싱글 노드 모드

환경설정

가상머신 설치

➡ https://www.virtualbox.org/

VirtualBox

Welcome to VirtualBox.org!

irtualBox is a powerful x86 and AMD64/Intel64 virtualization product for enterprise as well as home use. Not only is VirtualBox an extre
or enterprise customers, it is also the only professional solution that is freely available as Open Source Software under the terms of the
ee "About VirtualBox" for an introduction.

resently, VirtualBox runs on Windows, Linux, Macintosh, and Solaris hosts and supports a large number of guest operating systems incl
000, XP, Server 2003, Vista, Windows 7, Windows 8, Windows 10), DOS/Windows 3.x, Linux (2.4, 2.6, 3.x and 4.x), Solaris and OpenS

irtualBox is being actively developed with frequent releases and has an ever growing list of features, supported guest operating system
ommunity effort backed by a dedicated company: everyone is encouraged to contribute while Oracle ensures the product always meets
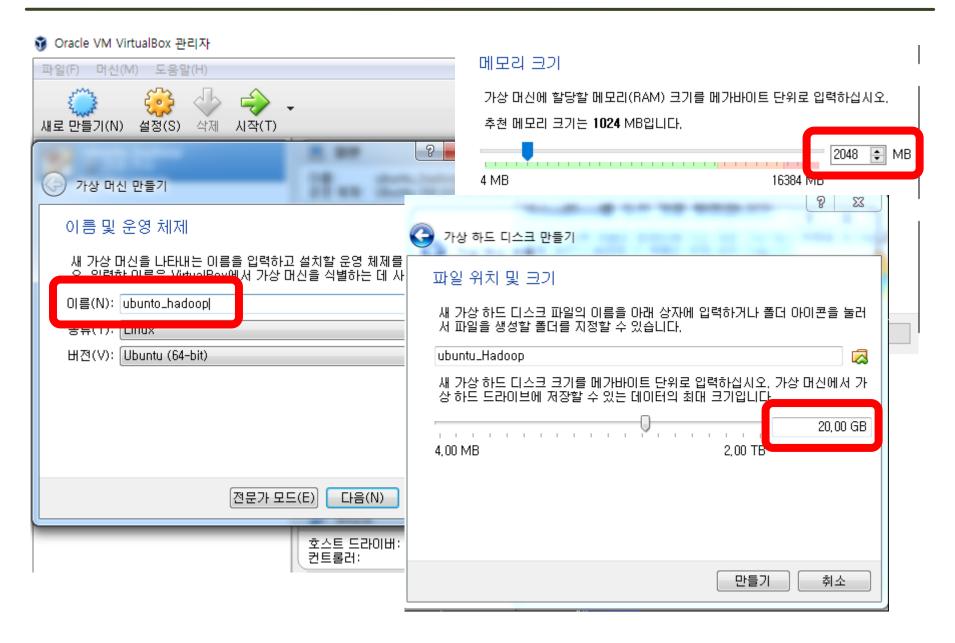
Download VirtualBox 6.0
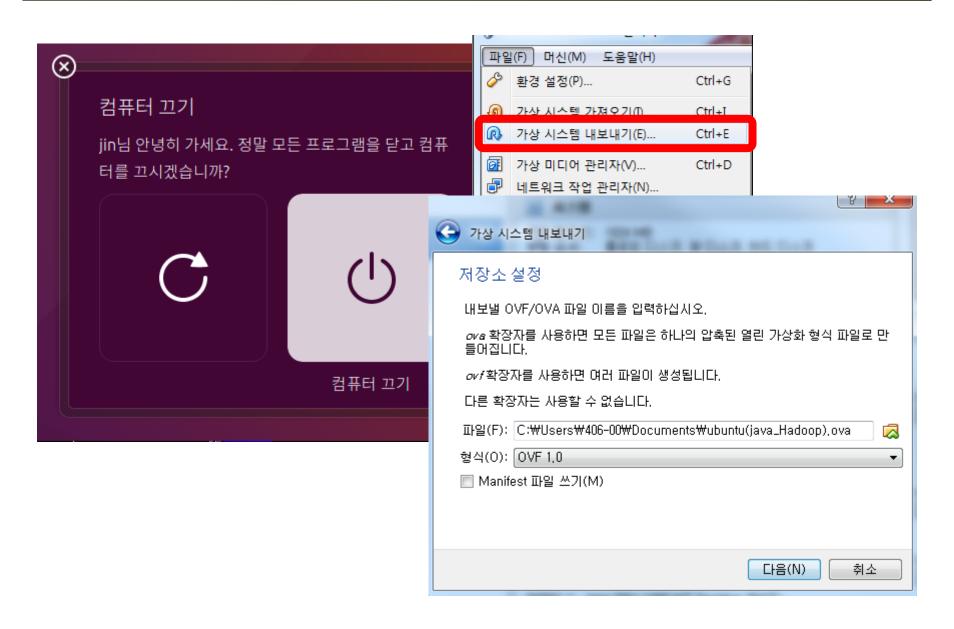
lot picks:

우분투 설치
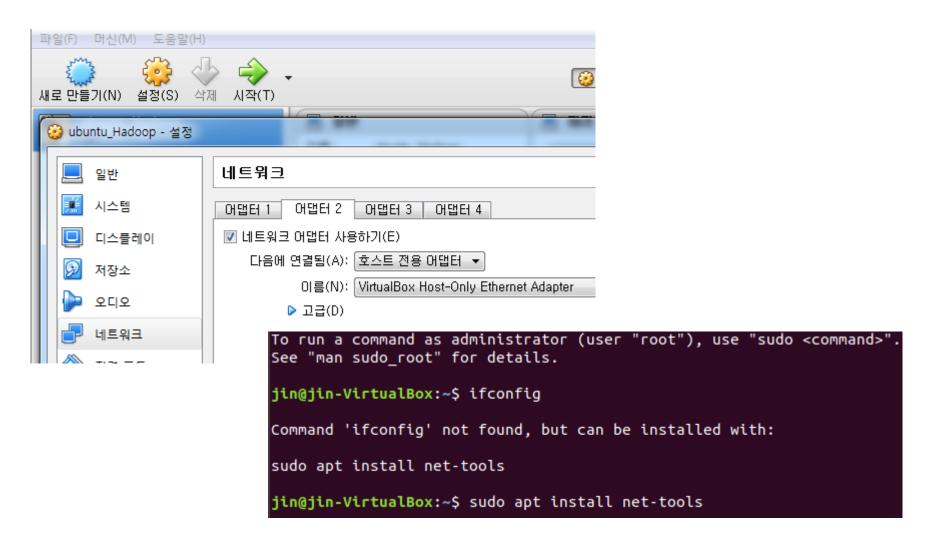
➡ https://ubuntu.com/

우분투 설치

복제하기

기본 환경 설정
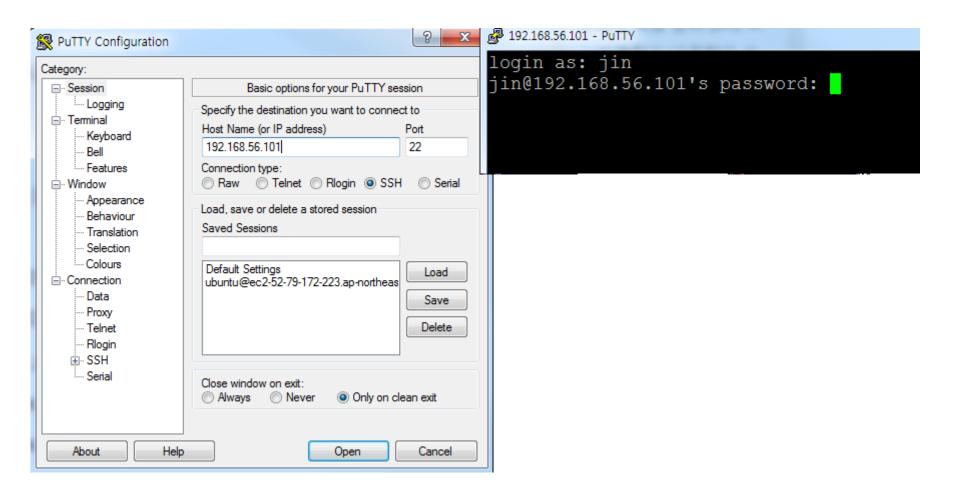
네트워크 설정

유틸 설치

```
jin@jin-VirtualBox:~$ sudo apt-get install ssh
패키지 목록을 읽는 중입니다... 완료
의존성 트리를 만드는 중입니다
상태 정보를 읽는 중입니다... 완료
The following additional packages will be installed:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
제안하는 패키지:
  ssh-askpass rssh molly-guard monkeysphere
다음 새 패키지를 설치할 것입니다:
```

```
jin@jin-VirtualBox:~$ sudo apt-get install vim -y
```

