

## 2.5 빅데이터 클러스터 구성

 HDFS/주키퍼 명령을 이용한 설치 확인

# 예제소스 및 파일 Download

- 파일럿 PC의 C 드라이브에 “예제소스”라는 폴더를 만들어 다운로드한 예제 소스를 복사해 놓는다.
    - 저장소 URL: <https://github.com/wikibook/bigdata2nd>
    - 우측 중앙의 [Clone or download] 버튼을 클릭한 후, [Download ZIP] 클릭
  - C://예제소스/bigdata2nd-master
-

- 샘플 파일 업로드

- 파일질라 FTP 클라이언트 실행

- Server02에 SFTP 접속

호스트: server02.hadoop.com

사용자명: bigdata

비밀번호: bigdata

포트: 22

[빠른연결] 버튼 클릭

- C://예제소스/bigdata2nd-master/CH02/Sample.txt 파일을 Server02의 /home/bigdata/ 경로에 업로드

## 2.5 빅데이터 클러스터 구성

 HDFS/주키퍼 명령을 이용한 설치 확인

# HDFS 명령어 사용

## ■ HDFS에 파일 저장

먼저 Server02에 SSH 접속을 하고 업로드한 샘플 파일이 있는 위치로 이동해서 HDFS의 put 명령을 실행한다.

### ■ PuTTY 프로그램 실행

### ■ Server02에 root 계정으로 SSH 접속

```
$ cd /home/bigdata
```

```
$ hdfs dfs -put Sample.txt /tmp
```

Sample.txt 파일이 HDFS의 /tmp 디렉터리로 저장된다.

## ■ HDFS에 저장한 파일 확인

```
$ hdfs dfs -ls /tmp
```

앞서 /tmp 디렉터리에 저장한 “Sample.txt” 파일의 목록이 조회된다.

## ■ HDFS에 저장한 파일 내용 보기

```
$ hdfs dfs -cat /tmp/Sample.txt
```

---

- HDFS에 저장한 파일 상태 확인

```
$ hdfs dfs -stat '%b %o %r %u %n' /tmp/Sample.txt
```

파일 크기(%b), 파일 블록 크기(%o), 복제 수(%r), 소유자명(%u), 파일명(%n) 정보를 보여준다.

- HDFS에 저장한 파일의 이름 바꾸기

```
$ hdfs dfs -mv /tmp/Sample.txt /tmp/Sample2.txt
```

기존 파일명인 "Sample.txt"를 "Sample2.txt"로 변경한다.

- HDFS의 파일 시스템 상태 검사

```
$ hdfs fsck /
```

전체 크기, 디렉터리 수, 파일 수, 노드 수 등 파일 시스템의 전체 상태를 보여준다.

```
$ hdfs dfsadmin -report
```

하둡 파일시스템의 기본 정보 및 통계를 보여준다.

---

- HDFS에 저장된 파일을 로컬 파일시스템으로 가져오기

```
$ hdfs dfs -get /tmp/Sample2.txt
```

그럼 로컬의 /home/pilot-pjt/ 디렉터리에 Sample2.txt 파일이 생성된다.

- HDFS의 저장한 파일 삭제(휴지통)

```
$ hdfs dfs -rm /tmp/Sample2.txt
```

삭제 명령을 실행하면 우선 휴지통에 임시 삭제되며, 복구가 가능하다. 휴지통으로 임시 삭제된 파일은 특정 시간(24시간, CDH 기준)이 지나면 자동으로 완전 삭제된다. 휴지통에 임시 삭제가 필요 없을 때는 -skipTrash 옵션을 이용한다.

## Tip \_ HDFS 파일의 비정상 상태

---

HDFS 점검 명령을 실행할 때 하둡의 파일시스템에 문제가 발생할 경우 “CORRUPT FILES”, “MISSING BLOCKS”, “MISSING SIZE”, “CORRUPT BLOCKS” 등의 항목에 숫자가 표기된다. 이 같은 상태가 지속되면 하둡은 물론 HBase, 하이브 등에 부분적인 장애가 발생할 수 있다. 특히 강제 셧다운이 빈번하고 리소스가 부족한 테스트 환경에서 자주 발생할 수 있는 현상이다. 원래 HDFS는 비정상적인 파일 블록을 발견할 경우 다른 노드에 복구하려고 시도하며, 다음과 같이 사용자가 직접 삭제/이동 명령으로 조치할 수 있다.

안전 모드 상태로 전환됐다면 강제로 안전 모드를 해제한다.

```
$ hdfs dfsadmin -safemode leave
```

손상된 파일을 강제로 삭제한다.

```
$ hdfs fsck / -delete
```

손상된 파일을 /lost + found 디렉터리로 옮긴다.

```
$ hdfs fsck / -move
```



## 2.5 빅데이터 클러스터 구성

 HDFS/주키퍼 명령을 이용한 설치 확인

# 주키퍼 명령어 사용

- PuTTY 프로그램 실행
- Server02에 root 계정으로 SSH 접속
- zookeeper-client 실행

Server02에 root 계정으로 로그인한 후 다음 명령을 실행한다.

```
$ zookeeper-client
```

- 주키퍼 Z노드 등록/조회/삭제

주키퍼의 Z노드 생성과 조회, 삭제가 정상적으로 수행되는지 확인한다.

```
$ [zk: localhost:2181(CONNECTED) 0]: create /pilot-pjt bigdata
```

```
$ [zk: localhost:2181(CONNECTED) 0]: ls /
```

```
$ [zk: localhost:2181(CONNECTED) 0]: get /pilot-pjt
```

```
$ [zk: localhost:2181(CONNECTED) 0]: delete /pilot-pjt
```

---

## 2.5 빅데이터 클러스터 구성

 HDFS/주키퍼 명령을 이용한 설치 확인

# HDFS 및 주키퍼 명령어 사용 (실습)