**Assignment 3.4**

**Problem Statement 1:**

1. Explain the importance of Namenode in Hadoop cluster.

**Answer:**

**Importance of NameNode:**

The Hadoop cluster mainly comprises of:

* The NameNode is the master daemon in HDFS.
* It stores metadata about the data stored in the Hadoop file system.
* It is single point of failure and the Hadoop storage is managed by the Namenode.
* The name node is responsible for the synchronization between the data nodes.
* Whenever a file is placed in the Hadoop cluster a corresponding entry of it location is maintained by the NameNode and hence to ensure the high availability of data in HDFS name node plays very important role.

**Problem Statement 2:**

1. Practice the beginners commands for HDFS. Share the screenshot of the commands used.

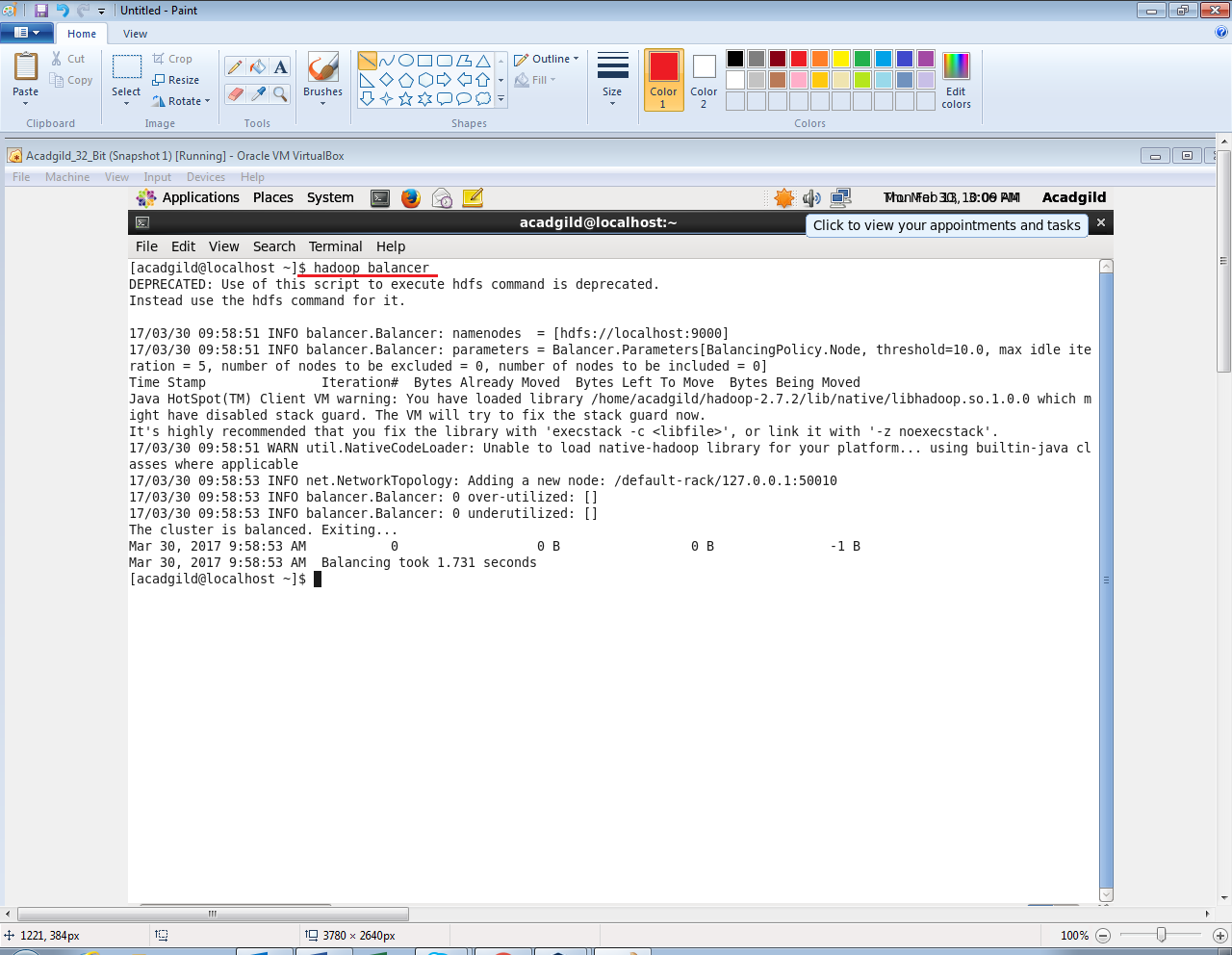
**Data set:**

The data set can be downloaded from the following link

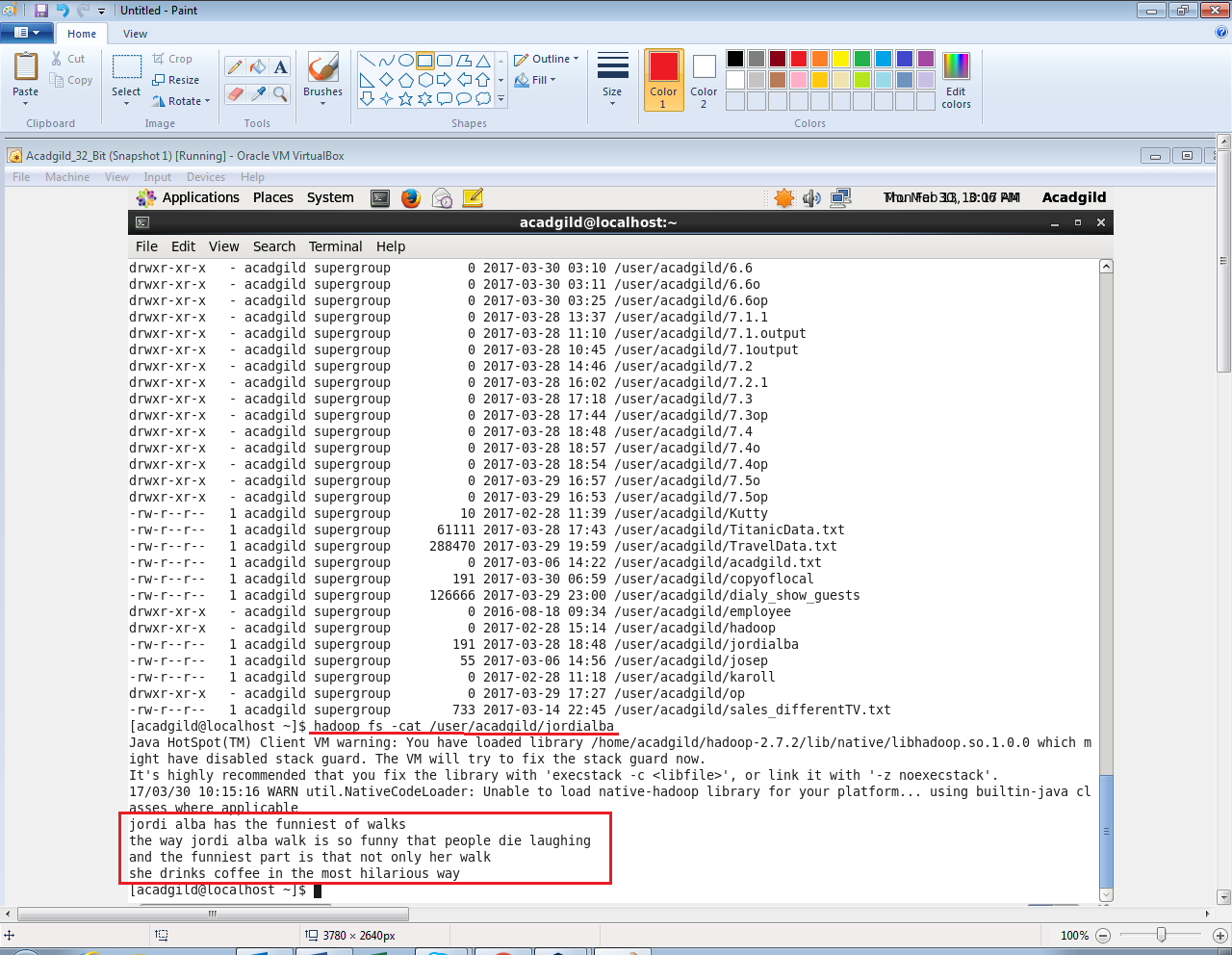
<https://acadgild.com/blog/hdfs-commands-for-beginners/>

**Screenshots:**

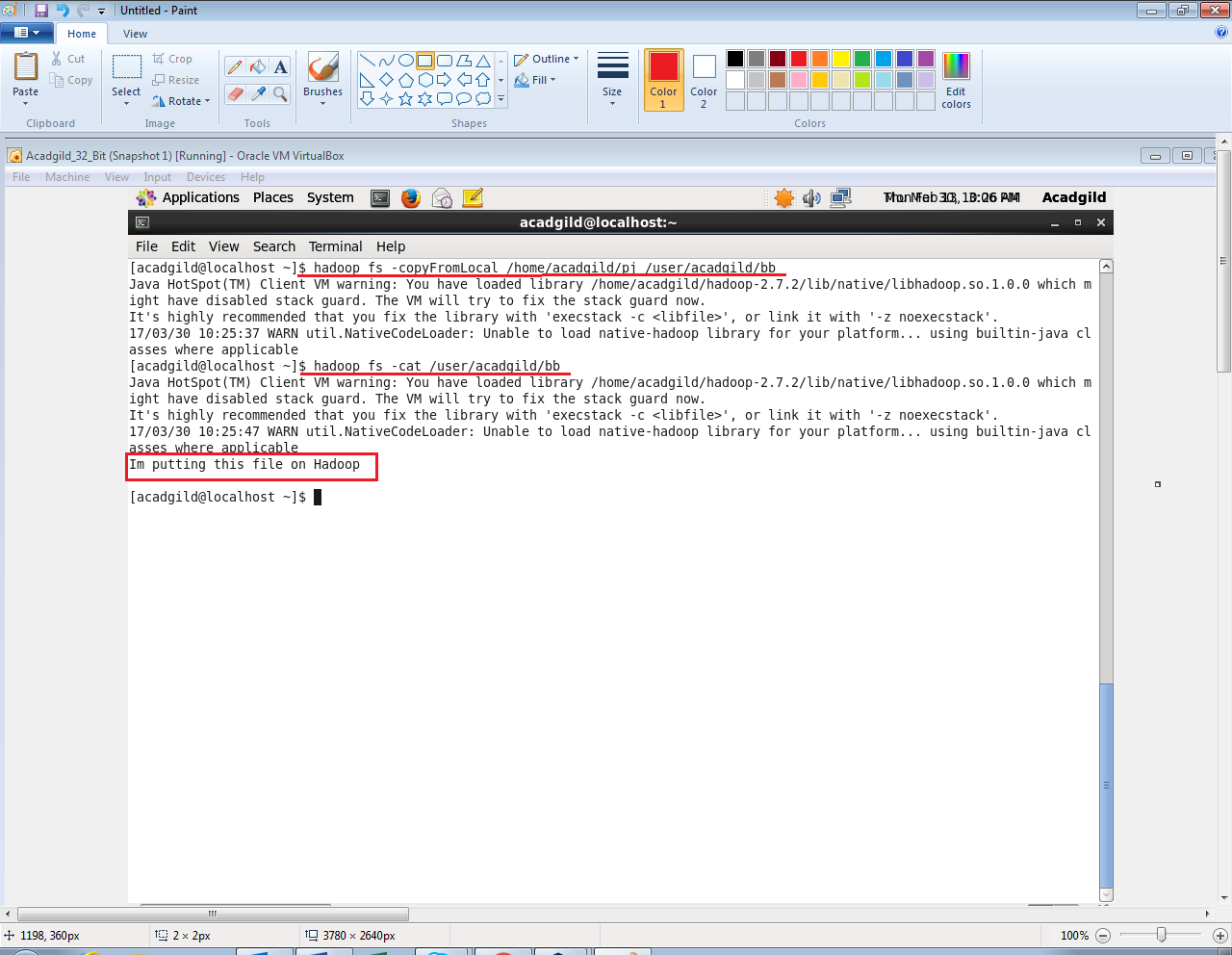
**Hadoop Balancer:**



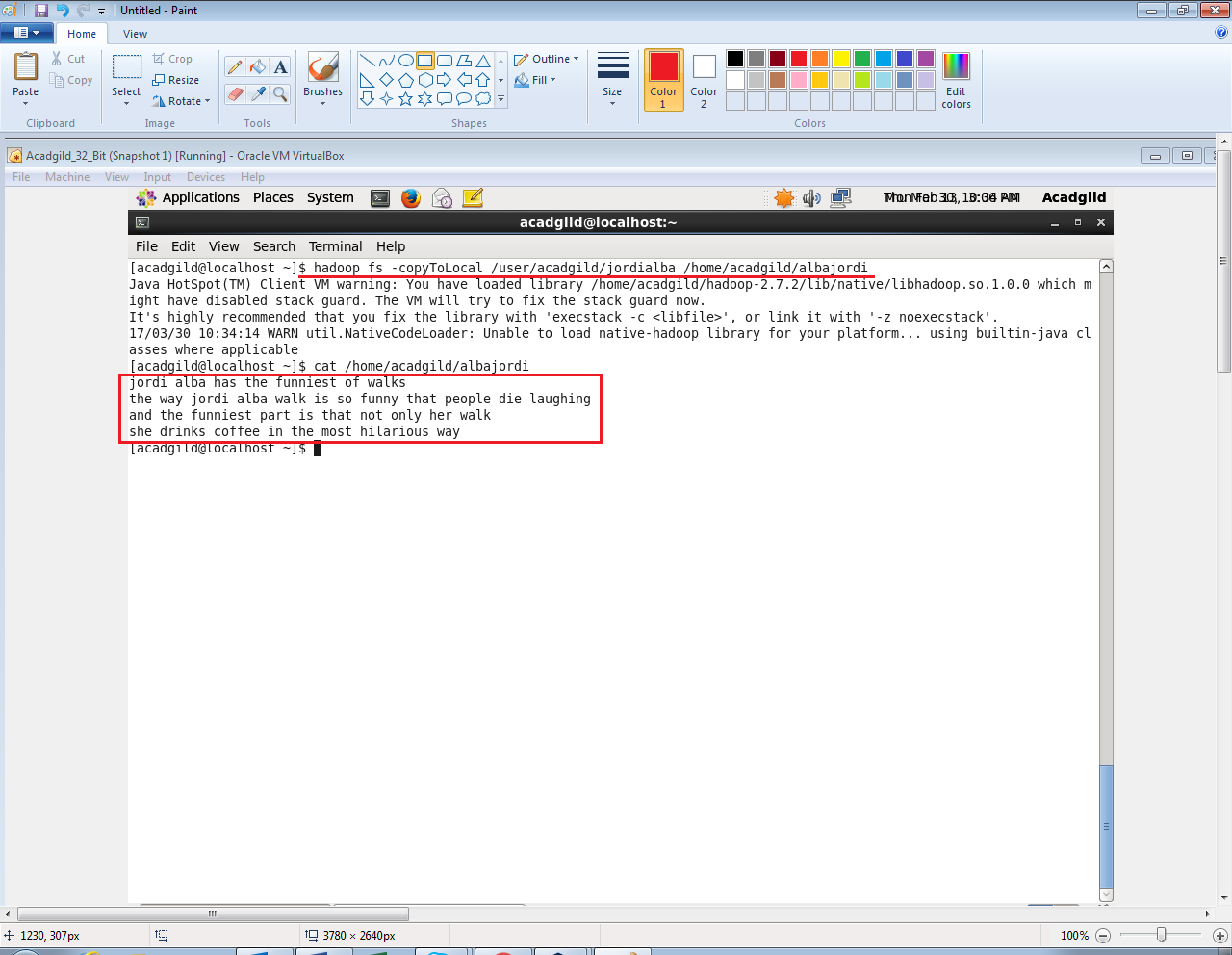
**Hadoop cat:**



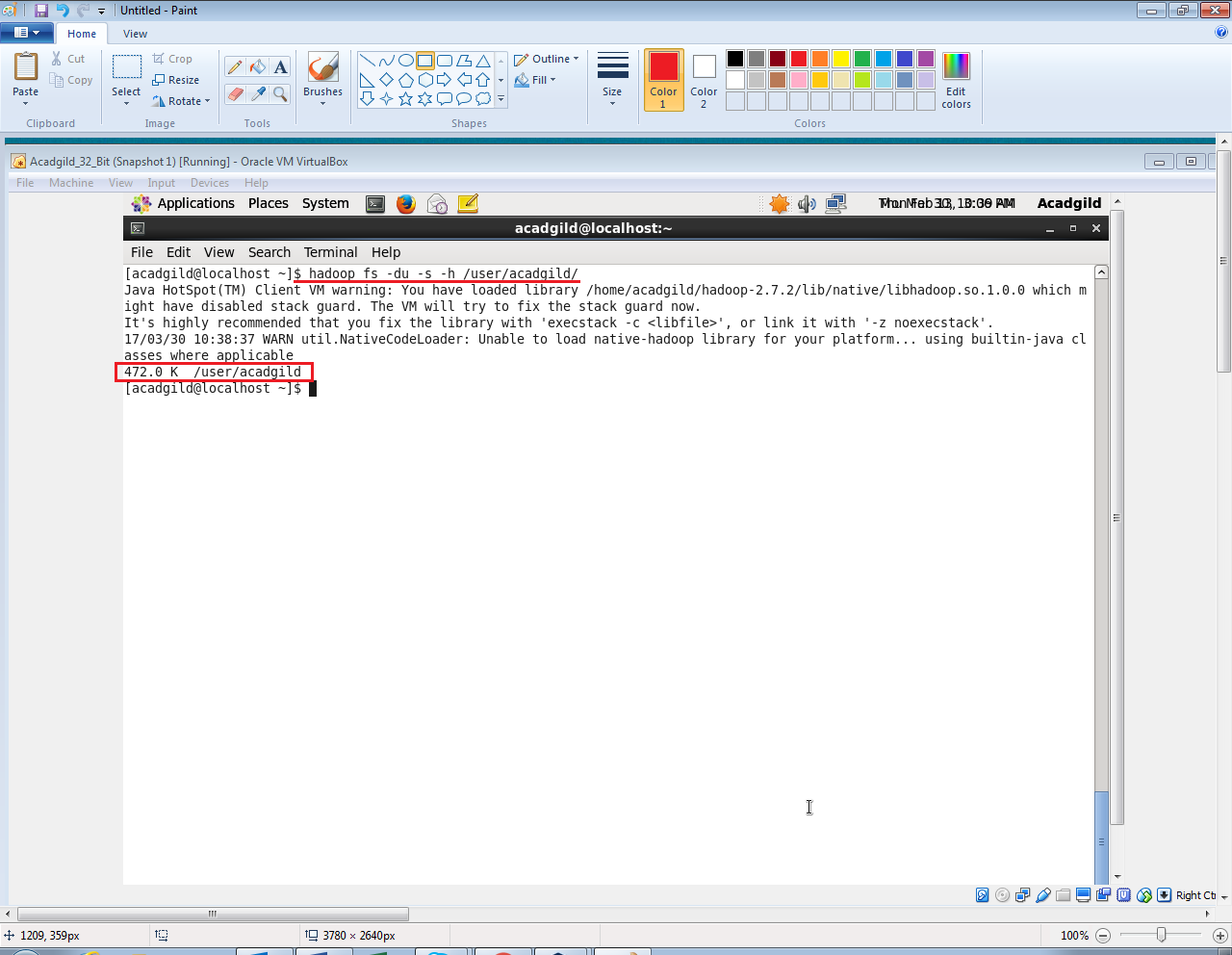
**Hadoop copy From Local:**



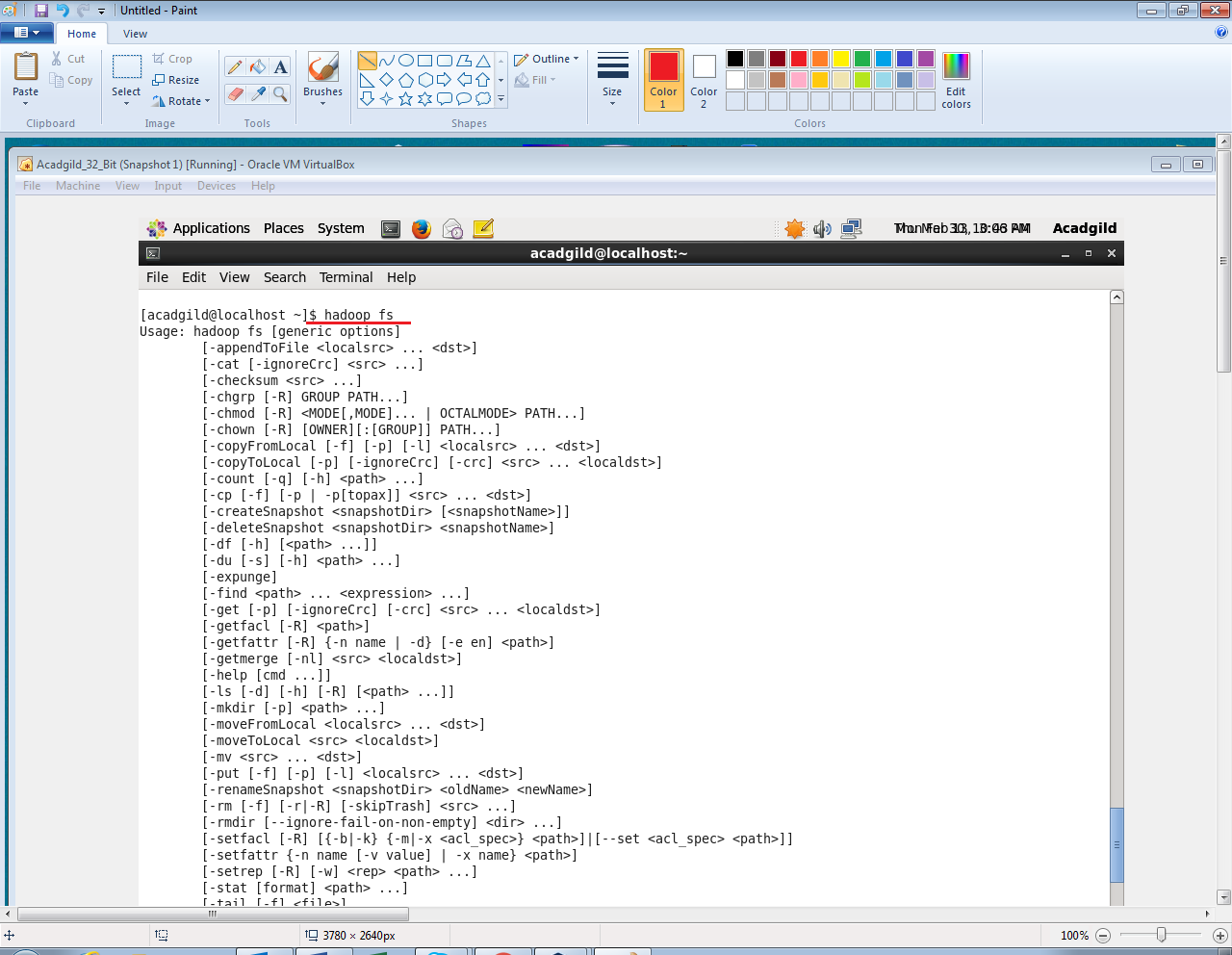
**Hadoop copy To Local:**

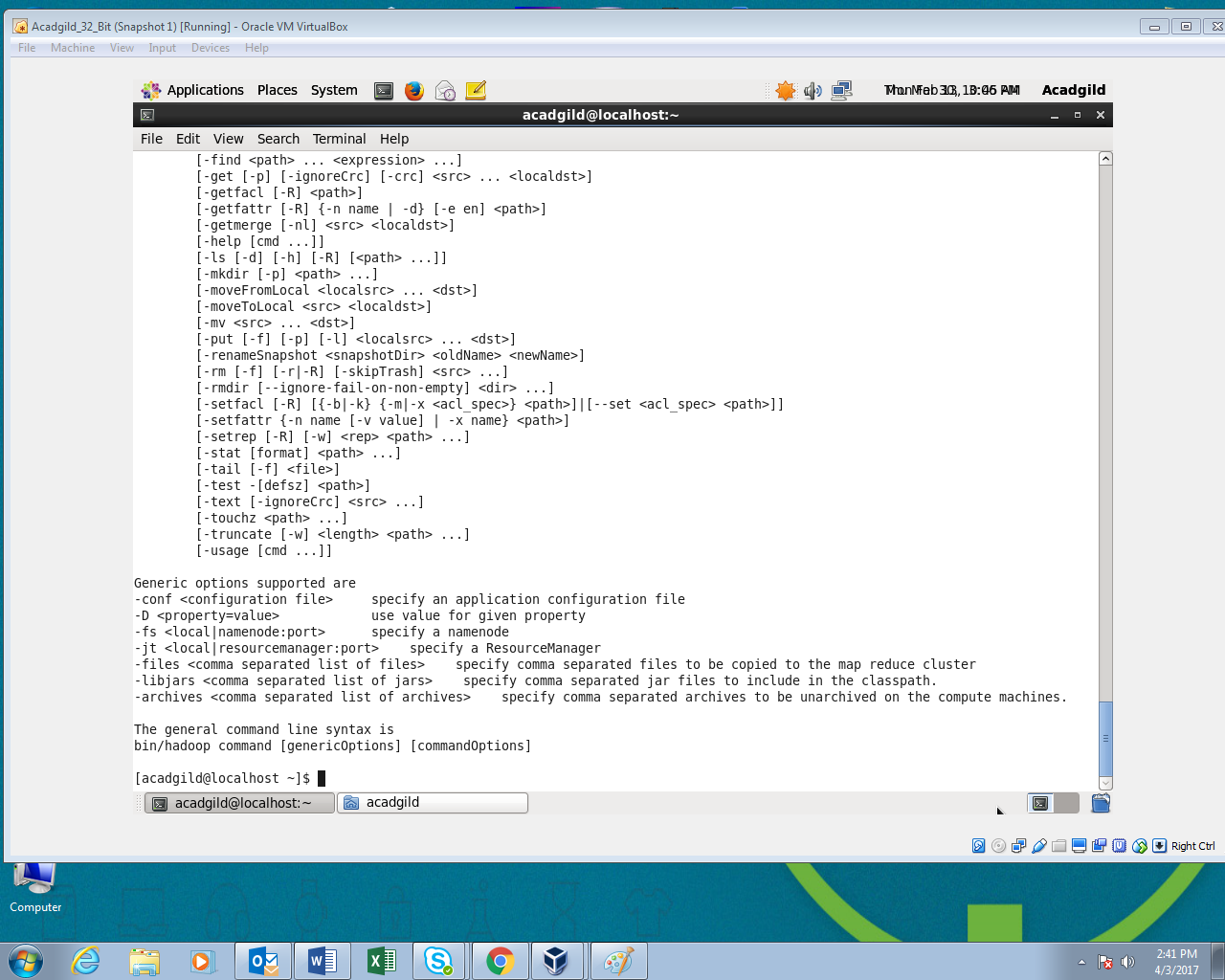


