Challenges in installation and Running Program

The benefits of Hadoop distributed processing are obvious and do not need to be stated yet again.

For big data processing this tool is both a huge asset and a competitive necessity. But what does deserve much focus is how hard it is to install, implement and optimize an environment successfully.

The first step to avoid repeating that fate is to understand what kinds of challenges are commonly experienced.

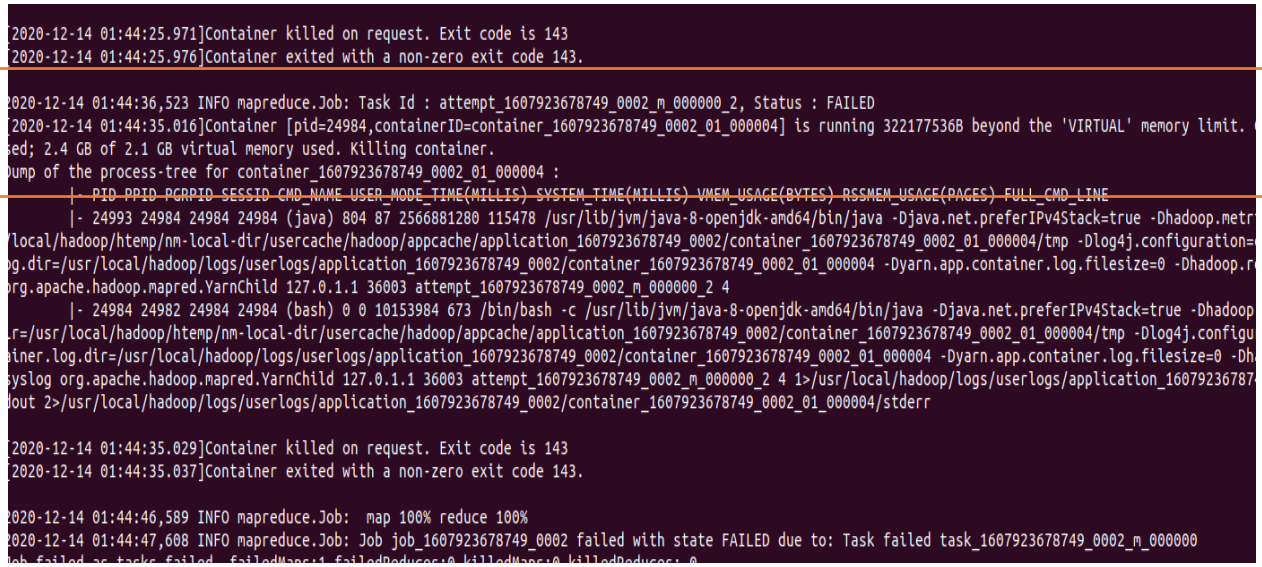
Here are 3 challenges that I encountered in installing and running local Hadoop program,

1. Path variables getting reset after VMware Restart.

   Ubuntu is installed on VMware. I observed that after machine restart path variables are getting lost in setup files. This issue is due to VMware.

2. Limited virtual memory allocation.

   Hadoop on local machine assigns limited virtual memory to each container. This makes difficult to process files having size more than 250MB on local machine.



3. Restart issue with Name node and data node.

Datanode and namenode process stops running after machine restart. I tried re-configuring it with clearing htemp folder and formatting namenode. However, this issue still persists, and final solution is to re-install Hadoop on machine