**Assignment 18.3**

Merge the small files in hadoop by following the below blog.

<https://acadgild.com/blog/merging-files-hdfs/>

Ans. Merging multiple files is useful when you want to retrieve the output of a [MapReduce](https://acadgild.com/big-data/big-data-development-training-certification) computation with multiple reducers, where each reducer produces a part of the output.

The HDFS **getmerge** command can copy the files present in a given path in HDFS to a single concatenated file in the local filesystem.

**hadoop fs -getmerge /user/hadoop/demo\_files merged.txt**

The **getmerge** command has the following syntax:

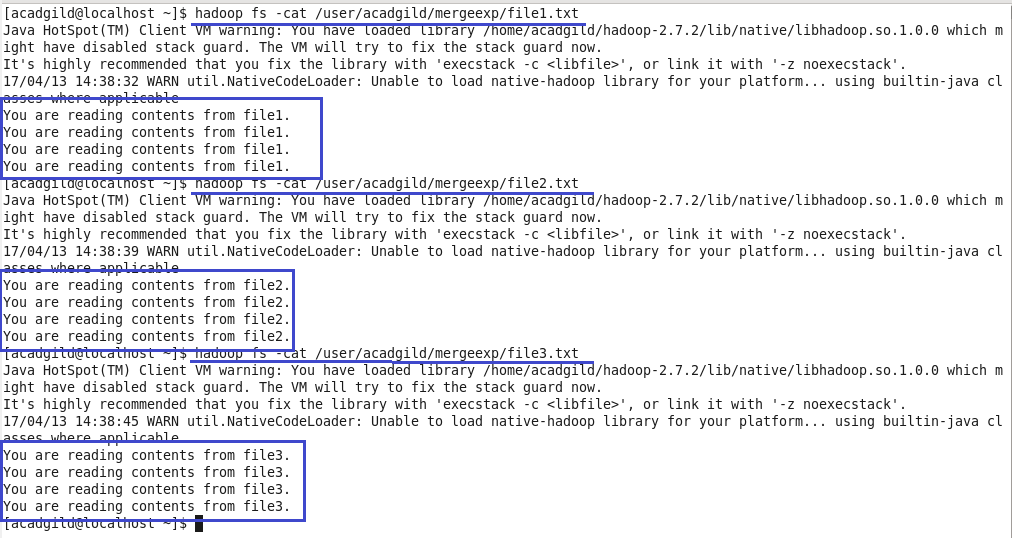
**Hadoop fs -getmerge -nl <source file path> <local system destination path>**

The **getmerge**command has three parameters:

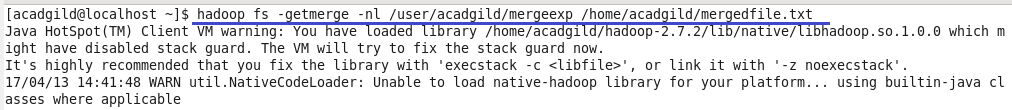
* <src files> is the HDFS path to the directory that contains the files to be concatenated
* <dist file> is the local filename of the merged file
* [-nl] is an optional parameter that adds a new line in the result file.

**Steps to merge the files**

There are 3 text files in directory mergeexp as file1.txt, file2.txt and file3.txt. The contents of these files are:



Merging of all the files present in mergeexp into one file merged\_file.txt on local file system.



Contents of the merged\_file.txt:

