

Window10에서

# Hadoop 설치 & 환경설정

작성자 : 김진성

# Hadoop Download

<https://www.barik.net/archive/2015/01/19/172716/>

Window에서 사용할 수 있도록

빌드 된 Hadoop 버전


Scala-IDE에서 Spark 셋팅과 Wo... | 하둡 설치 및 실행 - Windows... | Hadoop 2.6.0 Windows 64-bit... | +

← → ↻ barik.net/archive/2015/01/19/172716/

Titus Barik *I build tools for engineers and data scientists.* About CV GitHub @barik News Archives

## Hadoop 2.6.0 Windows 64-bit Binaries

Posted on January 19, 2015 by Titus Barik




The official release of [Apache Hadoop 2.6.0](#) does not include the required binaries (e.g., `winutils.exe`) necessary to run hadoop. In order to use Hadoop on Windows, it must be compiled from source. This takes a bit of effort, so I've provided a pre-compiled, unofficial distribution below:

- [hadoop-2.6.0.tar.gz](#)  
SHA1: 205b235d77213b958f126647241a5092845a01f8
- [hadoop-dist-2.6.0-javadoc.jar](#)  
SHA1: cbccbc5d7eeb03261e533cbe0b1b367fce83b181
- [hadoop-2.6.0-src.tar.gz](#)  
SHA1: 4a1dfa9bd34d5efb7f2a0cd4dcf03db3eab46a5d

I compiled the source using:

- [Apache Maven 3.2.5](#)

About



You've reached the site of **Titus Barik**. I'm a researcher at Microsoft. My research interests are in HCI and software engineering.

Contact Information

Personal: [titus@barik.net](mailto:titus@barik.net)  
Work: [titus.barik@microsoft.com](mailto:titus.barik@microsoft.com)

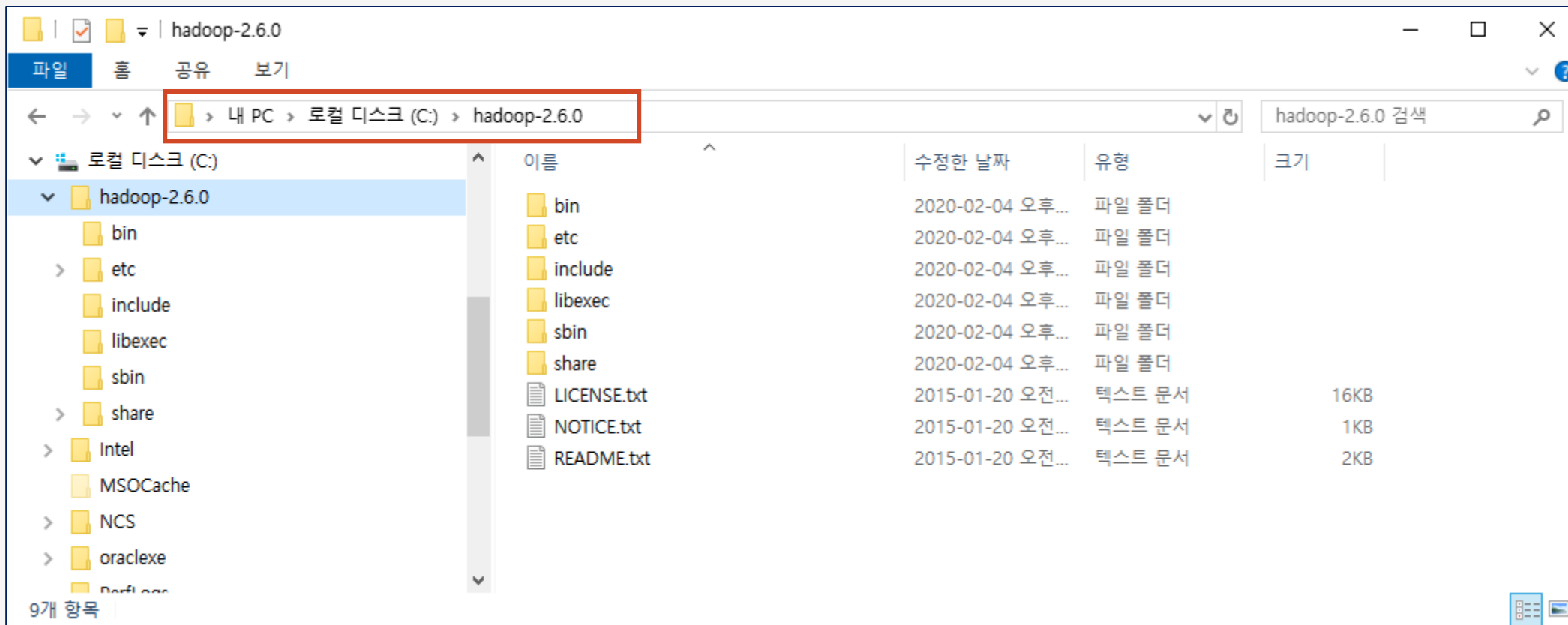
Google: [tbarik@gmail.com](mailto:tbarik@gmail.com)  
Skype: [tbarik](#)  
Twitter: [@barik](#)

Search

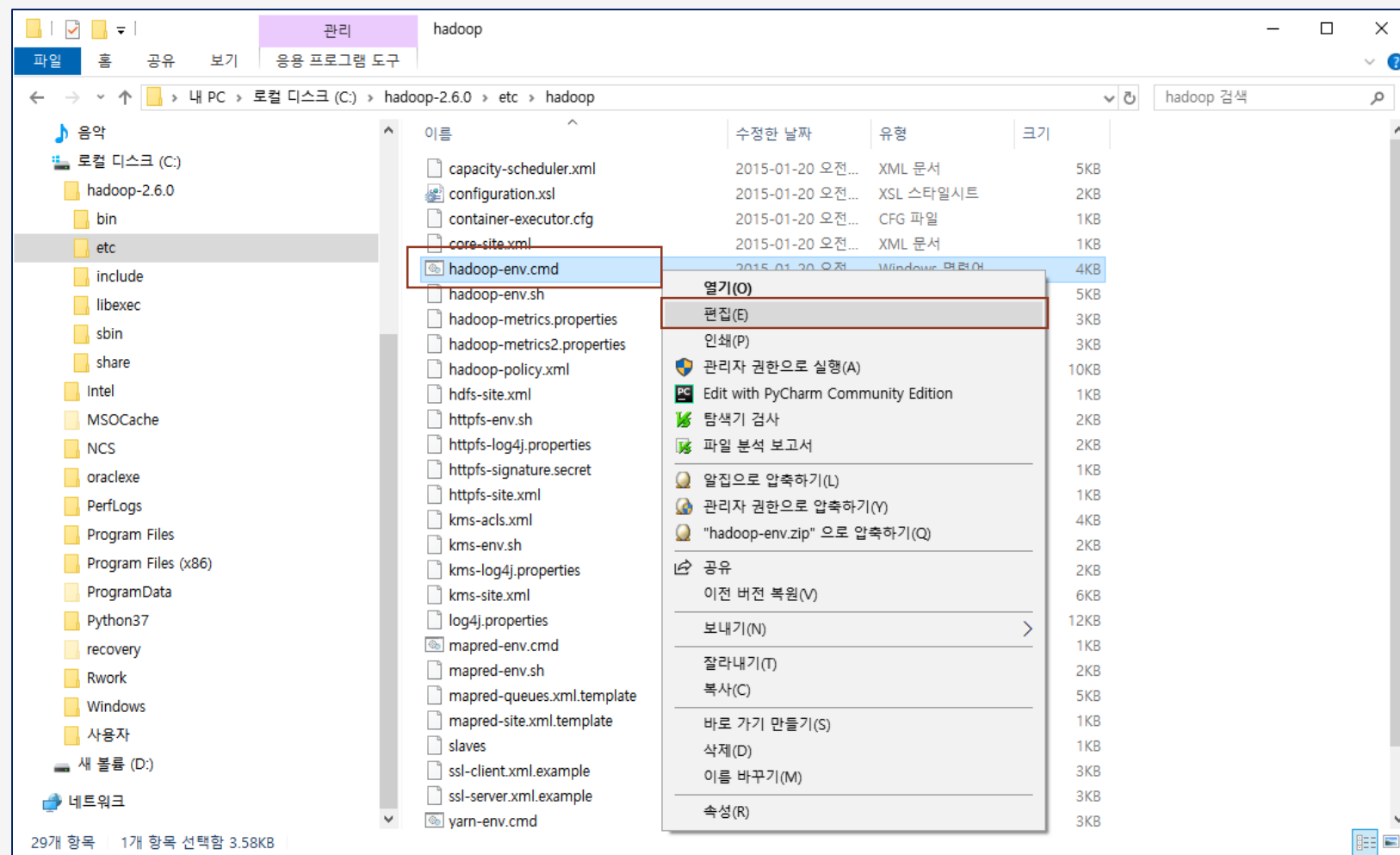
Search ...

Windows 정품 인증  
[설정]으로 이동하여 Windows를 정품 인증합니다.

# 압축 풀기



# 환경설정1 : Hadoop-env.cmd 설정



# 환경설정1

The image shows a Windows environment with two windows. On the left is a Notepad++ window titled 'hadoop-env.cmd - 메모장' containing Hadoop environment configuration text. On the right is a File Explorer window showing the path 'C:\Program Files\Java\jdk1.8.0\_151'.

**hadoop-env.cmd - 메모장**

```
@rem WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either  
@rem See the License for the specific language governing  
@rem limitations under the License.  
  
@rem Set Hadoop-specific environment variables here.  
  
@rem The only required environment variable is JAVA_HOME.  
@rem optional. When running a distributed configuration tool like  
@rem set JAVA_HOME in this file, so that it is correctly set on  
@rem remote nodes.  
  
@rem The java implementation to use. Required.  
set JAVA_HOME="C:\Program Files\Java\jdk1.8.0_151"  
  
@rem The jsvc implementation to use. Jsvc is required to run  
@rem set JSVC_HOME=%JSVC_HOME%  
  
@rem set HADOOP_CONF_DIR=  
  
@rem Extra Java CLASSPATH elements. Automatically insert capacity-scheduler.  
if exist %HADOOP_HOME%\contrib\capacity-scheduler (  
  if not defined HADOOP_CLASSPATH (  
    set HADOOP_CLASSPATH=%HADOOP_HOME%\contrib\capacity-scheduler\*.jar  
  ) else (  
    set HADOOP_CLASSPATH=%HADOOP_CLASSPATH%;%HADOOP_HOME%\contrib\capacity-scheduler\*.jar  
  )  
)
```

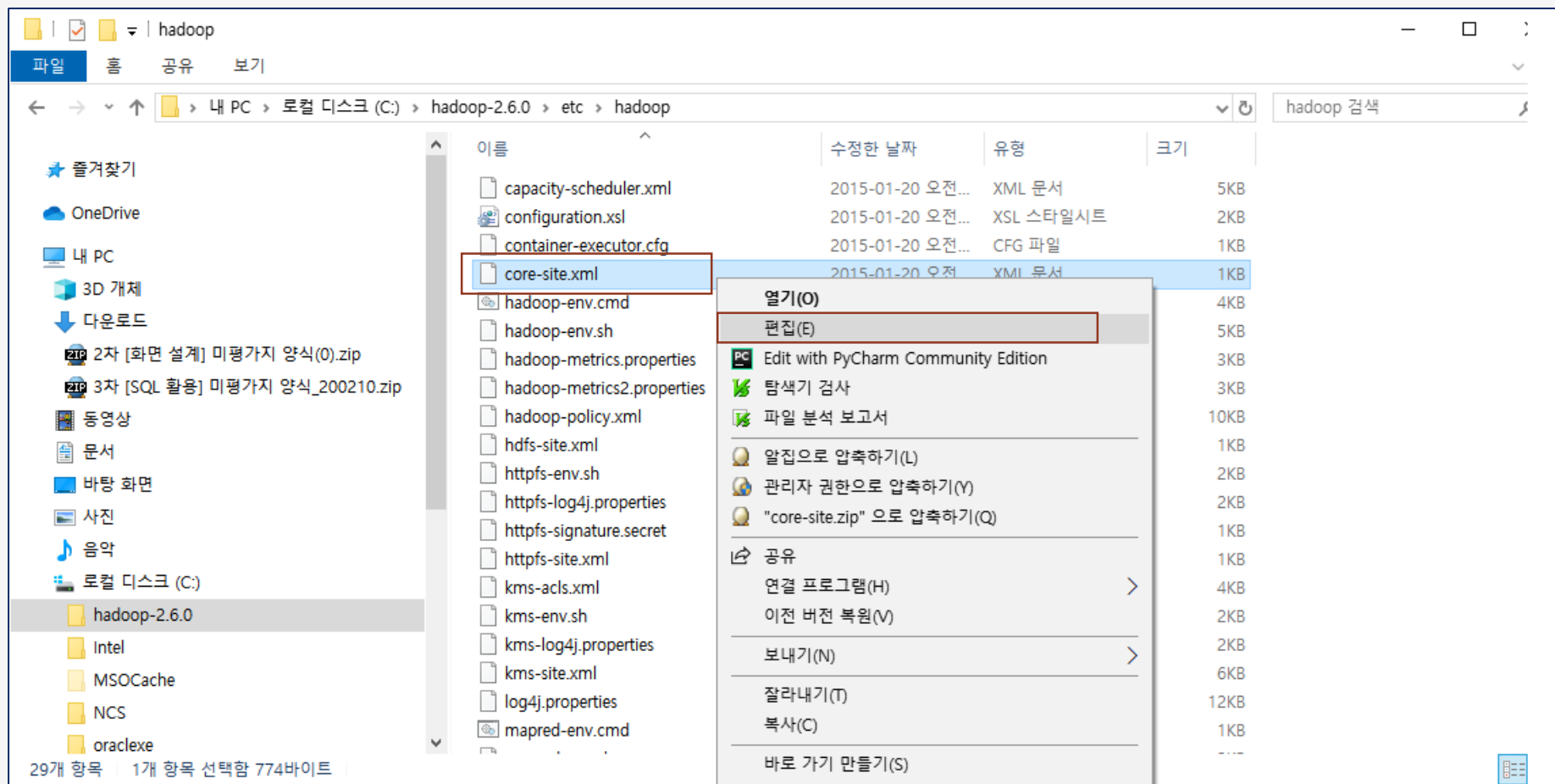
**File Explorer: jdk1.8.0\_151**

Path: > 내 PC > 로컬 디스크 (C:) > Program Files > Java > jdk1.8.0\_151

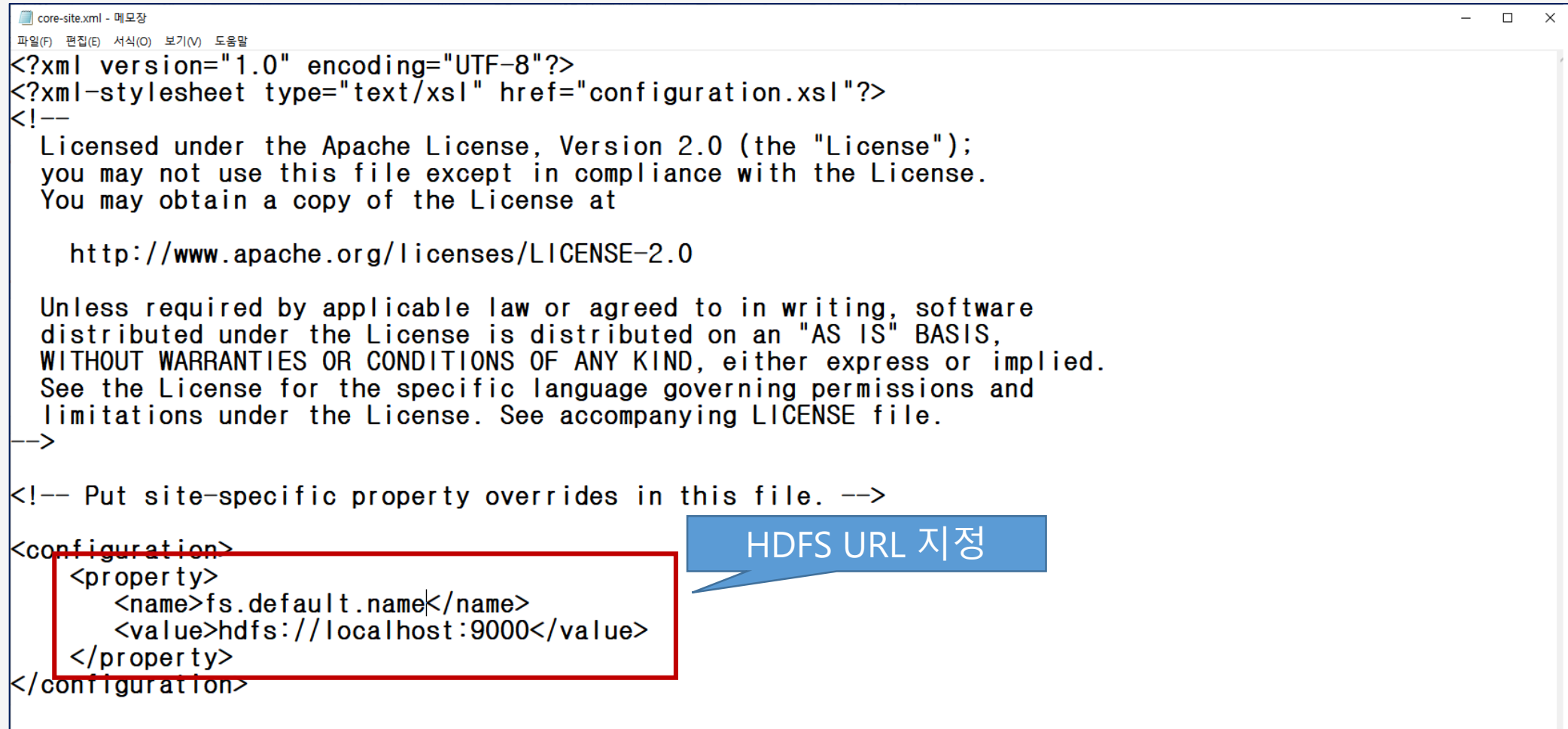
이름	수정한 날짜	유형	크기
bin	2019-11-12 오전...	파일 폴더	
db	2019-11-12 오전...	파일 폴더	
include	2019-11-12 오전...	파일 폴더	
jre	2019-11-12 오전...	파일 폴더	
lib	2019-11-12 오전...	파일 폴더	
COPYRIGHT	2017-09-05 오후...	파일	4KB
javaafx-src.zip	2019-11-12 오전...	ALZip ZIP File	5,079KB
LICENSE	2019-11-12 오전...	파일	1KB
README.html	2019-11-12 오전...	Chrome HTML D...	1KB
release	2019-11-12 오전...	파일	1KB
src.zip	2017-09-05 오후...	ALZip ZIP File	20,762KB
THIRDPARTYLICENSEREADME.txt	2019-11-12 오전...	텍스트 문서	142KB
THIRDPARTYLICENSEREADME-JAVAFX.txt	2019-11-12 오전...	텍스트 문서	63KB

❖ JAVA\_HOME에 jdk 경로를 지정한다. 만약 공백이 있다면 "\""로 감싸준다.

## 환경설정2 : core-site.xml 설정



## 환경설정2



```
core-site.xml - 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말
<?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

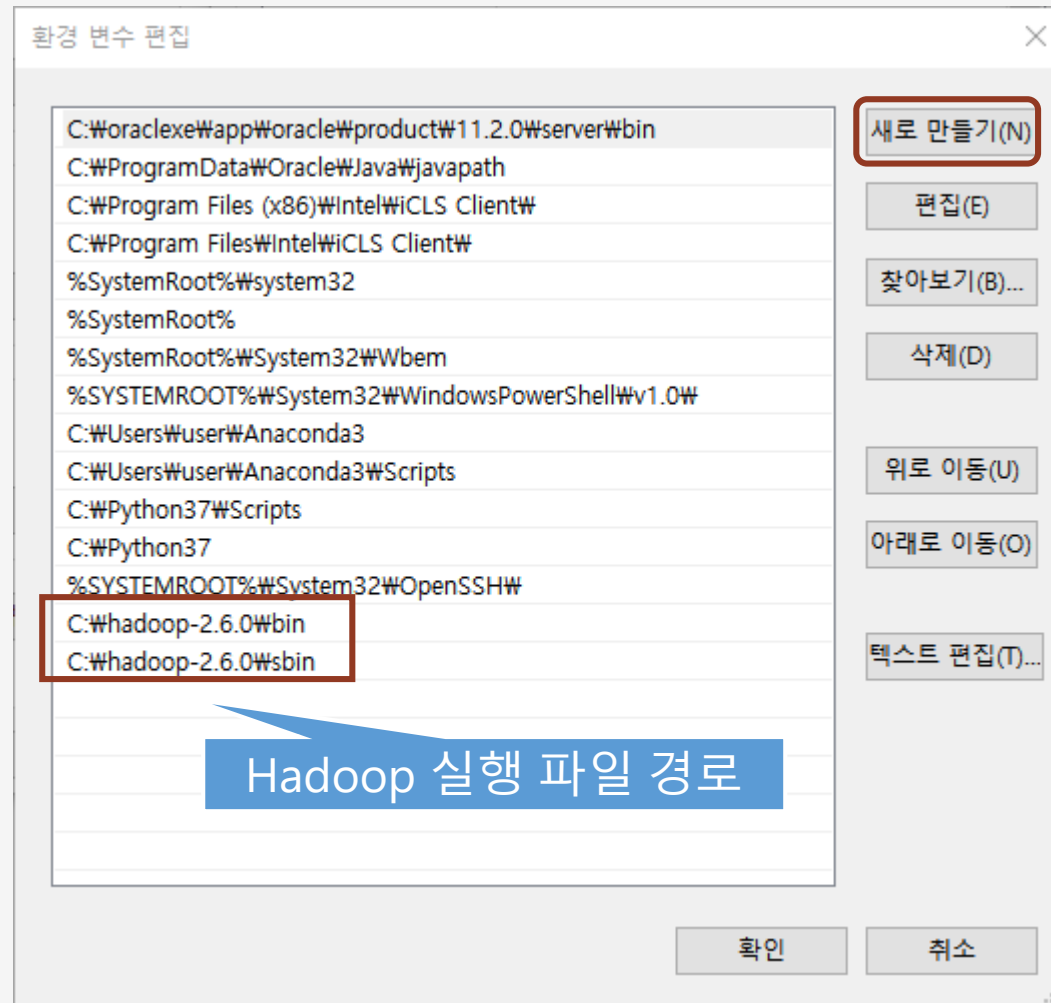
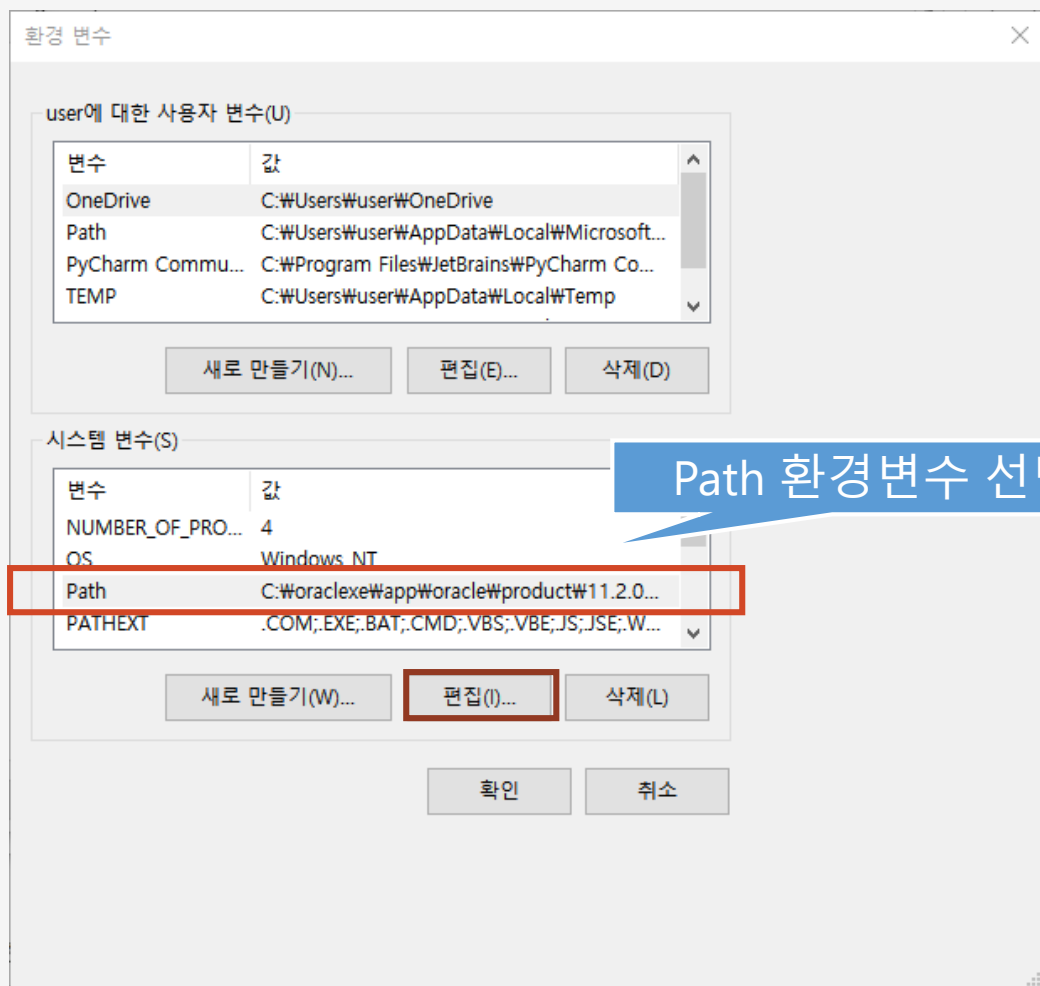
  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under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

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

<configuration>
  <property>
    <name>fs.default.name</name>
    <value>hdfs://localhost:9000</value>
  </property>
</configuration>
```

HDFS URL 지정

## 환경설정3 : Path 환경설정





# Name node Format

명령 프롬프트

C:\Users\User>

C:\Users\User>

C:\Users\User>hdfs namenode -format

'C:\Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는

배치 파일이 아닙니다.

20/02/04 23:58:29 INFO namenode.NameNode: ST

/\*\*\*\*\*

STARTUP\_MSG: Starting NameNode

STARTUP\_MSG: host = LAPTOP-7JT8HEBE/192.168

STARTUP\_MSG: args = [-format]

STARTUP\_MSG: version = 2.6.0

STARTUP\_MSG: classpath = C:\hadoop-2.6.0\share\hadoop\com

op-2.6.0\share\hadoop\common\lib\apacheds-1.4.7

odec-2.0.0-M15.jar;C:\hadoop-2.6.0\share\hadoop\com

n\lib\api-util-1.0.0-M20.jar;C:\hadoop-2.6.0\share\hadoop\com

avro-1.7.4.jar;C:\hadoop-2.6.0\share\hadoop\com

lib\commons-beanutils-core-1.8.0.jar;C:\hadoop-2.6.0\share\hadoop\com

hadoop\common\lib\commons-codec-1.4.jar;C:\hadoop-2.6.0\share\hadoop\com

.6.0\share\hadoop\common\lib\commons-compress-1.2.1.jar;C:\hadoop-2.6.0\share\hadoop\com

6.jar;C:\hadoop-2.6.0\share\hadoop\common\lib\commons-io-2.4.jar;C:\hadoop-2.6.0\share\hadoop\com

mons-io-2.4.jar;C:\hadoop-2.6.0\share\hadoop\com

mons-logging-1.1.3.jar;C:\hadoop-2.6.0\share\hadoop\com

on\lib\commons-net-3.1.jar;C:\hadoop-2.6.0\share\hadoop\com

hadoop\common\lib\curator-framework-2.6.0.jar;C:\hadoop-2.6.0\share\hadoop\com

0\share\hadoop\common\lib\gson-2.2.4.jar;C:\hadoop-2.6.0\share\hadoop\com

hadoop\common\lib\hadoop-annotations-2.6.0.jar;C:\hadoop-2.6.0\share\hadoop\com

.6.0\share\hadoop\common\lib\hamcrest-core-1.3.jar;C:\hadoop-2.6.0\share\hadoop\com

op-2.6.0\share\hadoop\common\lib\httpclient-4.5.1.jar;C:\hadoop-2.6.0\share\hadoop\com

