Virtual box – Ubuntu –Hadoop 설치

고정 ip / Ssh 설치

```
ubuntu-VirtualBox:~$ service
Uservice < option > | --status-all | [ service_name [ command | --full-resta
rt ]]
ubuntu@ubuntu-VirtualBox:~$ service ssh
* Usage: /etc/init.d/ssh {start|stop|reload|force-reload|restart|try-restart|stat
us}
ubuntu@ubuntu-VirtualBox:~$ service ssh status
● ssh.service - OpenBSD Secure Shell server
     Loaded: loaded (/lib/systemd/system/ssh.service; enabled; vendor preset: ena>
     Active: active (running) since Thu 2021-05-27 12:45:07 KST; 2min 11s ago
       Docs: man:sshd(8)
             man:sshd_config(5)
    Process: 667 ExecStartPre=/usr/sbin/sshd -t (code=exited, status=0/SUCCESS)
   Main PID: 688 (sshd)
      Tasks: 1 (limit: 9256)
     Memory: 2.5M
     CGroup: /system.slice/ssh.service
             └─688 sshd: /usr/sbin/sshd -D [listener] 0 of 10-100 startups

5월 27 12:45:06 ubuntu-VirtualBox systemd[1]: Starting OpenBSD Secure Shell serv>
5월 27 12:45:07 ubuntu-VirtualBox sshd[688]: Server listening on 0.0.0.0 port 22.
5월 27 12:45:07 ubuntu-VirtualBox sshd[688]: Server listening on :: port 22.
5월 27 12:45:07 ubuntu-VirtualBox systemd[1]: Started OpenBSD Secure Shell serve>
lines 1-16/16 (END)
```

```
ubuntu@ubuntu-VirtualBox:~$ ls
 hadoop-2.9.1    hadoop-2.9.1.tar.gz   '스크린샷, 2021-05-27 11-58-18.png'
ubuntu@ubuntu-VirtualBox:~$ cd hadoop-2.9.1/
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ pwd
/home/ubuntu/hadoop-2.9.1
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ sudo mkdir /usr/l
lib/    lib32/  lib64/  libexec/ libx32/  local/
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ sudo mkdir /usr/l
lib/    lib32/  lib64/  libexec/ libx32/  local/
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ sudo mkdir /usr/l
lib/    lib32/  lib64/  libexec/ libx32/  local/
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ sudo mkdir /usr/local/hadoop/
[sudo] ubuntu의 암호:
mkdir: `/usr/local/hadoop/' 디렉토리를 만들 수 없습니다: 파일이 있습니다
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ sudo mv * /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ ls
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ cd ../
ubuntu@ubuntu-VirtualBox:~$ cd /usr/l
lib/    lib32/  lib64/  libexec/ libx32/  local/
ubuntu@ubuntu-VirtualBox:~$ cd /usr/local/hadoop/
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ ls
LICENSE.txt   bin          etc      libexec  snap    문서     사진
NOTICE.txt    eclipse      include  sbin     공개    바탕화면  음악
README.txt    eclipse-workspace  lib  share  다운로드  비디오  템플릿
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$
```
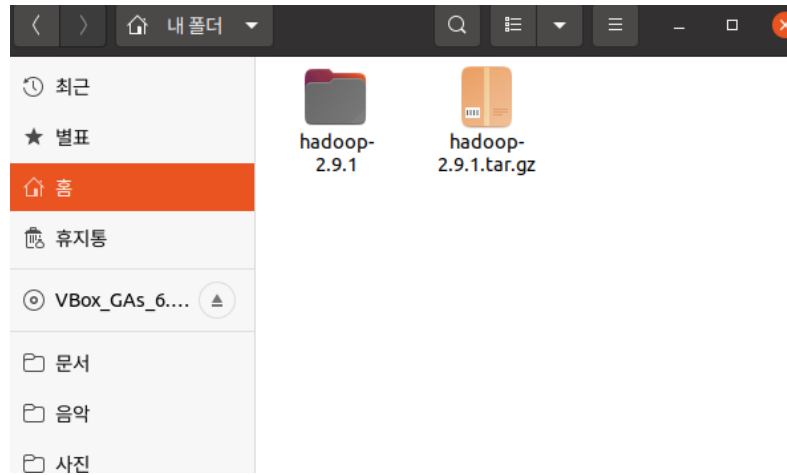
Hadoop 압축 해제


```
ubuntu@ubuntu-VirtualBox:~$ sudo mv * /usr/local/hadoop
```

```
ubuntu@ubuntu-VirtualBox:~$ sudo mv * /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:~$ cd /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ ls
eclipse             hadoop-2.9.1        snap   다운로드  바탕화면  사진   템플릿
eclipse-workspace   hadoop-2.9.1.tar.gz  공개   문서      비디오    음악
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$
```

```
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ ls
LICENSE.txt  NOTICE.txt  README.txt  bin  etc  include  lib  libexec  sb
in  share
```

```
ubuntu@ubuntu-VirtualBox:~/hadoop-2.9.1$ cd ~
ubuntu@ubuntu-VirtualBox:~$ ls -al
합계 353020
drwxr-xr-x 24 ubuntu  ubuntu   4096  5월 28 15:15 .
drwxr-xr-x  4 root    root     4096  5월 28 15:01 ..
-rw-------  1 ubuntu  ubuntu    583  5월 28 15:16 .bash_history
-rw-r--r--  1 ubuntu  ubuntu    220  5월 17 13:00 .bash_logout
-rw-r--r--  1 ubuntu  ubuntu   3771  5월 17 13:00 .bashrc
```

