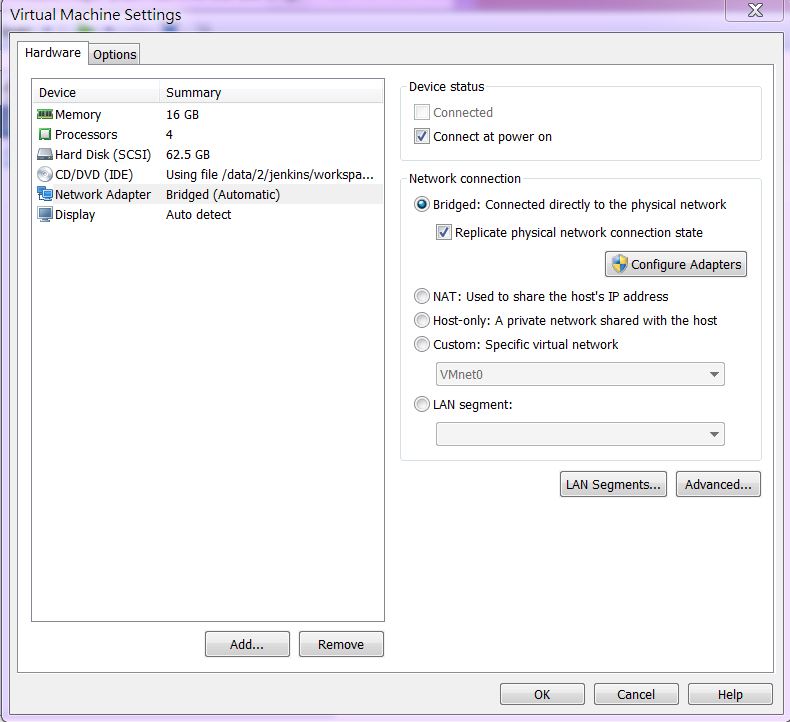
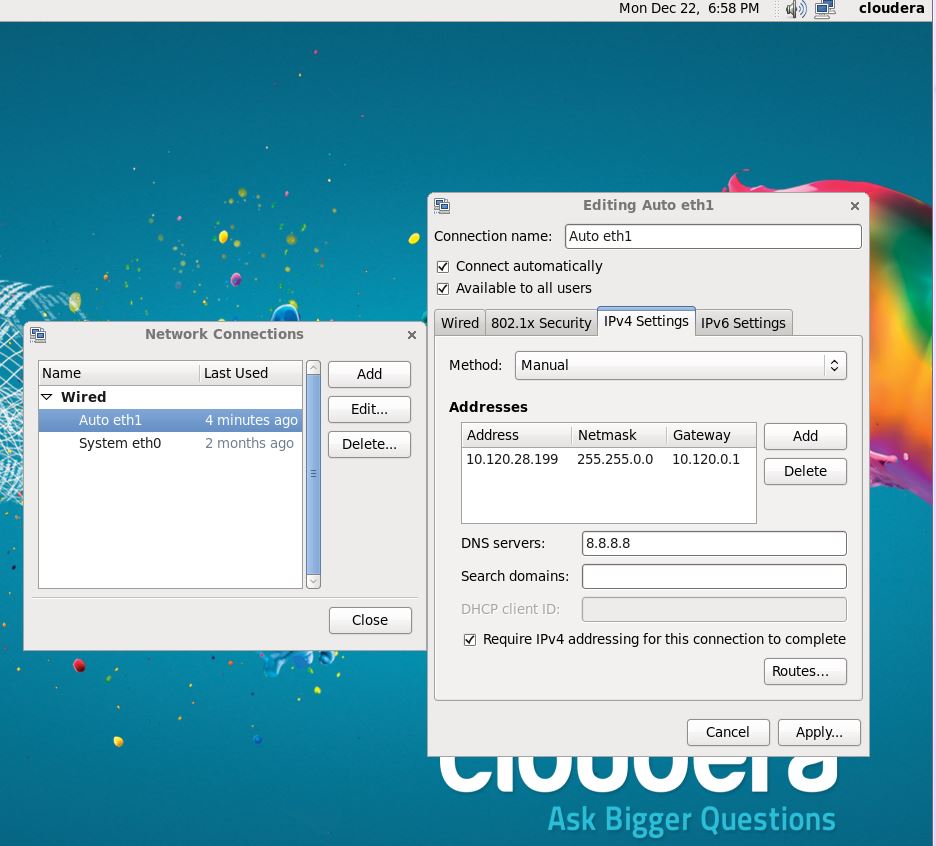
* 1. **(事前下載) Download for VMWare**

<http://www.cloudera.com/content/cloudera/en/downloads/quickstart_vms/cdh-5-3-x.html>

* 1. **(事前設定)** 開一個新VM, Network Adapter Bridged(Automatic)