**Hadoop du:**

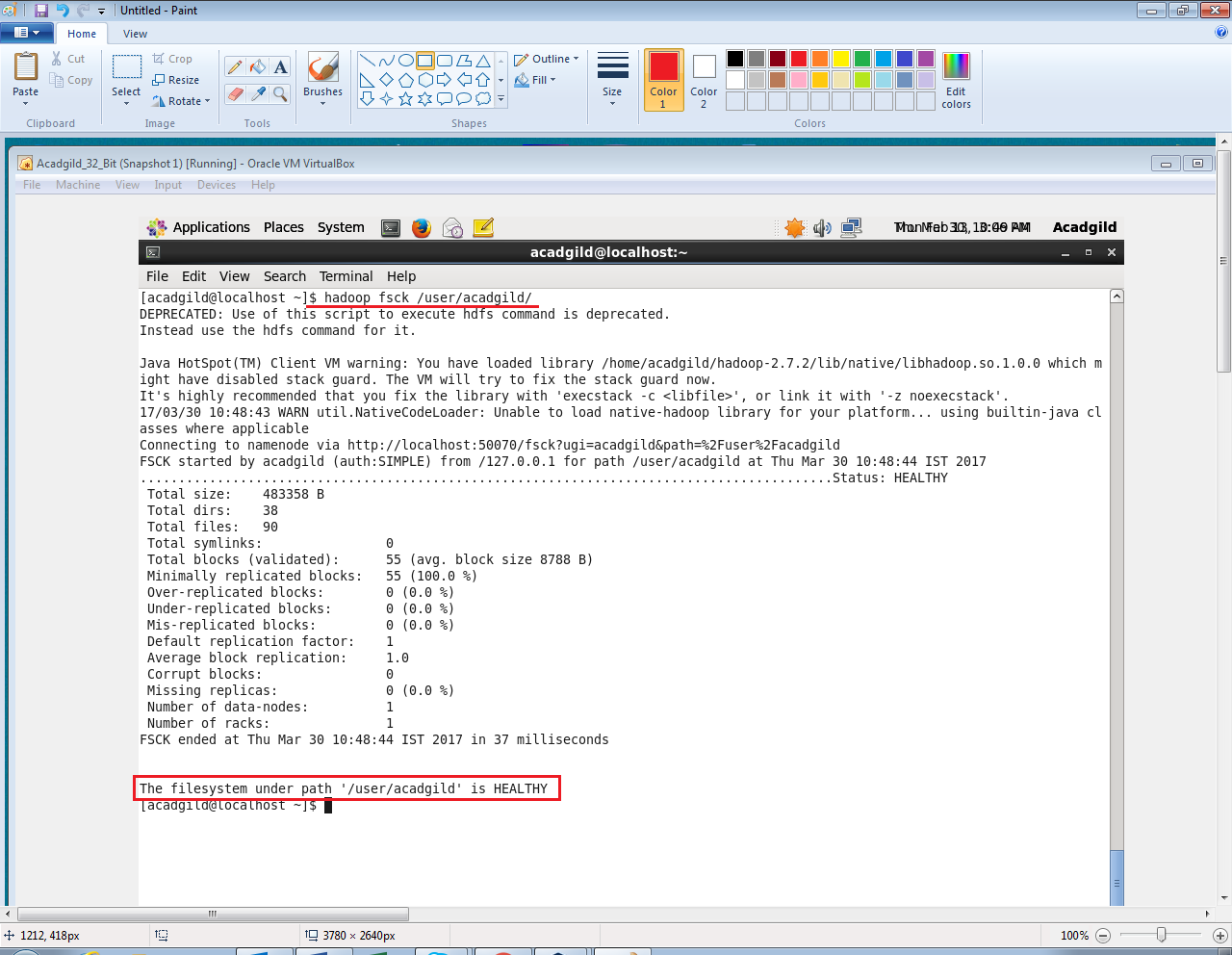


**Hadoop fs:**

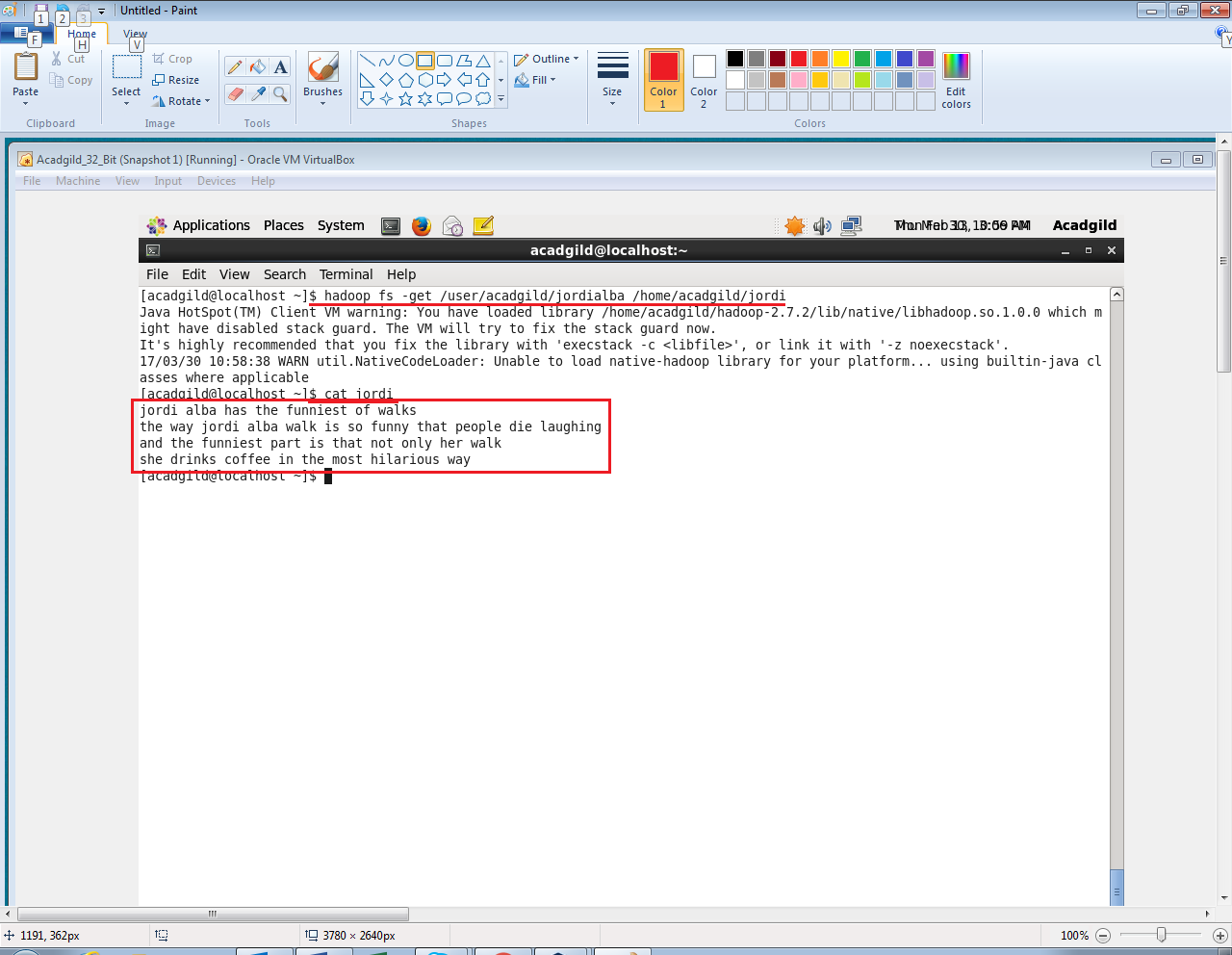




