

HDFS 명령어

HDFS 명령어

- HDFS은 하둡 분산 파일 시스템의 파일일들을 조작하고 관리할 수 있는 명령어를 제공함
 - `$ hadoop fs -명령어 [파일명 또는 디렉토리명]`
 - `$ hadoop fs -help`
- `ls`
 - 지정한 디렉토리에 파일의 정보를 출력하거나, 특정한 파일을 선택해서 출력하는 명령어
 - `$ bin/hadoop fs -ls /`
 - `$ bin/hadoop fs -ls wordcount`

HDFS 명령어

- `lsr`
 - `lsr` 명령어는 현재 디렉토리의 하위 디렉토리 정보까지 출력하는 명령어
 - `$ bin/hadoop fs -lsr wordcount`
- `du`
 - 지정한 디렉토리나 파일의 사용량을 확인하는 명령어로, 바이트 단위로 결과를 출력하는 명령어
 - `$ bin/hadoop fs -du wordcount`
- `dus`
 - `du` 명령어는 디렉토리와 파일별로 용량을 출력하지만 `dus`는 전체 합계 용량만 출력하는 명령어
 - `$ bin/hadoop fs -dus wordcount`

HDFS 명령어

- cat
 - 지정한 파일의 내용을 화면에 출력하는 명령어
 - `$ bin/hadoop fs -cat conf/hadoop-env.sh`
- text
 - Cat 명령어는 텍스트 파일만 출력할 수 있지만, text명령어는 zip파일 형태로 압축된 파일도 텍스트 형태로 화면에 출력하는 명령어
 - `$ bin/hadoop fs -text conf/hadoop-env.sh`
- mkdir
 - 지정한 경로에 디렉토리를 생성하는 명령어
 - `$ bin/hadoop fs -mkdir tempDir`

HDFS 명령어

- put
 - 지정한 로컬 파일 시스템의 파일 및 디렉터리를 목적지 경로로 복사하는 명령어
 - `$ bin/hadoop fs -put conf tempDir`
- copyFromLocal
 - put 명령어와 동일한 기능을 제공하는 명령어
 - `$ bin/hadoop fs -copyFromLocal [로컬 디렉터리, 파일] [목적지 디렉터리, 파일]`
- get
 - hdfs에서 저장된 데이터를 로컬 파일 시스템으로 복사하는 명령어
 - `$ bin/hadoop fs -get wordcount`

HDFS 명령어

- getmerge
 - 지정한 경로에 있는 모든 파일의 내용을 합친 후, 로컬 파일 시스템에 단 하나의 파일로 복사하는 명령어
 - `$ bin/hadoop fs -getmerge wordcount/_logs/history_wordlog`
- cp
 - 지정한 소스 디렉터리 및 파일을 목적지 경로로 복사하는 기능의 명령어
 - `$ bin/hadoop fs -cp etc/hadoop/hadoop-env.sh /tmp`
- copyToLocal
 - get명령어와 동일한 기능을 제공하는 명령어
 - `$ bin/hadoop fs -copyToLocal [소스 디렉터리,파일][로컬 디렉터리, 파일]`

HDFS 명령어

- mv
 - 소스 디렉터리 및 파일을 목적지 경로로 옮기는 명령어
 - `$ bin/hadoop fs -mv etc /tmp/etc`
- moveFromLocal
 - put명령어와 동일하게 동작하지만 로컬 파일 시스템으로 파일이 복사된 후 소스 경로의 파일은 삭제되는 명령어
 - `$ bin/hadoop fs -moveFormLocal [소스 디렉터리, 파일][로컬 디렉터리,파일]`
- rm
 - 지정한 디렉터리나 파일을 삭제하는 명령어
 - `$ bin/hadoop fs -rm etc`

HDFS 명령어

- rmr

- 지정한 파일 및 디렉터리를 삭제하는 명령어. 비어있지 않은 디렉터리도 삭제되는 명령어
- `$ bin/hadoop fs -rmr /tmp/*`

- count

- 기본적으로 지정한 경로에 대한 전체 디렉터리 개수, 전체 파일 개수, 전체 파일 크기, 지정한 경로명을 출력하는 명령어
- `$ bin/hadoop fs -count wordcount`

- tail

- 지정한 파일의 마지막 1KB에 해당하는 내용을 화면에 출력하는 명령어
- `$ bin/hadoop fs -tail wordcount`

HDFS 명령어

- `chmod`
 - 지정한 경로에 대한 권한을 변경하는 명령어
 - 권한 모드는 숫자로 표기하는 8진수 표기법 및 영분으로 표시하는 심볼릭 표기법으로 설정함
 - `$ bin/hadoop fs -chmod 777 test.csv`
 - `$ bin/hadoop fs -chmod 755 sample.csv`
- `chown`
 - 지정한 파일과 디렉터리에 대한 소유권을 변경하는 명령어
 - `-R` 옵션을 사용하는 경우 재귀적으로 수행되어 하위디렉터리의 설정도 모두 변경됨
 - `$ bin/hadoop fs -chown hadoop:hadoopGroup test.csv`

HDFS 명령어

- touchz
 - 크기가 0 Byte인 파일을 생성하는 명령어
 - `$ bin/hadoop fs -touchz test.txt`
- stat
 - 지정한 경로에 대한 통계 정보를 조회하는 명령어
 - %b : 블록단위의 파일크기
 - %F : 디렉터리일 경우 "directory", 일반파일일 경우 "regular file"으로 출력
 - %n : 디렉터리명이나 파일명으로 출력
 - %o : 블록크기
 - %r : 복제 파일 개수
 - %y : 디렉터리 및 갱신일자를 `yyy-MM-dd HH:mm:ss`형식으로 출력
 - %Y : 디렉터리 및 파일 갱신일자를 유닉스 time stamp형식으로 출력
- `$ bin/hadoop fs -stat wordcount`
- `$ bin/hadoop fs -stat %b-%F-%n-%o-%r-%y-%Y wordcount`

HDFS 명령어

- setrep
 - 설정한 파일의 복제 데이터 개수를 변경할 수 있는 명령어
 - -R 옵션을 사용할 경우 하위 디렉터리에 있는 파일의 복제 데이터 개수도 변경되는 명령어
 - `$ bin/hadoop fs -setrep -w 1 test.csv`
- expunge
 - 휴지통을 비우는 명령어
 - .Trash/라는 임시 디렉터리에 저장한 후 일정시간이 지난후 완전히 삭제됨
 - `$ bin/hadoop fs -expunge`
- test
 - 지정한 경로에 대한 [-ezd] 옵션으로 경로가 이미 존재하는지 확인하고 파일크기가 0이거나 디렉터리인지 확인함
 - 체크 결과가 맞는 경우 0을 출력함