Hadoop 폴더 패스 설정

```
ubuntu@ubuntu-VirtualBox:~$ sido gedit .bashrc
```

```
1 # ~/.bashrc: executed by bash(1) for non-login shells.
2 # see /usr/share/doc/bash/examples/startup-files (in the package bash-doc)
3 # for examples
4 export HADOOP_HOME=/usr/local/hadoop
5
6 export PATH=$PATH:$HADOOP_HOME/bin
7
8 export PATH=$PATH:$HADOOP_HOME/sbin
9
10 export HADOOP_MAPRED_HOME=$HADOOP_HOME
11
12 export HADOOP_COMMON_HOME=$HADOOP_HOME
13
14 export HADOOP_HDFS_HOME=$HADOOP_HOME
15
16 export YARN_HOME=$HADOOP_HOME
17
18 export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
19
20 export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"
21
22 # If not running interactively, don't do anything
23 case $- in
24    *i*) ;;
25      *) return;;
```

```
ubuntu@ubuntu-VirtualBox:~$ source .bashrc
ubuntu@ubuntu-VirtualBox:~$ hadoop
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]
  CLASSNAME            run the class named CLASSNAME
 or
  where COMMAND is one of:
  fs                   run a generic filesystem user client
  version              print the version
  jar <jar>            run a jar file
                       note: please use "yarn jar" to launch
                       YARN applications, not this command.
  checknative [-a|-h]  check native hadoop and compression libraries availability
  distcp <srcurl> <desturl> copy file or directories recursively
  archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archive
  classpath            prints the class path needed to get the
                       Hadoop jar and the required libraries
  credential           interact with credential providers
  daemonlog            get/set the log level for each daemon
  trace                view and modify Hadoop tracing settings

Most commands print help when invoked w/o parameters.
```

Xml 설정 파일

이름	크기	수정 시각
`</>` capacity-scheduler.xml	7.9 kB	2018년 4월 16일
configuration.xsl	1.3 kB	2018년 4월 16일
container-executor.cfg	1.2 kB	2018년 4월 16일
`</>` core-site.xml	774 바이트	2018년 4월 16일
hadoop-env.cmd	4.1 kB	2018년 4월 16일
hadoop-env.sh	5.0 kB	2018년 4월 16일
hadoop-metrics.properties	2.5 kB	2018년 4월 16일
hadoop-metrics2.properties	2.6 kB	2018년 4월 16일
`</>` hadoop-policy.xml	10.2 kB	2018년 4월 16일
`</>` hdfs-site.xml	775 바이트	2018년 4월 16일
httpfs-env.sh	2.2 kB	2018년 4월 16일

```xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

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

```xml
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<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>
<property>
    <name>dfs.http.address</name>
    <value>localhost:50070</value>
</property>
<property>
    <name>dfs.secondary.http.address</name>
    <value>localhost:50090</value>
</property>
</configuration>
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<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>
```

hadoop-env.sh — /usr/local/hadoop/etc/hadoop