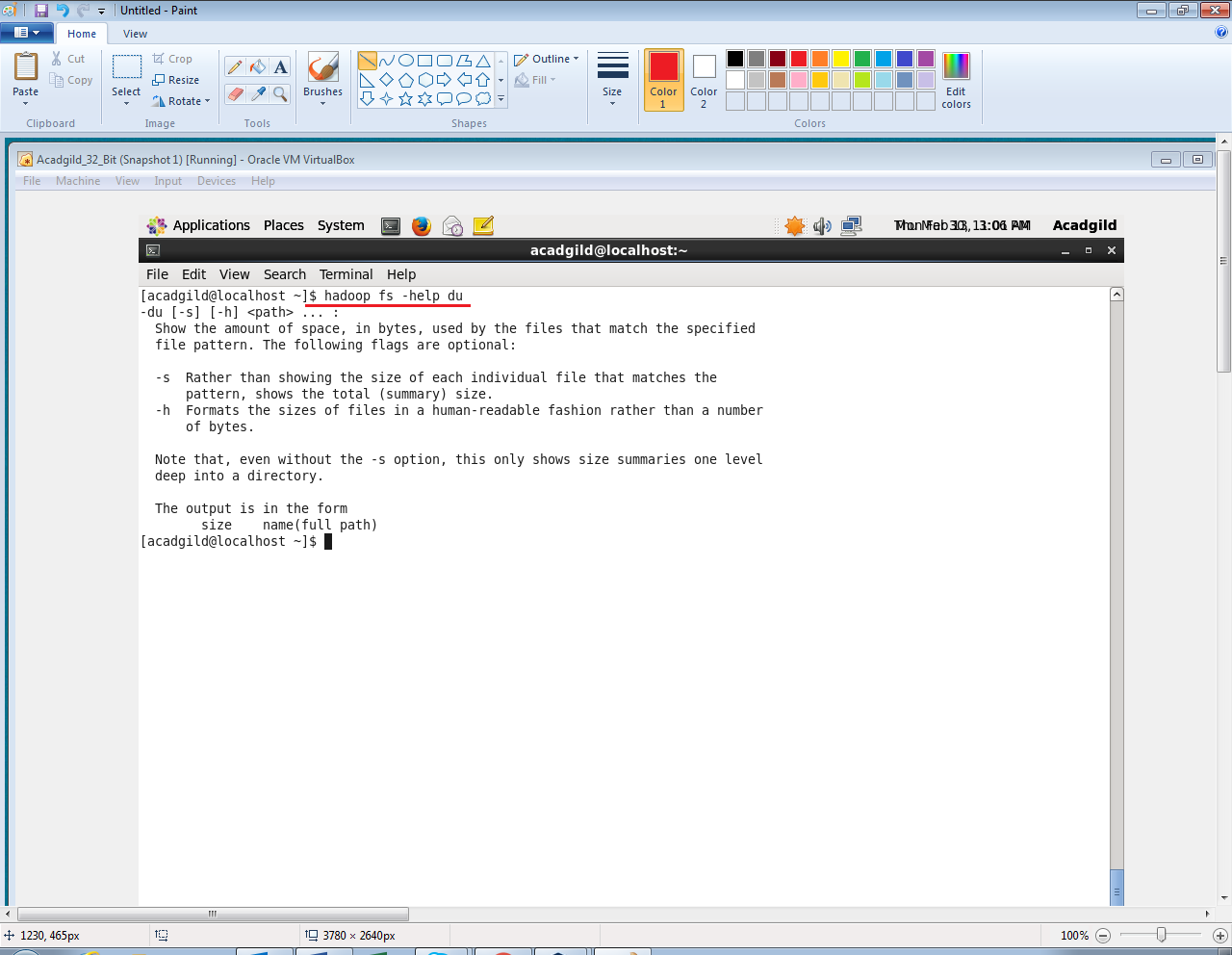
**Hadoop fsck:**



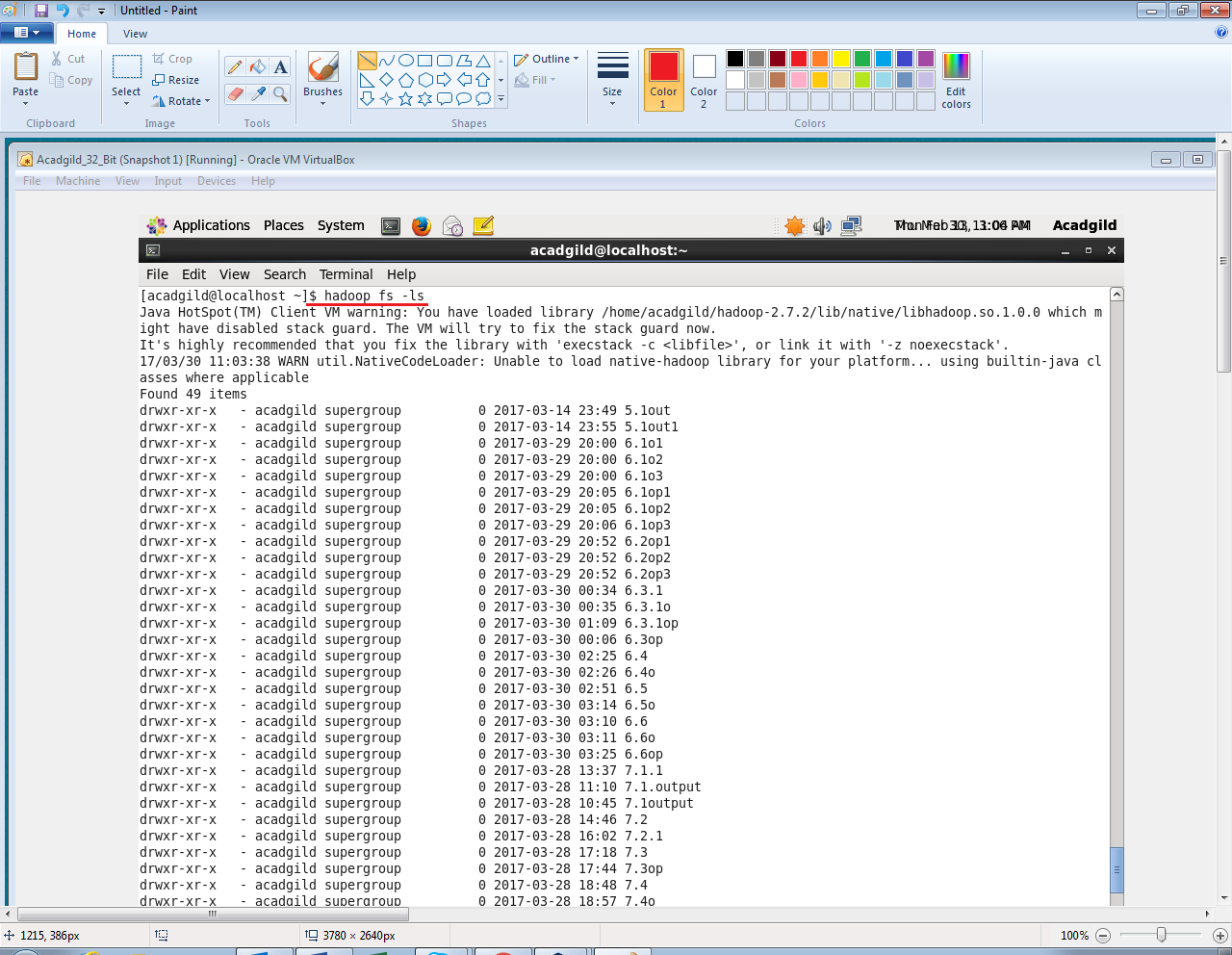
