

2024.06.25

오전

제어판-네트워크연결-어댑터연결

wifi 와 vmware 이더넷2를 네트워크 브릿지연결

nat로 해도 애랑 통신이 가능하다..

속성들어가면 enternet2과 wife체크되어있는데 wife끄기

worker를 4기기로 줄인다

kafka1_2_3 키고

master의

vim ~/.bashrc

alias spark_start="/opt/spark/sbin/start-all.sh"

alias spark_stop="/opt/spark/sbin/stop-all.sh"

한다음 source ~/.bashrc

3 대 컴퓨터에서 hosts에 아래 주소 확인

이전 사용했던 도메인 이름 있으면 주석처리!!!

namenode, secondnode, datanode1, datanode2 datanode3, datanode4, datanode5,
datanode6, datanode7

master에서

cd \$SPARK_HOME

SPARK -> conf -> vim spark-env.sh

추가하기

export SPARK_WORKER_MEMORY=3g

설정 파일 보내기

scp ./spark-env.sh worker1:/opt/spark/conf/

scp ./spark-env.sh worker2:/opt/spark/conf/

주석할때 중복된 hostname이 있으면 큰일난다..

spark_start로 jps 확인하기

자꾸 주피터연결에서 에러가 떠서 memory를 2g로 줄임

에러뜨는 이유가 kafka와 zookeeper가 켜져있어서 그렇다
무조건 kstop zstop하자!!

[Errno 13] Permission denied: '/home/hadoop/hadoop/bin/hdfs'

이 에러는...

계정충돌이어서 chown으로 옮겨도 보고...

miniconda 설치와 jupyter설치..

환경변수 설치...함수도 넣어주고

>>> conda initialize >>>

**!! Contents within this block are managed by
'conda init' !!**

conda_setup="

(' /home/hadoop/miniconda3/bin/conda"shell.bash"hook'2 > /dev/null)'if[\$? - eq0]; theneval''

conda_setup"

else

if [-f "/home/hadoop/miniconda3/etc/profile.d/conda.sh"]; then

. "/home/hadoop/miniconda3/etc/profile.d/conda.sh"

else

export PATH="/home/hadoop/miniconda3/bin:\$PATH"

fi

fi

unset __conda_setup

<<< conda initialize <<<

결국 네트워크로..브리지로 끊자

서버로 쓰기로했던 노트북을 spark연결중

log를 보기위해서

/opt/spark/logs에서 tail -f로 파일명

결론을보자

cd sbin/에 가서 ./start-worker spark://client:7077 기존의 프로세스 죽이고 간다

client:8888에 들어간다

ip:port 디코에 있는거치고 비밀번호치면
병렬처리할수있다