명령 프롬프트

20/02/04 23:58:36 INFO util.GSet: 1.0% max memory 889 MB = 8.9 MB

20/02/04 23:58:36 INFO util.GSet: capacity = 2^20 = 1048576 entries

20/02/04 23:58:36 INFO namenode.NameNode: Caching file names occurring more than 10 times

20/02/04 23:58:36 INFO util.GSet: Computing capacity for map cachedBlocks

20/02/04 23:58:36 INFO util.GSet: VM type = 64-bit

20/02/04 23:58:36 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB

20/02/04 23:58:36 INFO util.GSet: capacity = 2^18 = 262144 entries

20/02/04 23:58:36 INFO namenode.FSNamesystem: dfs.namenode.safemode.threshold-pct = 0.9990000128746033

20/02/04 23:58:36 INFO namenode.FSNamesystem: dfs.namenode.safemode.min.datanodes = 0

20/02/04 23:58:36 INFO namenode.FSNamesystem: dfs.namenode.safemode.extension = 30000

20/02/04 23:58:36 INFO namenode.FSNamesystem: Retry cache on namenode is enabled

20/02/04 23:58:36 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis

20/02/04 23:58:36 INFO util.GSet: Computing capacity for map NameNodeRetryCache

20/02/04 23:58:36 INFO util.GSet: VM type = 64-bit

20/02/04 23:58:36 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB

20/02/04 23:58:36 INFO util.GSet: capacity = 2^15 = 32768 entries

20/02/04 23:58:36 INFO namenode.NNConf: ACLs enabled? false

20/02/04 23:58:36 INFO namenode.NNConf: XAttrs enabled? true

20/02/04 23:58:36 INFO namenode.NNConf: Maximum size of an xattr: 16384

20/02/04 23:58:37 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1233990877-192.168.200.157-1580828316993

20/02/04 23:58:37 INFO common.Storage: Storage directory \tmp\hadoop-user\dfs\name has been successfully formatted.

20/02/04 23:58:37 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0

20/02/04 23:58:37 INFO util.ExitUtil: Exiting with status 0

20/02/04 23:58:37 INFO namenode.NameNode: SHUTDOWN\_MSG:

SHUTDOWN\_MSG: Shutting down NameNode at LAPTOP-7JT8HEBE/192.168.200.157

\*\*\*\*

C:\Users\User>

# hdfs와 yarn 실행

명령 프롬프트

C:\Users\User>

C:\Users\User>start-all.cmd

This script is Deprecated. Instead use start-dfs.cmd and start-yarn.  
'C:\Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는  
배치 파일이 아닙니다.  
'C:\Program'은(는) 내부 또는 외부 명령, 실행할  
배치 파일이 아닙니다.  
starting yarn daemons  
'C:\Program'은(는) 내부 또는 외부 명령, 실행할  
배치 파일이 아닙니다.

C:\Users\User>

namenode, datanode, nodemanager, resourcemanager가  
모두 열리고, 모두 종료되지 않은 상태로 있어야 성공한 것

Apache Hadoop Distribution - hadoop datanode

Apache Hadoop Distribution - hadoop namenode

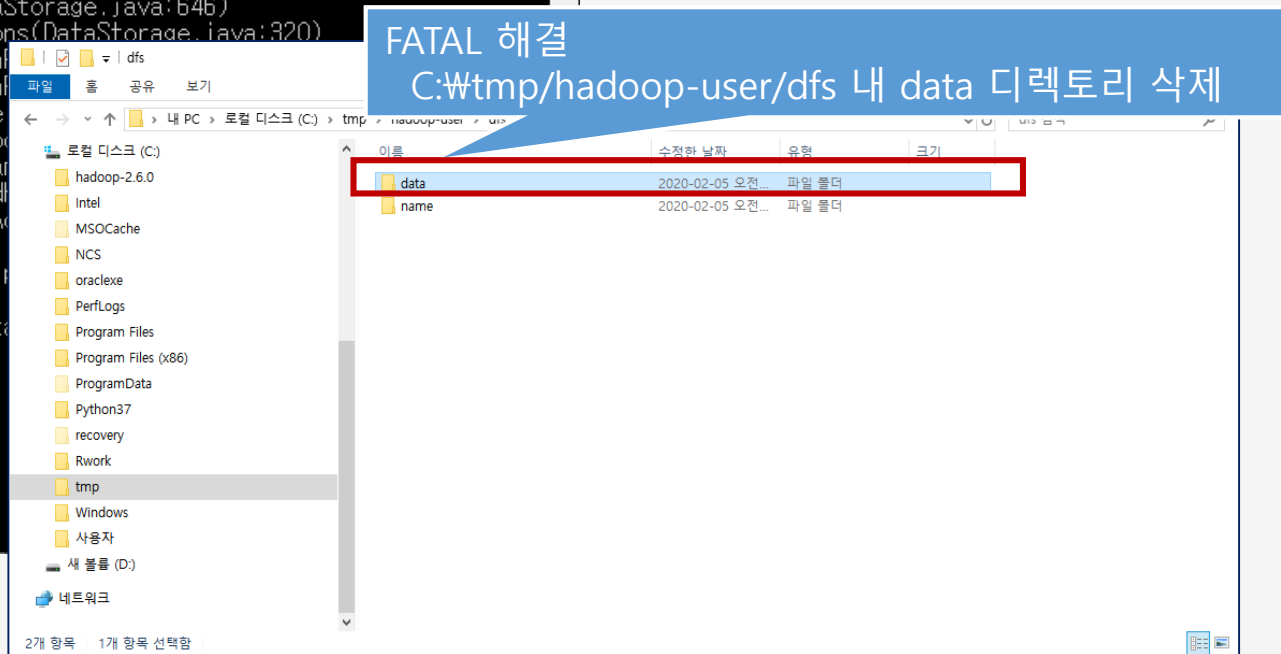
Apache Hadoop Distribution - yarn resourcemanager

