

# 자바쌤~!

https://cafe.naver.com/javassem

검색

63기-자바 & 빅데이터 >

## [ 스파크 ] 설치



자바쌤 1:1 채팅  
2020.06.25. 11:16 조회 3

댓글 0   URL 복사

```
<dn01의 root 계정에서 >
[root@dn01 ~]# cd /tmp
[root@dn01 tmp]# wget http://apache.mirror.cdnetworks.com/spark/spark-2.4.5/spark-2.4.5-bin-hadoop2.7.tgz
[root@dn01 tmp]# tar xzvf spark-2.4.5-bin-hadoop2.7.tgz
[root@dn01 tmp]# mkdir -p /opt/spark/2.4.5
[root@dn01 tmp]# mv spark-2.4.5-bin-hadoop2.7/* /opt/spark/2.4.5/
[root@dn01 tmp]# ln -s /opt/spark/2.4.5 /opt/spark/current
[root@dn01 tmp]# chown -R hadoop:hadoop /opt/spark/
[root@dn01 tmp]# su - hadoop
[hadoop@dn01 ~]$ vi ~/.bash_profile
##### spark #####
export SPARK_HOME=/opt/spark/current
export PATH=$PATH:$SPARK_HOME/bin
export PATH=$PATH:$SPARK_HOME/sbin
#### spark #####
[hadoop@dn01 ~]$ source ~/.bash_profile
[hadoop@dn01 conf]$ cd $SPARK_HOME/conf
[slaves]
nn01
dn02
( 기존 localhost는 지우자 )
```

```
[spark-defaults.conf]
spark.yarn.jars /opt/spark/current/jars/*

[log4j.properties]
#INFO -> ERROR로 바꿔 줌 -> Spark 로그에 정신없는 INFO가 안나타남
log4j.rootCategory=ERROR, console
```

```
[spark-env.sh]
SPARK_MASTER_HOST=dn01
export JAVA_HOME=/opt/jdk/current
export HADOOP_HOME=/opt/hadoop/current
export SPARK_HOME=/opt/spark/current
export HADOOP_CONF_DIR=$HADOOP_HOME/etc/hadoop
export YARN_CONF_DIR=$HADOOP_HOME/etc/hadoop
export SPARK_DRIVER_MEMORY=2g
export SPARK_EXECUTOR_INSTANCES=2
export SPARK_EXECUTOR_CORES=1
export SPARK_EXECUTOR_MEMORY=2g
export SPARK_MASTER_IP=192.168.56.102
#export SPARK_WORKER_DIR=/spark_data/spwork
#export SPARK_PID_DIR=/spark_data/sptmp
export SPARK_DIST_CLASSPATH=$(/opt/hadoop/current/bin/hadoop classpath):/opt/spark/current/jars/*
```

### 카페정보   나의활동



매니저 자바쌤  
since 2013.03.15.  
카페소개

새싹2단계	
436	초대하기
즐거찾는 멤버	64명
게시판 구독수	4회
우리카페업 수	0회

카페 글쓰기

카페 채팅

### 즐거찾는 게시판

전체글보기 4,082

### KOSMO

63기-자바 & 빅데이터

### 팁

- 심심문제
- 기타등등
- JAVA
- Oracle/MySQL
- JSP
- WebUI
- 파이썬
- Spring
- 리눅스
- 하둡+eco
- 프로젝트

### 공유해요

- 알고리즘도전
- 기술면접 준비하기
- 실존 면접
- 기회는 준비된자에게
- 프로젝트 팁

### 쉬어가기

### 청년아카데미

### 전문가과정

공지사항  
카페회칙

카페탈퇴하기



```
#export PYTHONPATH=/opt/python/current/python3
#export PYSARK_PYTHON=/opt/python/current/python3
```

```
[hadoop@dn01 ~] spark-shell
```

```
spark-shell 확인
```

```
scala> sc.setLogLevel("WARN")
scala> val f = sc.textFile("file:///etc/hosts")
scala> f.count
scala> f.first
scala> f.collect
scala> :quit
```

```
=====
```

dn01 노드에서 dn02와 nn01로 전송 복사

```
$ sudo scp -r /opt/spark dn02:/opt/spark
```

```
$ sudo scp -r /opt/spark nn01:/opt/spark
```

scp로 전송하면 softlink는 안잡힌다.

current가 디렉토리로 된 거 확인하고 디렉토리를 지우고 softlink로 설정

dn02와 nn01 [ root 계정 ] 에서 확인

```
[root@dn02 ~]# rm -rf /opt/spark/current
```

```
[root@dn02 ~]# ln -s /opt/spark/2.4.5 /opt/spark/current
```

```
[root@dn02 ~]# ll /opt/spark/
```

```
[root@dn02 ~]# chown -R hadoop:hadoop /opt/spark/
```

```
[root@dn02 ~]# su - hadoop
```

아래처럼 안하려면 dn01에서 scp로 전송

```
[hadoop@dn02 ~]$ vi ~/.bash_profile
```

```
##### spark #####
```

```
export SPARK_HOME=/opt/spark/current
```

```
export PATH=$PATH:$SPARK_HOME/bin
```

```
export PATH=$PATH:$SPARK_HOME/sbin
```

```
#### spark #####
```

```
[hadoop@dn02 ~]$ source ~/.bash_profile
```

```
[hadoop@nn01 ~]$ vi ~/.bash_profile
```

```
##### spark #####
```

```
export SPARK_HOME=/opt/spark/current
```

```
export PATH=$PATH:$SPARK_HOME/bin
```

```
export PATH=$PATH:$SPARK_HOME/sbin
```

```
#### spark #####
```

```
[hadoop@nn01 ~]$ source ~/.bash_profile
```

## 댓글

댓글알림

김용선

댓글을 남겨보세요



등록

---

자바쌤~!    <https://cafe.naver.com/javassem>