Packages skipped in install for some reason

Gcc

Kernel-devel

Patch

Turn on networking

/etc/sysconfig/network-scripts for ifcfg-eth

I had to change the ONBOOT=no to ONBOOT=yes.

Disable KDUMP

chkconfig kdump off

When running the command "hadoop  jar  hadoop-­‐mapreduce-­‐examples-­‐2.7.1.jar  wordcount   shakespeare/input                                   shakespeare/output" I was getting an error message about an invalid jar. I think somewhere my path is not correct. To get around this I copied the jar from hadoop/share/hadoop/mapreduce to my working directory. No errors after that. Anyone know what environment variable I wiffed on? I would like to avoid copying every jar file we use to my working directory.

Noting a problem I ran into in case others bump into a similar issue with not being able to connect to the web interface via port 50070.

I ran into a problem during the Hadoop setup.  I believe it was my own issue that I created.  When I finished the firewall setup commands I couldn't connect to the web interface from my core OS (port 50070).  Long story short, I then ran the "hdfs namenode -format" command after reading a post on stackoverflow.  This caused more problems and when I restarted services the datanode was no longer there.  When this happens you need to run the following command to remove the old datanode data<?>

[hadoop@localhost datanode]$ rm -Rf /home/hadoop/hadoopdata/hdfs/datanode/\*

Then when I re-ran the start-dfs.sh and start-yarn.sh commands  and JPS showed the datanode again, and I could also connect to the web interface.

You probably wont create this same problem,,, but if your datanode goes missing you might need to run the command above to resolve.

Thanks,