

```

1  import java.io.IOException;
2  import java.net.URI;
3
4
5  import org.apache.hadoop.conf.Configuration;
6  import org.apache.hadoop.fs.FSDataInputStream;
7  import org.apache.hadoop.fs.FSDataOutputStream;
8  import org.apache.hadoop.fs.FileStatus;
9  import org.apache.hadoop.fs.FileSystem;
10 import org.apache.hadoop.fs.Path;
11
12 // The program merges all files located at a folder on a local file system and
13 // loads the outcomes of merge to HDFS as a single file
14
15 public class solution1 {
16
17     public static void main(String[] args) throws IOException {
18
19         // The program has two parameters:
20         // a path to folder on a local file system with the files to be merged
21         // a path and a name of file that contains the results of merge in HDFS
22         String localStr = args[0];
23         String hdfsStr = args[1];
24
25         // We start from creation of a an object with HDFS configuration
26         Configuration conf = new Configuration();
27
28         // Next, we create handles for input folder at a local file system and
29         // and handle for output file in HDFS
30         FileSystem hdfs = FileSystem.get(URI.create(hdfsStr), conf);
31         FileSystem local = FileSystem.getLocal(conf);
32
33         // Next, we create a string with a path a name of a folder with input files and ..
34         Path inputDir = new Path(localStr);
35         String folderName = inputDir.getName();
36         // ... a path to a file in HDFS
37         Path hdfsFile = new Path(hdfsStr, folderName);
38
39         try {
40             // Next, we create a list of names of files located in a folder on a local file system
41             // and ...
42             FileStatus[] inputFiles = local.listStatus(inputDir);
43             // ... and a handle output file in HDFS
44             FSDataOutputStream out = hdfs.create(hdfsFile);
45
46             // Next, we iterate over the files in a folder on a local file system and we copy the
47             // files
48             // to a buffer and buffer is immediately written to an output file in HDFS
49             for (int i=0; i<inputFiles.length; i++) {
50                 System.out.println(inputFiles[i].getPath().getName());
51                 FSDataInputStream in = local.open(inputFiles[i].getPath());
52                 byte buffer[] = new byte[256];
53                 int bytesRead = 0;
54                 while( (bytesRead = in.read(buffer)) > 0) {
55                     out.write(buffer, 0, bytesRead);
56                 }
57                 in.close();
58             }
59             out.close();
60         } catch (IOException e) {
61             e.printStackTrace();
62         }
63     }
64 }

```