package com.twitter.search.earlybird.partition;

import java.io.IOException;

import org.apache.hadoop.conf.Configuration;

import org.apache.hadoop.fs.FileStatus;

import org.apache.hadoop.fs.FileSystem;

import org.apache.hadoop.fs.Path;

public final class HdfsUtil {

private HdfsUtil() {

}

public static FileSystem getHdfsFileSystem() throws IOException {

Configuration config = new Configuration();

// Since earlybird uses hdfs from different threads, and closes the FileSystem from

// them independently, we want each thread to have its own, new FileSystem.

return FileSystem.newInstance(config);

}

/\*\*

\* Checks if the given segment is present on HDFS

\*/

public static boolean segmentExistsOnHdfs(FileSystem fs, SegmentInfo segmentInfo)

throws IOException {

String hdfsBaseDirPrefix = segmentInfo.getSyncInfo().getHdfsUploadDirPrefix();

FileStatus[] statuses = fs.globStatus(new Path(hdfsBaseDirPrefix));

return statuses != null && statuses.length > 0;

}

}