

## [Hadoop 클러스터 구축을 위한 설정II]

### 2. 설정 순서

(1) `hadoop-env.sh` (jdk 경로, 클래스 패스, 데몬 실행 옵션 등 설정)

모든 노드 `export JAVA_HOME=/home/nova/jdk1.8.0_181`

(2) `core-site.xml` (hdfs와 맵리듀스에서 공통적으로 사용할 환경 정보 설정)

모든 노드 `fs.defaultFS ; hdfs://master:9000`

(3) `hdfs-site.xml` (하둡 분산 파일 시스템 설정을 오버라이드하는 파일)

네임노드① `dfs.replication ; 2` (복제수)

네임노드② `dfs.namenode.name.dir ; file:///dfs/name` (파일시스템 이미지 저장될 경로)

네임노드③ `dfs.namenode.edits.dir ; file:///dfs/edits` (에디트로그가 저장될 경로)

네임노드④ `dfs.namenode.secondary.http-address ; backup:9868` (보조네임노드와 포트번호)

데이터노드 `dfs.datanode.data.dir ; file:///dfs/data` (데이터 파일이 저장될 경로)

보조네임노드 `dfs.namenode.checkpoint.dir ; file:///dfs/namesecondary`

(파일시스템이미지, 에디트로그 사본이 저장될 경로)

(4) `mapred-site.xml`(맵 리듀스 프레임워크 이름 지정)

네임노드 `mapreduce.framework.name ; yarn`

(5) `yarn-site.xml` (맵리듀스 프레임워크에서 사용할 셔플 서비스 지정)

리소스매니저 `yarn.nodemanager.aux-services ; mapreduce_shuffle` (맵리듀스가 사용할 보조서비스)

노드매니저 `yarn.resourcemanager.hostname ; master`

(6) `workers` (데이터 노드들의 host이름들 지정)

(7) 데이터노드와 보조네임노드의 `hdfs-site.xml`

데이터노드 `dfs.datanode.data.dir ; file:///dfs/data` (데이터 파일이 저장될 경로)

보조네임노드 `dfs.namenode.checkpoint.dir ; file:///dfs/namesecondary`

(파일시스템이미지, 에디트로그 사본이 저장될 경로)

(8) 데이터노드의 `yarn-site.xml`

노드매니저 `yarn.resourcemanager.hostname ; master`

3. 네임노드 포맷 (한번만 실행하도록 한다.)

4. 하둡 클러스터 실행

`start-all.sh` (反) `stop-all.sh`