```bash
# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements.  See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership.  The ASF licenses this file
# to you under the Apache License, Version 2.0 (the
# "License"); you may not use this file except in compliance
# with the License.  You may obtain a copy of the License at
#
#     http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME.  All others are
# optional.  When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use.
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
```
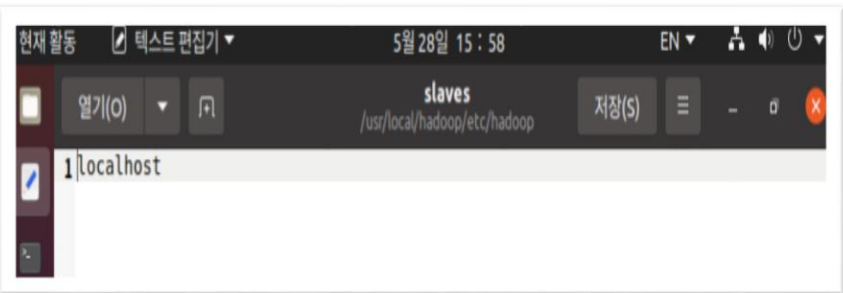
Hadoop 그룹 내 Manager 계정 생성, 권한 부여, 폴더 신규 개설

```
ubuntu@ubuntu-VirtualBox:~$ sudo chown -R manager:hadoop /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:~$ sudo mv * /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:~$
```

```
ubuntu@ubuntu-VirtualBox:~$ groups manager
manager : manager hadoop
ubuntu@ubuntu-VirtualBox:~$ ls -l /usr/local
합계 36
drwxr-xr-x  2 root    root   4096  2월 10 03:47 bin
drwxr-xr-x  2 root    root   4096  2월 10 03:47 etc
drwxr-xr-x  2 root    root   4096  2월 10 03:47 games
drwxr-xr-x 22 manager hadoop 4096  5월 28 15:54 hadoop
drwxr-xr-x  2 root    root   4096  2월 10 03:47 include
drwxr-xr-x  3 root    root   4096  2월 10 03:47 lib
lrwxrwxrwx  1 root    root      9  5월 17 12:54 man -> share/man
drwxr-xr-x  2 root    root   4096  2월 10 03:47 sbin
drwxr-xr-x  7 root    root   4096  2월 10 03:51 share
drwxr-xr-x  2 root    root   4096  2월 10 03:47 src
ubuntu@ubuntu-VirtualBox:~$
```

```
ubuntu@ubuntu-VirtualBox:~$ sudo gedit /usr/local/hadoop/etc/hadoop/slaves
```

현재 활동 | 텍스트 편집기 ▼ | 5월 28일 15 : 58 | EN ▼

열기(O) — slaves /usr/local/hadoop/etc/hadoop — 저장(S)

```
1 localhost
```

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ sudo chown -R ubuntu:hadoop /usr/local/hadoop
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ mkdir hdfs
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ ls
LICENSE.txt  eclipse-installer   hdfs      share    바탕화면
NOTICE.txt   eclipse-workspace   include   snap     비디오
README.txt   etc                 lib       공개     사진
bin          hadoop-2.9.1        libexec   다운로드  음악
eclipse      hadoop-2.9.1.tar.gz sbin      문서     템플릿
```

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ mkdir tmp
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop$ cd hdfs/
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ mkdir datanode namenode
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ ls
datanode  namenode
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$
```

Name node 포맷 후 start-all.sh 실행

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ hadoop namenode -format
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

21/05/28 16:06:28 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = ubuntu-VirtualBox/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.9.1
STARTUP_MSG:   classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/
doop/common/lib/commons-net-3.1.jar:/usr/local/hadoop/share/hadoop/common/
```

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
21/05/28 16:07:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library
 for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:i9L1VJdSF3zq/CU34SsHgpPJ7gukw9RApZ6/PSGIAo4.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
localhost: Warning: Permanently added 'localhost' (ECDSA) to the list of known hos
ts.
ubuntu@localhost's password:
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-ubuntu-name
node-ubuntu-VirtualBox.out
```

Jps 실행 : 6개 동작중 확인

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ jps
33665 DataNode
34294 NodeManager
33495 NameNode
34408 Jps
33979 ResourceManager
33855 SecondaryNameNode
```

```
ubuntu@ubuntu-VirtualBox:/usr/local/hadoop/hdfs$ stop-all.sh
This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh
21/05/28 16:10:05 WARN util.NativeCodeLoader: Unable to load native-hadoop library
 for your platform... using builtin-java classes where applicable
Stopping namenodes on [localhost]
ubuntu@localhost's password:
localhost: stopping namenode
ubuntu@localhost's password:
localhost: stopping datanode
Stopping secondary namenodes [localhost]
ubuntu@localhost's password:
localhost: stopping secondarynamenode
21/05/28 16:10:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library
 for your platform... using builtin-java classes where applicable
stopping yarn daemons
stopping resourcemanager
ubuntu@localhost's password:
localhost: stopping nodemanager
localhost: nodemanager did not stop gracefully after 5 seconds: killing with kill
-9
no proxyserver to stop
```