win R과 리눅스 Hive연결

winJava+winR+linuxHive

--> Thrift 하이브랑 외부랑 연결하는게 thrift이고 포트번호 10000번을 쓴다.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

리눅스에서 jps해서 6개 안뜨면 start-all.sh 가동

10000번 포트 잡기

#hive -- service hiveserver2 --hiveconf hive.server2.thrift.port=10000 --hiveconf hive.root.logger=INFO,console 가동

~~~뜨고 다음과 같이 끝나면 정상

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

SHUTDOWN\_MSG: Shutting down HiveServer2 at hadoopserver/192.168.111.100

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

10000번 포트 잘 잡혔는지 확인

#netstat -nl | grep 10000

안되면 # . /etc/profile 한다음 다시 해보기

방화벽을 해제해야한다.

#systemctl stop firewalld → firewqlldemon

→ 윈도우에서 거침없이 접근 가능하단 것

firewall-cmd --permanent --zone=public --add-port=10000/tcp

→ 방화벽 해제 하지 말고 10000번 포트 개방

firewall-cmd --reload

→ 방화벽 설정 다시 불러와서 적용

java\_hive\_lib안에는 linuxjava-hive에서 사용한 jar들 모음이다.

add external jar로 삽입

그런데 이 jar들 윈도우에서도 필요 : winjava-linuxhive를 위해 필요하기 때문

윈오두 R에다음 내용 설치

install.packages(“RJDBC”, dep=TRUE)

install.packages(‘rJava’)

Rserve 패키지를 설치

winR에서 linuxHive와 연동 (이때 사용되는게 6311포트)

R 에서 함수 만들기

rhive2 <-function(){

library(RJDBC);

library(DBI);

library(rJava);

#hive\_lib<-"/usr/local.hive/lib";

#hive\_lib <- ''

hive\_lib <- 'C:/lib/java\_hive\_lib';

.jinit();

.jaddClassPath(dir(hive\_lib,full.names=TRUE));

.jclassPath();

drv=JDBC(driverClass='org.apache.hive.jdbc.HiveDriver','hive-jdbc-1.0.1.jar');

conn <- dbConnect(drv,"jdbc:hive2://192.168.111.100:10000","hive","111111");

query <- "select \* from hdi2";

hdi <- dbGetQuery(conn,query);

dbDisconnect(conn);

return(hdi);

}

hdi2 확인하려면

hive> dfs -ls /user/hive/warehouse; → 하이브안에서는 hadoop 안붙이고 쓴다

해보면

drwxr-xr-x - root supergroup 0 2018-03-13 15:38 /user/hive/warehouse/hdi2

이클립스 bulid path 라이브러이

C:\Program Files\R\R-3.4.3\library\Rserve\java 경로에 있는 jar 파일 두개 추가