**최신기술프로젝트1 5주차 과제**

**HDFS Assignment**

컴퓨터공학과

2017103948 고다경

1. Load data on HDFS

텍스트이(가) 표시된 사진

자동 생성된 설명

* input2.txt란 파일을 업로드한 후, 파일 목록 확인을 하였습니다.

1. Replication Configuration : 1 -> 3
2. Block-size Configuration: 128MB -> 64MB

텍스트이(가) 표시된 사진

자동 생성된 설명

* hdfs-site.xml 파일에서 Replication을 3으로, Block-size를 67108864로 수정하였습니다.



* <path>에 있는 파일/디렉토리에 대한 통계를 print합니다.
* %b : bytes로 된 파일 사이즈
* %F : 유형
* %g : 소유자의 그룹 이름
* %n : 이름
* %o : 블록 크기  
  현재 블록 크기는 정해준 대로 6710884인 것을 볼 수 있습니다.
* %r : replication  
  현재 replication이 3개 인 것을 볼 수 있습니다.
* %u : owner의 USERNAME
* %y : 수정한 날짜(UTC 날짜로 표기)
* %Y : 수정한 날짜(1970년 1월 1일 UTC 이후의 milliseconds 표시)

텍스트이(가) 표시된 사진

자동 생성된 설명

* 현재 HDFS의 용량과 전체 노드의 용량 상태를 확인했습니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

* Replication이 1일 때, Default replication factor가 1인 것을 볼 수 있습니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

* Replication을 3으로 수정한 후, Default replication factor가 3으로 변경된 것을 확인했습니다.