Apache Hadoop Distribution - yarn nodemanager

```
20/2/5 00:05:59 INFO http.HttpRequestLog: Http request log for http.requests.nodemanager is not defined
20/2/5 00:05:59 INFO http.HttpServer2: Added global filter 'safety' (class=org.apache.hadoop.http.HttpServer2$QuotingInputFilter)
20/2/5 00:05:59 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context node
20/2/5 00:05:59 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context static
20/2/5 00:05:59 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context logs
20/2/5 00:05:59 INFO http.HttpServer2: adding path spec: /node/*
20/2/5 00:05:59 INFO http.HttpServer2: adding path spec: /ws/*
20/2/5 00:05:59 INFO http.HttpServer2: Jetty bound to port 8042
20/2/5 00:05:59 INFO mortbay.log: jetty-6.1.26
20/2/5 00:05:59 INFO mortbay.log: Extract jar:file:/C:/hadoop-2.6.0/share/hadoop/yarn/hadoop-yarn-common-2.6.0.jar!/webapps/node to C:\Users\User\AppData\Local\Temp\Jetty_0_0_0_8042_node_19tj0xwebapp
20/2/5 00:06:00 INFO mortbay.log: Started HttpServer2$SelectChannelConnectorWithSafeStartup@0.0.0.0:8042
20/2/5 00:06:00 INFO webapp.WebApps: Web app /node started at 8042
20/2/5 00:06:00 INFO webapp.WebApps: Registered webapp guice modules
20/2/5 00:06:00 INFO client.RMPProxy: Connecting to ResourceManager at /0.0.0.0:8031
20/2/5 00:06:01 INFO nodemanager.NodeStatusUpdaterImpl: Sending out 0 NM container statuses: []
20/2/5 00:06:01 INFO nodemanager.NodeStatusUpdaterImpl: Registering with RM using containers: []
20/2/5 00:06:01 INFO security.NMContainerTokenSecretManager: Rolling master-key for container-tokens, got key with id 1284112573
20/2/5 00:06:01 INFO security.NMTokenSecretManagerInNM: Rolling master-key for container-tokens, got key with id 832695464
20/2/5 00:06:01 INFO nodemanager.NodeStatusUpdaterImpl: Registered with ResourceManager as LAPTOP-7JT8HEBE:56723 with total resource of <memory:8192, vCores:8>
20/2/5 00:06:01 INFO nodemanager.NodeStatusUpdaterImpl: Notifying ContainerManager to unblock new container-requests
```

# DataNode FATAL 오류 해결

```
Apache Hadoop Distribution
20/02/05 00:41:54 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 0 time(s); retry
policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
20/02/05 00:41:54 INFO common.Storage: DataNode version: -56 and NameNode layout version: -60
20/02/05 00:41:54 INFO common.Storage: Lock on #tmp#hadoop-user#dfs#data#in_use.lock acquired by nodename 9708@LAPTOP-7J
T8HEBE
20/02/05 00:41:54 FATAL datanode.DataNode: Initialization failed for Block pool <registering> (Datanode Uuid unassigned)
service to localhost/127.0.0.1:9000. Exiting.
java.io.IOException: Incompatible clusterIDs in C:\tmp\hadoop-user\dfs\data: namenode clusterID = CID-b57bc9b5-bef5-4e33
911b-37571d1b2f38; datanode clusterID = CID-fa033a44-2b04-45bb-b448-7e5b4f188188
    at org.apache.hadoop.hdfs.server.datanode.DataStorage.doTransition(DataStorage.java:646)
    at org.apache.hadoop.hdfs.server.datanode.DataStorage.addStorageLocations(DataStorage.java:320)
    at org.apache.hadoop.hdfs.server.datanode.DataStorage.recoverTransitionInfo(DataStorage.java:100)
    at org.apache.hadoop.hdfs.server.datanode.DataNode.initStorage(DataNode.java:100)
    at org.apache.hadoop.hdfs.server.datanode.DataNode.initBlockPool(DataNode.java:100)
    at org.apache.hadoop.hdfs.server.datanode.BPOfferService.verifyAndSetNameNode(BPOfferService.java:100)
    at org.apache.hadoop.hdfs.server.datanode.BPServiceActor.connectToNameNode(BPServiceActor.java:100)
    at org.apache.hadoop.hdfs.server.datanode.BPServiceActor.run(BPServiceActor.java:100)
    at java.lang.Thread.run(Thread.java:748)
20/02/05 00:41:54 WARN datanode.DataNode: Ending block pool service for: Block pool <registering> (Data
node) service to localhost/127.0.0.1:9000
20/02/05 00:41:54 INFO datanode.DataNode: Removed Block pool <registering> (Data
node) service to localhost/127.0.0.1:9000
20/02/05 00:41:56 WARN datanode.DataNode: Exiting Datanode
20/02/05 00:41:56 INFO util.ExitUtil: Exiting with status 0
20/02/05 00:41:57 INFO datanode.DataNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down DataNode at LAPTOP-7JT8HEBE/192.168.200.157
*****/
C:\Users\User>
```



# HDFS 테스트

**명령 프롬프트**

'c:\#'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는 배치 파일이 아닙니다.