* 1. **(事前設定)** VM內部設定, IPv4同windows網卡, Address末2碼不能相同



**2-1. Linux (RRO8.0.1) centos 6版本(全程以root身分進行)**

[cloudera@quickstart ~]$ su

[root@quickstart cloudera]# sudo yum clean all

# yum list make gcc gcc-gfortran

# yum install gcc-c++.x86\_64

# wget –no-check-certificate <http://mran.revolutionanalytics.com/install/RRO-8.0.1-Beta-el6.x86_64.rpm>

# yum --nogpgcheck localinstall RRO-8.0.1-Beta-el6.x86\_64.rpm

**2-2. 套件下載及安裝**

# which hadoop

# export HADOOP\_CMD=/usr/bin/hadoop

# locate streaming | grep jar | more

# export HADOOP\_STREAMING=/usr/lib/hadoop-0.20-mapreduce/contrib/streaming/hadoop-streaming.jar

# echo $JAVA\_HOME

# export JAVA\_HOME=/usr/java/jdk1.7.0\_67-cloudera

# R CMD javareconf

# sudo R

> install.packages(c("codetools", "Rcpp", "RJSONIO", "bitops", "digest", "functional", "stringr", "plyr", "reshape2", "rJava", "caTools"))

> q()

**2-3. Download rmr2**

# wget –no-check-certificate http://goo.gl/Y5ytsm

# R CMD INSTALL Y5ytsm

**2-4.** **Download rhdfs**

# wget –no-check-certificate https://github.com/RevolutionAnalytics/rhdfs/blob/master/build/rhdfs\_1.0.8.tar.gz?raw=true

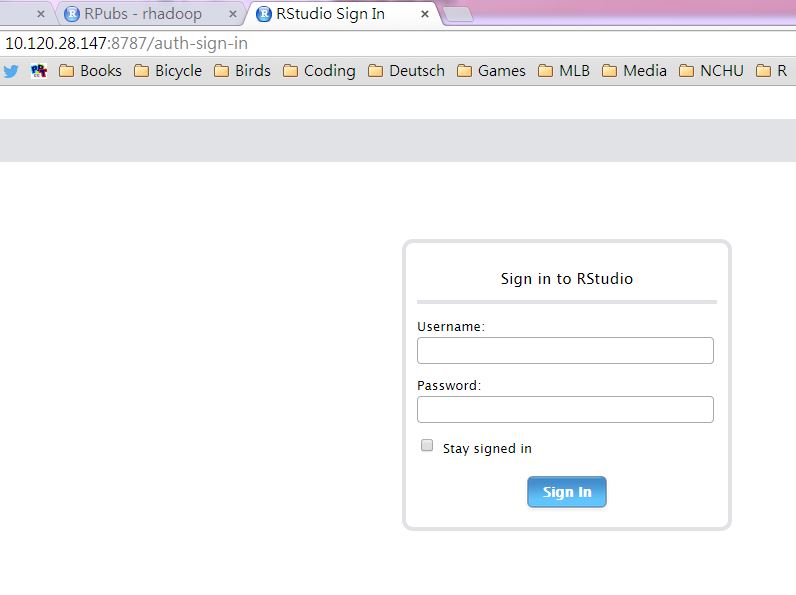
# R CMD INSTALL rhdfs\_1.0.8.tar.gz\?raw\=true

# 2-5. Download RStudio Server - RedHat/CentOS

<http://www.rstudio.com/products/rstudio/download-server/>

# sudo yum install openssl098e # Required only for RedHat/CentOS 6 and 7  
# wget http://download2.rstudio.org/rstudio-server-0.98.1091-x86\_64.rpm  
# sudo yum install --nogpgcheck rstudio-server-0.98.1091-x86\_64.rpm

# sudo rstudio-server restart



**2-6. 執行rmr2前的系統設定(建議用RScript儲存，每次開啟Rsession時需執行)**

Sys.setenv(HADOOP\_CMD="/usr/bin/hadoop")

Sys.setenv(HADOOP\_STREAMING="/usr/lib/hadoop-0.20-mapreduce/contrib/streaming/hadoop-streaming.jar")

Sys.setenv(JAVA\_HOME="/usr/java/jdk1.7.0\_67-cloudera")

library(rJava)

library(rhdfs)

library(rmr2)

backend.parameters =

list(

hadoop =

list(

D = "mapred.map.child.ulimit=2097152",

D = "mapred.reduce.child.ulimit=2097152",

D = "mapred.tasktracker.map.tasks.maximum=1",

D = "mapred.tasktracker.reduce.tasks.maximum=1"))

#-- test rhdfs

hdfs.init()

hdfs.ls("/")