**Hadoop get:**



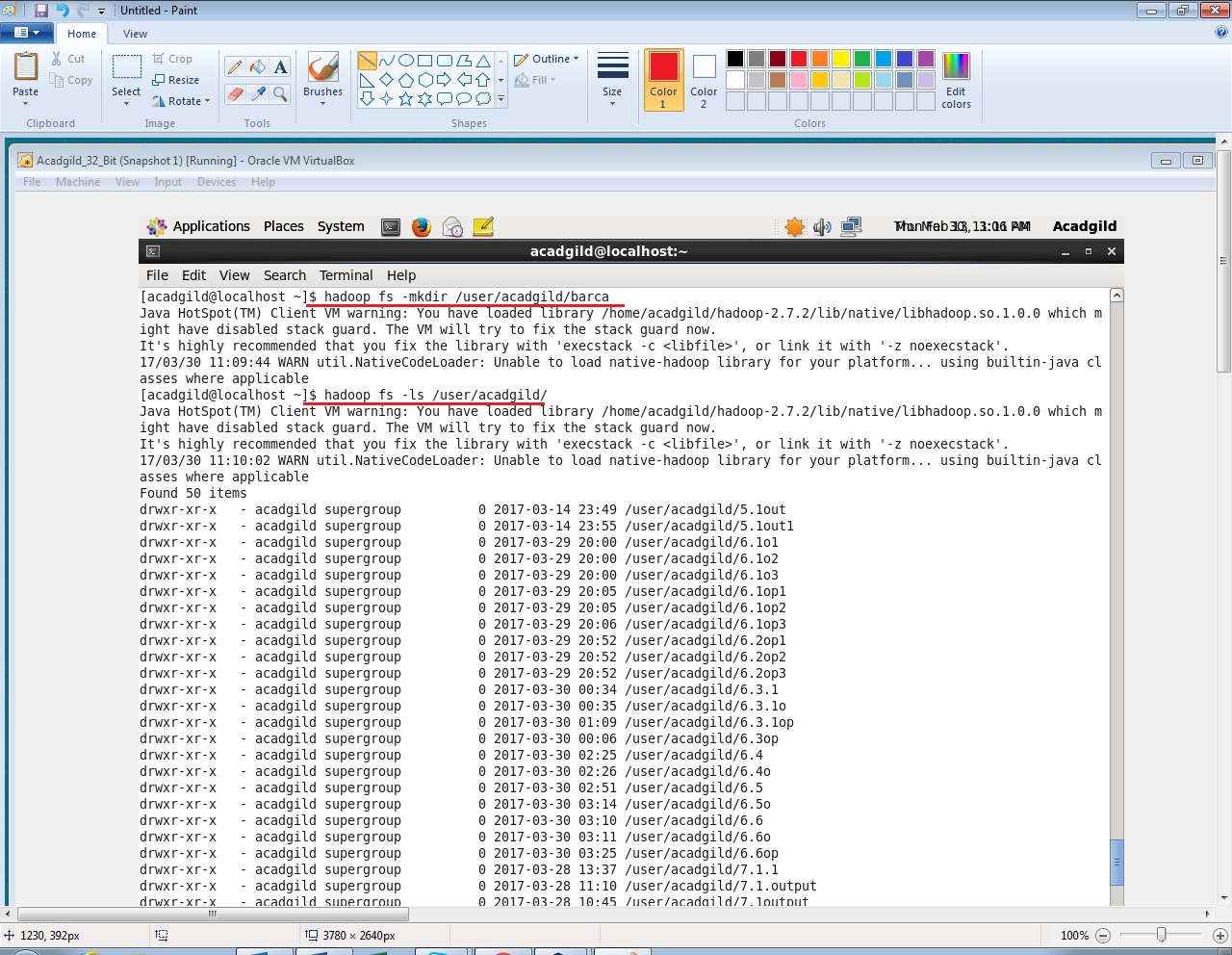
**Hadoop help:**

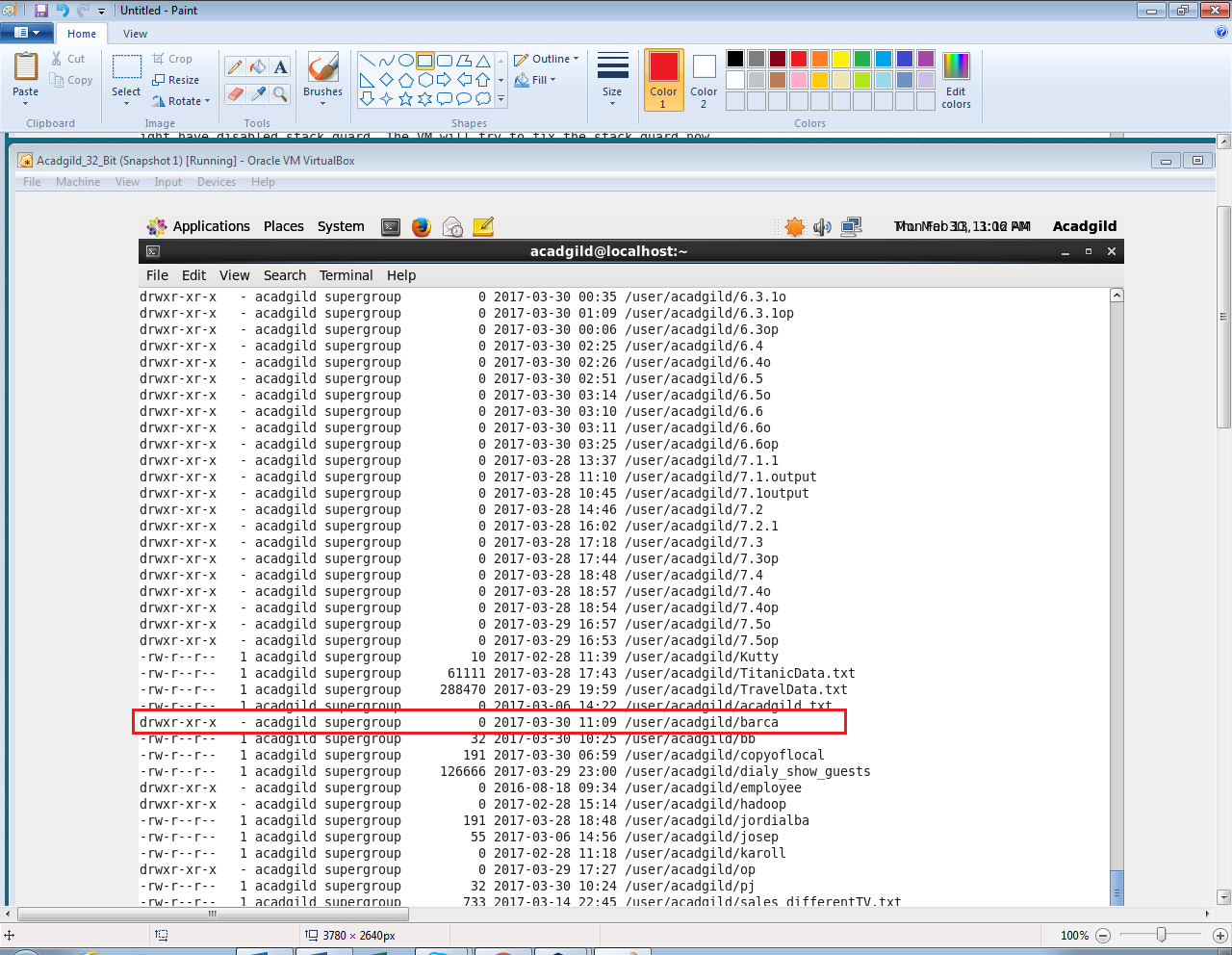