Putty 접속

JDK 설치

```
jin@jin-VirtualBox:~$ sudo apt-get update

jin@jin-VirtualBox:~$ java -version

jin@jin-VirtualBox:~$ sudo apt-get install default-jdk -y

jin@jin-VirtualBox:~$ which java
/usr/bin/java
jin@jin-VirtualBox:~$ readlink -f /usr/bin/java
/usr/lib/jvm/java-11-openjdk-amd64/bin/java

jin@jin-VirtualBox:~$ sudo vim /etc/profile
```

export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64

```
jin@jin-VirtualBox:~$ source /etc/profile
```

하둡 설치

다운로드 및 설치

```
jin@jin-VirtualBox:~$ wget http://apache.mirror.cdnetworks.com/hadoop/common/had
oop-3.2.0/hadoop-3.2.0.tar.gz
```

```
jin@jin-VirtualBox:~$ tar zxvf hadoop-2.9.0.tar.gz
```

```
jin@jin-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop
```

```
jin@jin-VirtualBox:~$ ls
examples.desktop  hadoop-2.9.0.tar.gz  다운로드  바탕화면  사진
hadoop-2.9.0      공개                 문서      비디오    음악
jin@jin-VirtualBox:~$ cd hadoop-2.9.0/
jin@jin-VirtualBox:~/hadoop-2.9.0$ sudo mv * /usr/local/hadoop/
```

```
jin@jin-VirtualBox:~/hadoop-2.9.0$ ls
jin@jin-VirtualBox:~/hadoop-2.9.0$ cd ..
jin@jin-VirtualBox:~$ rm -rf ha*
jin@jin-VirtualBox:~$ ls
examples.desktop  공개  다운로드  문서
```

사용자 등록 및 하둡 소유권 변경

```
jin@jin-VirtualBox:~$ sudo addgroup hadoop
그룹 `hadoop' (GID 1001) 추가 ...
완료.
jin@jin-VirtualBox:~$ sudo adduser --ingroup hadoop manager
'manager' 사용자를 추가 중...
새 사용자 'manager' (1001) 을(를) 그룹 'hadoop' (으)로 추가 ...
'/home/manager' 홈 디렉터리를 생성하는 중...
'/etc/skel'에서 파일들을 복사하는 중...
새 UNIX 암호 입력:
새 UNIX 암호 재입력:
```

```
jin@jin-VirtualBox:~$ groups manager
manager : hadoop
jin@jin-VirtualBox:~$ sudo adduser manager sudo
[sudo] password for jin:
'manager' 사용자를 'sudo' 그룹에 추가 중...
사용자 manager을(를) sudo 그룹에 등록 중
완료.
```

```
jin@jin-VirtualBox:~$ sudo chown -R manager:hadoop /usr/local/hadoop/
jin@jin-VirtualBox:~$ ls -l /usr/local
합계 36
drwxr-xr-x 2 root     root   4096  8월  1  2017 bin
drwxr-xr-x 2 root     root   4096  8월  1  2017 etc
drwxr-xr-x 2 root     root   4096  8월  1  2017 games
drwxr-xr-x 9 manager hadoop 4096  2월  9 16:20 hadoop
```

Static IP 설정

```
jin@jin-VirtualBox:~$ sudo vim /etc/network/interfaces
```

```
# interfaces(5) file used by ifup(8) and ifdown(8)
auto lo
iface lo inet loopback

auto enp0s8
iface enp0s8 inet static
address 192.168.56.101
netmask 255.255.255.0
gateway 192.168.56.1
dns-nameservers 168.126.63.1
```

```
jin@jin-VirtualBox:~$ sudo /etc/init.d/networking restart
[ ok ] Restarting networking (via systemctl): networking.service.
```

~/.bashrc 수정

```
manager@master:~$ sudo vim .bashrc
```

```
#HADOOP VARIABLES START
export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
```