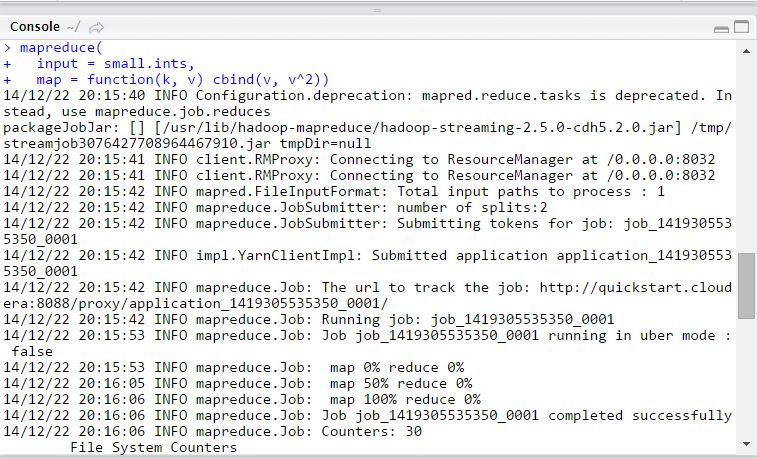
#-- test rmr

small.ints = to.dfs(1:1000)

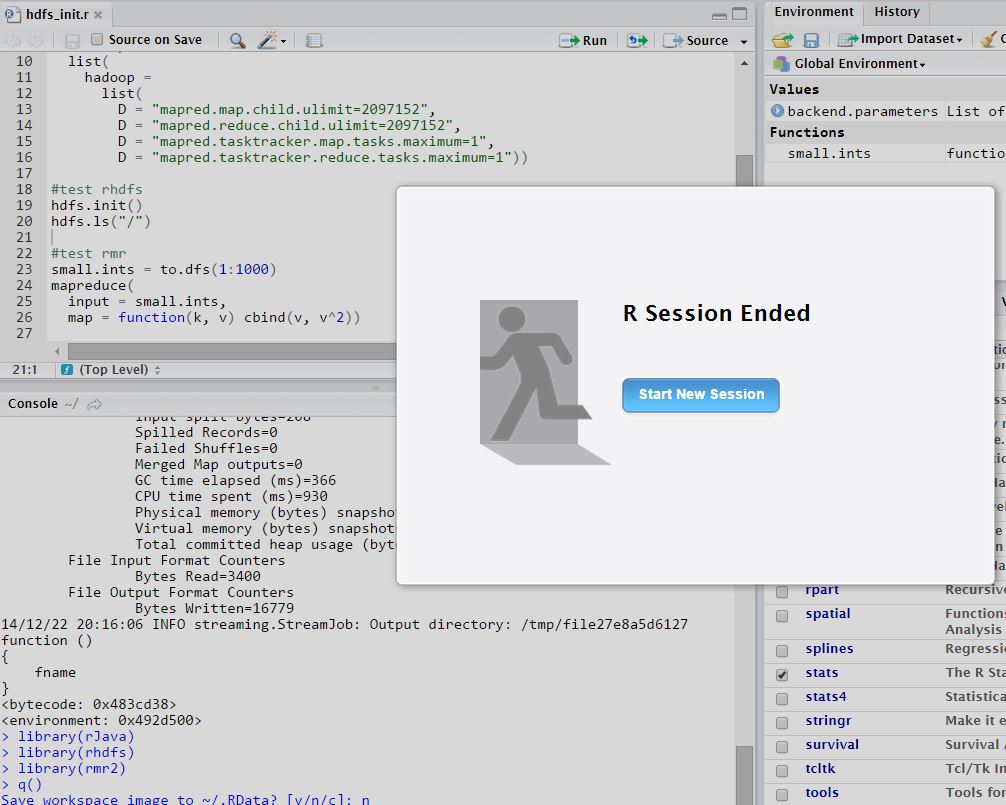
mapreduce(

input = small.ints,

map = function(k, v) cbind(v, v^2))



> q()



**3-1 SHINY-SERVER**

<http://www.rstudio.com/products/shiny/download-server/>

選取RedHat/CentOS

# sudo R

> install.packages('shiny', repos='http://cran.rstudio.com/')

> q()

**Download and Install**

# wget http://download3.rstudio.org/centos-5.9/x86\_64/shiny-server-1.2.3.368-x86\_64.rpm  
# sudo yum install --nogpgcheck shiny-server-1.2.3.368-x86\_64.rpm

# sudo rm -r starOpeningWeekendGross

# sudo cp -r /home/cloudera/Desktop/R/starOpeningWeekendGross /srv/shiny-server/sample-apps

**Reference:**

<http://rpubs.com/ywchiu/25570>

<https://support.rstudio.com/hc/en-us/articles/200552306-Getting-Started>

# RStudio-server

<http://10.120.28.xxx:8787/>

# shiny-server

<http://10.120.28.xxx:3838/>

version 20150111