**Hadoop list:**

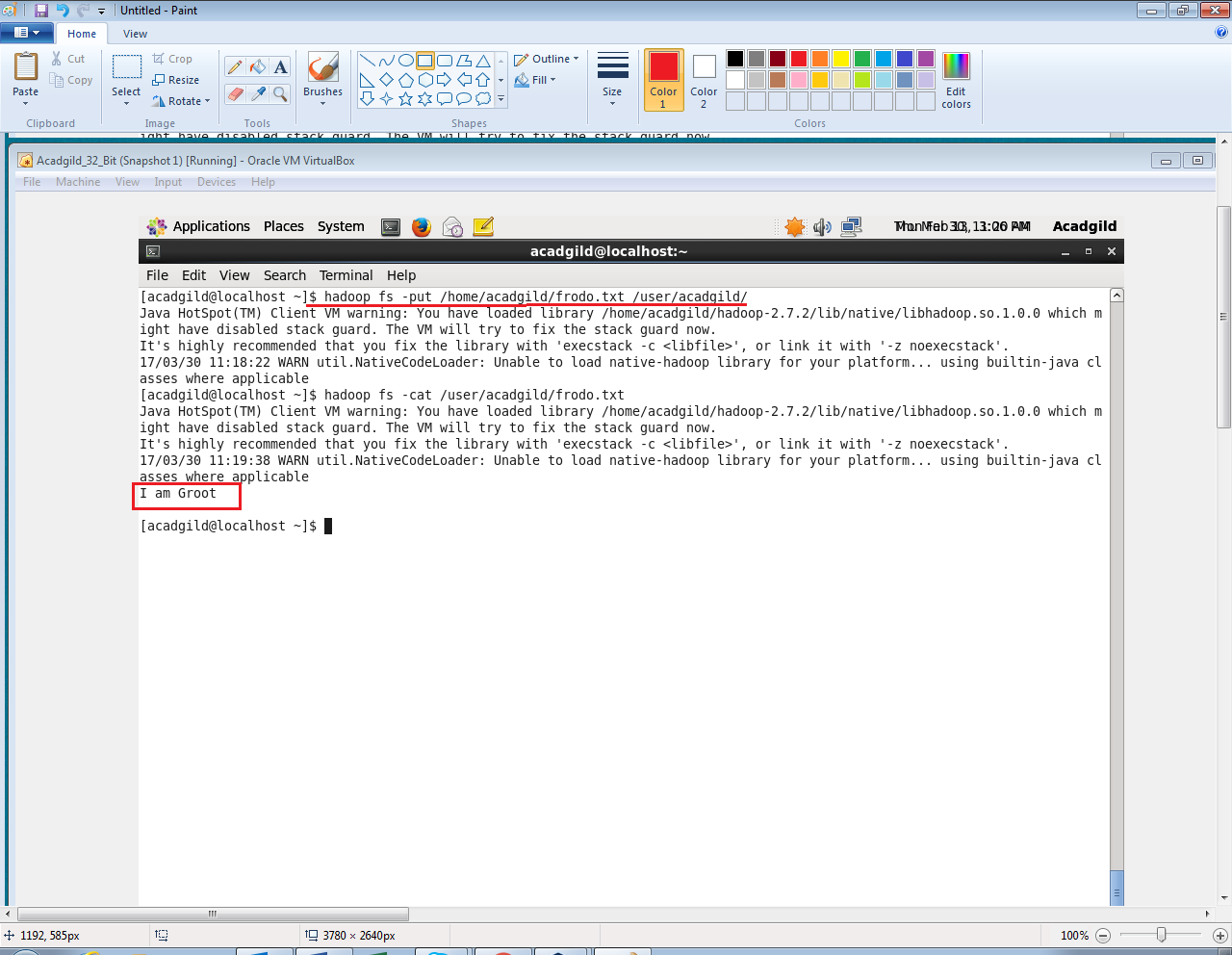


**Hadoop mkdir:**





**Hadoop put:**



**Hadoop rm:**

