Assignment 18.3

Merge the small files in Hadoop.

Merging multiple files is useful when you want to retrieve the output of a Map Reduce 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 file system.

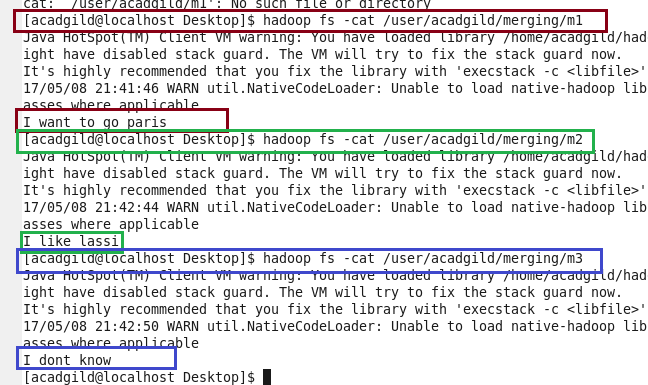
The **getmerge** command has the following syntax:

**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.

There are 3 text files in directory merging as m1, m2 and m3. The contents of these files are:



Merging of all the files present in merging directory into one file merging\_Merging.txt on local file system:

Output :

