map 32% reduce 2%
21/09/03 19:01:25 INFO mapred.JobClient: map 38% reduce 2%
21/09/03 19:01:27 INFO mapred.JobClient: map 43% reduce 2%
21/09/03 19:01:28 INFO mapred.JobClient: map 48% reduce 2%
21/09/03 19:01:30 INFO mapred.JobClient: map 50% reduce 4%
21/09/03 19:01:33 INFO mapred.JobClient: map 50% reduce 5%
21/09/03 19:01:37 INFO mapred.JobClient: map 58% reduce 5%
21/09/03 19:01:39 INFO mapred.JobClient: map 65% reduce 5%
21/09/03 19:01:40 INFO mapred.JobClient: map 68% reduce 5%
21/09/03 19:01:42 INFO mapred.JobClient: map 72% reduce 5%
21/09/03 19:01:44 INFO mapred.JobClient: map 75% reduce 5%
21/09/03 19:01:45 INFO mapred.JobClient: map 75% reduce 6%
21/09/03 19:01:46 INFO mapred.JobClient: map 75% reduce 8%
21/09/03 19:01:51 INFO mapred.JobClient: map 82% reduce 8%
21/09/03 19:01:53 INFO mapred.JobClient: map 95% reduce 8%
21/09/03 19:01:54 INFO mapred.JobClient: map 100% reduce 8%
21/09/03 19:01:55 INFO mapred.JobClient: map 100% reduce 9%
21/09/03 19:01:56 INFO mapred.JobClient: map 100% reduce 21%
21/09/03 19:01:58 INFO mapred.JobClient: map 100% reduce 22%
21/09/03 19:02:01 INFO mapred.JobClient: map 100% reduce 27%
21/09/03 19:02:03 INFO mapred.JobClient: map 100% reduce 44%
21/09/03 19:02:04 INFO mapred.JobClient: map 100% reduce 49%
21/09/03 19:02:06 INFO mapred.JobClient: map 100% reduce 50%

```

access\_log (~/Desktop) - gedit

training@localhost:~ Java - WordCount/sr... [training - File Brow... [HDFS:/user/trainin... access\_log (~/Desk...

**Output:**

**File: /user/training/statuslogs/part-r-00001**

Goto :

[Go back to dir listing](#)  
[Advanced view/download options](#)

[View Next chunk](#)

10.1.1.113	1
10.1.1.125	12
10.1.1.144	1
10.1.1.195	4
10.1.1.236	12
10.1.1.5	1
10.1.10.155	2
10.1.10.197	2
10.1.10.198	1
10.1.10.48	1
10.1.10.5	1
10.1.100.104	1
10.1.100.13	1
10.1.100.138	1
10.1.100.183	14
10.1.100.199	35
10.1.100.5	1
10.1.101.135	29
10.1.101.140	1
10.1.101.141	1
10.1.101.204	1
10.1.101.206	1
10.1.101.208	1
10.1.101.31	1

**File: /user/training/statuslogs/part-r-00002**

Goto :

[Go back to dir listing](#)  
[Advanced view/download options](#)

[View Next chunk](#)

10.1.101.63	1
10.1.102.106	1
10.1.106.230	5
10.1.11.81	3
10.1.120.214	10
10.1.122.130	2
10.1.122.57	1
10.1.126.21	3
10.1.127.186	1
10.1.129.142	2
10.1.143.54	2
10.1.146.91	20
10.1.147.37	1
10.1.151.231	1
10.1.151.249	1
10.1.152.49	6
10.1.156.219	3
10.1.158.223	1
10.1.160.24	2
10.1.165.151	1
10.1.165.27	1
10.1.166.226	1
10.1.166.60	11
10.1.166.93	3

File: [/user/training/statuslogs/part-r-00003](#)

Goto :

[Go back to dir listing](#)  
[Advanced view/download options](#)

[View Next chunk](#)

10.1.123.80	1
10.1.124.195	1
10.1.129.86	3
10.1.131.79	1
10.1.133.67	1
10.1.138.3	1
10.1.141.120	1
10.1.146.91	8
10.1.148.181	1
10.1.169.153	2
10.1.169.95	2
10.1.181.142	42
10.1.186.241	7
10.1.187.149	1
10.1.197.227	6
10.1.2.113	1
10.1.219.14	4
10.1.219.212	25
10.1.224.169	1
10.1.232.249	1
10.1.232.31	16
10.1.232.59	1
10.1.234.164	1
10.1.245.107	46

File: [/user/training/statuslogs/part-r-00004](#)

Goto :

[Go back to dir listing](#)  
[Advanced view/download options](#)

10.97.165.10	5
--------------	---