**디렉터리 이동**

C:\Users\User>cd c:\#

c:\#>dir

C 드라이브의 볼륨에는 이름이 없습니다.  
볼륨 일련 번호: FOAC-9395

c:\# 디렉터리

날짜	시간	타입	이름	크기
2020-02-04	오후 11:58	<DIR>		
2016-12-23	오후 01:20	<DIR>		
2019-11-22	오후 03:16	<DIR>		
2020-01-27	오후 04:05	<DIR>		
2018-09-15	오후 04:33	<DIR>		
2019-12-27	오후 05:42	<DIR>		
2019-12-16	오후 11:58	<DIR>		
2019-09-02	오후 12:41	<DIR>		
2016-12-23	오후 01:20	<DIR>		
2019-10-30	오후 12:03	<DIR>		
2019-06-22	오후 06:42	<DIR>		
2020-02-04	오후 11:58	<DIR>		
2019-07-25	오전 12:44	<DIR>		
2020-02-04	오전 10:35	<DIR>		

2개 파일  
12개 디렉터리 79,618

**명령 프롬프트**

2018-09-15 오후 04:33 <DIR> PerfLogs  
2019-12-27 오후 05:42 <DIR> Program Files  
2019-12-16 오후 11:58 <DIR> Program Files (x86)  
2019-09-02 오후 12:41 <DIR> Python37  
2016-12-23 오후 01:20 2,699 rhdssetup.log  
2019-10-30 오후 12:03 <DIR> Rwork  
2019-06-22 오후 06:42 201 setup.log  
2020-02-04 오후 11:58 <DIR> tmp  
2019-07-25 오전 12:44 <DIR> Users  
2020-02-04 오전 10:35 <DIR> Windows

2개 파일 2,900 바이트  
12개 디렉터리 79,618,764,800 바이트 남음

**Hdfs에 파일 업로드**

c:\#>hdfs dfs -put setup.log /

'C:\#Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는 배치 파일이 아닙니다.

**Hdfs 업로드 파일 검색**

c:\#>hdfs dfs -ls /

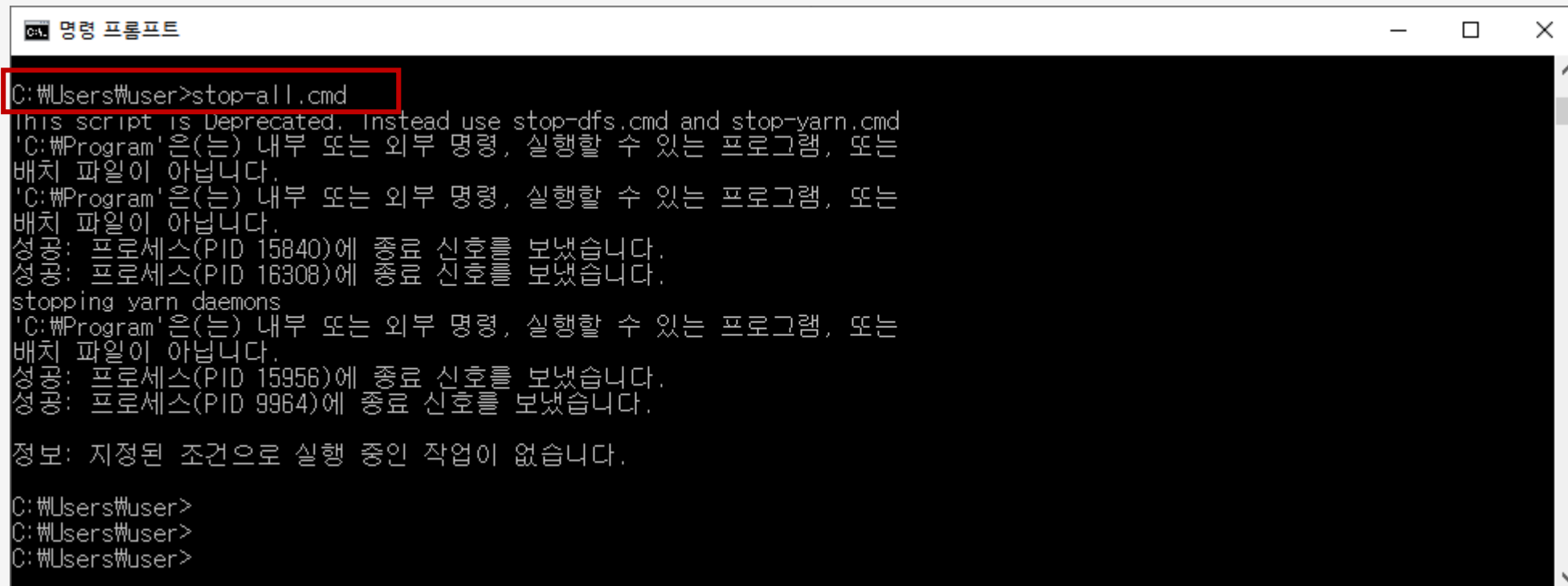
'C:\#Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는 배치 파일이 아닙니다.

Found 1 items

권한	유저	그룹	크기	날짜	경로
-rw-r--r--	3 user	supergroup	201	2020-02-05 00:50	/setup.log

c:\#>

# hdfs와 yarn 종료



```
CA. 명령 프롬프트
C:\Users\user>stop-all.cmd
This script is Deprecated. Instead use stop-dfs.cmd and stop-yarn.cmd
'C:\Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는
배치 파일이 아닙니다.
'C:\Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는
배치 파일이 아닙니다.
성공: 프로세스(PID 15840)에 종료 신호를 보냈습니다.
성공: 프로세스(PID 16308)에 종료 신호를 보냈습니다.
stopping yarn daemons
'C:\Program'은(는) 내부 또는 외부 명령, 실행할 수 있는 프로그램, 또는
배치 파일이 아닙니다.
성공: 프로세스(PID 15956)에 종료 신호를 보냈습니다.
성공: 프로세스(PID 9964)에 종료 신호를 보냈습니다.

정보: 지정된 조건으로 실행 중인 작업이 없습니다.

C:\Users\user>
C:\Users\user>
C:\Users\user>
```