Processing of large files. Do the following tasks and provide the commands you used. No need to provide the sample out files and results.

The Files location: https://drive.google.com/drive/u/1/folders/15TX9w69hJIe-LIp_k6aSVWmZTbiOkdvT

1. Get the line count of the given log file wso2carbon.log
2. Filter the log file for lines having keyword ERROR. Hint: use grep command
3. Get the line count which have the keyword ERROR.
4. Extract all the lines which includes the keyword ERROR and write those lines into separate text file called out1.log.
5. Get the number of DataEndpointAuthenticationException errors appear in the log file?
6. Extract all the log lines between 2019-07-17 09:03:00 and 2019-07-17 09:08:00 and add to a file called out2.log.
7. Get number of errors occurred between the 2019-07-17 09:03:00 - 2019-07-17 09:08:00 time range.
8. Split the wso2carbon.log file into multiple files. Each file should be sized less than 500kb.
9. Extract all distinct IP addresses mentioned in the log file (same IP address should not repeat twice). Hint: use grep command with regex pattern to extract ip addresses and pipe to get unique IPs

Assignment 01 - Answers

Go to the place where wso2carbon.log file is located and open a terninal there and type,

a) wc -l wso2carbon.log

b) grep **ERROR** wso2carbon.log

c) grep **ERROR** wso2carbon.log | wc -l

d) grep ERROR wso2carbon.log > out1.log

e) grep DataEndpointAuthenticationException wso2carbon.log | wc -l

f) sed -n '/2019-07-17 09:03/,/2019-07-17 09:08/p' wso2carbon.log > out1.log

g) sed -n '/2019-07-17 09:03/,/2019-07-17 09:08/p' wso2carbon.log | wc -l

h) split -b 500KB wso2carbon.log

i) grep -o '[0-9]\{1,3\}\.[0-9]\{1,3\}\.[0-9]\{1,3\}\.[0-9]\{1,3\}' wso2carbon.log |sort -u| cut -d ' ' -f 4 | sort -u