

[Spark-I]

1. <https://spark.apache.org/downloads.html> 사이트에서 [spark-2.3.1-bin-hadoop2.8.tgz](#) 다운로드

2. 설치

```
[nova@master ~]$ cd 다운로드/
[nova@master 다운로드]$ ls
2007.csv.bz2 2008.csv.bz2 57 hadoop-3.0.3.tar.gz
spark-2.3.1-bin-hadoop2.7.tgz jdk-8u181-linux-x64.tar.gz
[nova@master 다운로드]$ cd
[nova@master ~]$ tar -xf ~/다운로드/spark-2.3.1-bin-hadoop2.7.tgz
[nova@master ~]$ mv spark-2.3.1-bin-hadoop2.7/ spark-2.3.1/
[nova@master ~]$ ls # 설치 확인
hadoop-3.0.3 spark-2.3.1 다운로드 jdk1.8.0_181 공개 문서 바탕화면 사진 음악 비디오 서식
```

3. 환경 설정

```
[nova@master ~]$ cd spark-2.3.1/conf/
[nova@master conf]$ cp spark-env.sh.template spark-env.sh
[nova@master conf]$ vi spark-env.sh
#!/usr/bin/env bash
export SPARK_DIST_CLASSPATH=$(~/hadoop-3.0.3/bin/hadoop classpath) export JAVA_HOME=~/.jdk1.8.0_181
[nova@master conf]$ cp log4j.properties.template log4j.properties
[nova@master conf]$ vi log4j.properties
# Set everything to be logged to the console log4j.root log4j.rootCategory=WARN, # 변경: INFO -> WARN
consolelog4j.appender.console=org.apache.log4j.ConsoleAppender
```

4. 실행

```
[nova@master conf]$ ~/spark-2.3.1/sbin/start-all.sh
starting org.apache.spark.deploy.master.Master, logging to /home/nova/spark2.3.1/logs/spark-
nova-org.apache.spark.deploy.master.Master-1-master.out
localhost: starting org.apache.spark.deploy.worker.Worker, logging to /home/nova/spark-
2.3.1/logs/spark-nova-org.apache.spark.deploy.worker.Worker-1- master.out
[nova@master conf]$ jps
26803 Worker
26868 Jps
25156 ResourceManager
26644 Master
24778 NameNode
```