RegexAcceptPathFilter

## 1. RegexAcceptPathFilter

package hdfs.uploadfile;  
  
import org.apache.hadoop.fs.Path;  
import org.apache.hadoop.fs.PathFilter;  
  
public class RegexAcceptPathFilter implements PathFilter {  
  
 private final String regex;  
 public static final String REGEX1 = "^.\*txt$";  
 public static final String REGEX2 = "^.\*\\.svn$";  
  
 public RegexAcceptPathFilter(String regex) {  
 if (regex.equals(REGEX1)) {  
 this.regex = REGEX1;  
 }else {  
 this.regex = REGEX2;  
 }  
 }  
  
 public boolean accept(Path path) {  
 boolean flag = path.toString().matches(regex);  
 if (regex.equals(REGEX1)) {  
 return flag;  
 }else {  
 return !flag;  
 }  
 }  
}

## 2. CopyManyFilesToHDFS

package hdfs.uploadfile;  
  
import org.apache.hadoop.conf.Configuration;  
import org.apache.hadoop.fs.\*;  
import org.apache.hadoop.io.IOUtils;  
  
import java.io.IOException;  
import java.net.URI;  
import java.net.URISyntaxException;  
  
public class CopyManyFilesToHDFS {  
 private static FileSystem fs = null;  
 private static FileSystem local = null;  
  
 public static void writeMergeFile(String path) throws IOException, URISyntaxException {  
 Configuration conf = new Configuration();  
 local = FileSystem.getLocal(conf);  
 //HDFS 接口  
 URI uri = new URI("hdfs://master:9000");  
 //获取文件系统对象  
 fs = FileSystem.get(uri, conf);  
 FileStatus[] localDirs = local.globStatus(new Path(path), new RegexAcceptPathFilter(RegexAcceptPathFilter.REGEX2));  
 Path[] dirs = FileUtil.stat2Paths(localDirs);  
  
 FSDataOutputStream out = null;  
  
 for (Path dir : dirs) {  
 String filename = dir.toUri().getPath();  
 //目录下的txt文件  
 FileStatus[] localStatuses = local.globStatus(new Path(filename+"/\*"), new RegexAcceptPathFilter(RegexAcceptPathFilter.REGEX1));  
 Path[] listPaths = FileUtil.stat2Paths(localStatuses);  
 //输出路径  
 Path block = new Path("hdfs://master:9000/middle/"+dir.getName());  
// Path block = new Path("C:\\Users\\dishui\\Downloads\\weather\\"+dir.getName()+".txt");  
 out = fs.create(block);  
 for (Path path1 : listPaths) {  
 FSDataInputStream open = local.open(path1);  
 IOUtils.copyBytes(open,out,1024);  
 }  
 // 写入本地目录  
// local.createNewFile(block);  
 fs.createNewFile(block);  
 }  
 }  
  
 public static void main(String[] args) throws IOException, URISyntaxException {  
 writeMergeFile("C:\\Users\\dishui\\Downloads\\73\\\*");  
 }  
  
}

3. 结果