환경 변수 설정

```
$ sudo vim /usr/local/hadoop/etc/hadoop/hadoop-env.sh

# The directory where pid files are stored. /tmp by
# NOTE: this should be set to a directory that can o
#       the user that will run the hadoop daemons.
#       potential for a symlink attack.
export HADOOP_PID_DIR=${HADOOP_PID_DIR}
export HADOOP_SECURE_DN_PID_DIR=${HADOOP_PID_DIR}

# A string representing this instance of hadoop. $US
export HADOOP_IDENT_STRING=$USER

export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
```

core-site

```
$ sudo mkdir -p /usr/local/hadoop/tmp

$ sudo chown -R manager:hadoop /usr/local/hadoop/tmp/
$ sudo vim /usr/local/hadoop/etc/hadoop/core-site.xml
```

```xml
<configuration>
        <property>
                <name>hadoop.tmp.dir</name>
                <value>/usr/local/hadoop/tmp</value>
        </property>

        <property>
                <name>fs.default.name</name>
                <value>hdfs://localhost:62350</value>
        </property>
</configuration>
```

mapred-site

```
manager@master:~$ cp /usr/local/hadoop/etc/hadoop/
mapred-site.xml.template /usr/local/hadoop/etc/had
oop/mapred-site.xml
```

```
manager@master:~$ sudo vim /usr/local/hadoop/etc/
hadoop/mapred-site.xml
```

```
<configuration>
        <property>
                <name>mapred.job.tracker</name>
                <value>localhost:62351</value>
        </property>
</configuration>
```

hdfs-site

```
$ sudo mkdir -p /usr/local/hadoop/hdfs/namenode
$ sudo mkdir -p /usr/local/hadoop/hdfs/datanode
$ sudo chown -R manager:hadoop /usr/local/hadoop
```

```
manager@master:~$ sudo vim /usr/local/hadoop/etc/
hadoop/hdfs-site.xml
```

```xml
<property>
        <name>dfs.replication</name>
        <value>1</value>
</property>
<property>
        <name>dfs.namenode.name.dir</name>
        <value>file:/usr/local/hadoop/hdfs/namenode</value>
</property>
<property>
        <name>dfs.datanode.data.dir</name>
        <value>file:/usr/local/hadoop/hdfs/datanode</value>
</property>
```

인증키 생성

```
manager@jin-VirtualBox:~$ ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/manager/.ssh/id_rsa):
Created directory '/home/manager/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/manager/.ssh/id_rsa.
Your public key has been saved in /home/manager/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:RDTlpfu/MtCvTeflGfJdu+0hEMWCoCwNL9UeDlCe5ZA manager@jin-Vi
The key's randomart image is:
+---[RSA 2048]----+
|    oo+o+=.o o.   |
|     BE*+ + +..   |
|    o B+.o o..    |
|     o  +   ..    |
|        S .o      |
|         ..o      |
|          ..+.o+| |
|           o+=+X| |
|           .++BB| |
+----[SHA256]-----+
```

인증키 배포

```
manager@jin-VirtualBox:~$ ssh-copy-id -i .ssh/id_rsa.pub manager@localhost
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: ".ssh/id_rsa.pub"
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:YjOvkT28J3USbt4j4LtaBwpeHxdWEJIvEkqe6Hc5Gys.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
manager@localhost's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'manager@localhost'"
and check to make sure that only the key(s) you wanted were added.
```

```
manager@jin-VirtualBox:~$ sudo reboot
```

하둡 실행

```
manager@jin-VirtualBox:~$ hadoop namenode -format

manager@master:~$ start-dfs.sh


manager@jin-VirtualBox:~$ jps
1841 SecondaryNameNode
1593 DataNode
1994 Jps
1437 NameNode
```

동작 확인

FileZilla 설치

➡ https://filezilla-project.org/

분석 폴더 생성 및 이동

```
$ sudo mkdir -p /usr/local/hadoop/jar
$ sudo chown -R manager:hadoop /usr/local/hadoop/jar/
```

리모트 사이트: /usr/local/hadoop/jar

- hadoop
  - ? bin
  - ? etc
  - ? hdfs
  - ? include
  - jar
  - ? lib
  - ? libexec
  - ? logs

파일명	크기	파일 유형
..		
CHANGES.txt	446,615	텍스트 문서
WordCount.jar	3,215	ALZip JAR ...

```
$ hadoop fs -mkdir -p /wordcount/input
manager@master:~$ hadoop fs -copyFromLocal /usr/local/
hadoop/jar/*.txt /wordcount/input/

manager@master:~$ hadoop jar /usr/local/hadoop/jar/
WordCount.jar com.care.WordCount.WordCount /wordcou
nt/input /wordcount/output

manager@master:~$ hadoop fs -ls /wordcount/output

manager@master:~$ hadoop fs -cat /wordcount/output/
part-00000
```