|  |
| --- |
| SeokRae Kim  2017-7-13 |

|  |
| --- |
| FICS |
| Hadoop |
| 설치 가이드 |

목 차

[**I.** **하둡 살펴보기** 1](#_Toc488069790)

[**1.** **하둡(Hadoop)이란?** 1](#_Toc488069791)

[**2.** **하둡 프로젝트 설치 프로그램** 1](#_Toc488069792)

[**II.** **하둡 프로젝트 개발 준비** 2](#_Toc488069793)

[**2.** **VMware 설정** 2](#_Toc488069794)

[**1)** **OS(Linux) 설치** 2](#_Toc488069795)

[**3.** 하둡 실행 계정 생성 18](#_Toc488069796)

[**4.** 호스트 수정 18](#_Toc488069797)

[**5.** 인코딩(Encoding)설정 18](#_Toc488069798)

[**6.** 자바 설치 18](#_Toc488069799)

[**1)** 자바 다운 18](#_Toc488069800)

[**2)** 자바 설치 18](#_Toc488069801)

[**3)** 자바 Path 설정 18](#_Toc488069802)

[**7.** VMware PC 4대 만들기 18](#_Toc488069803)

[**1)** Mac Address 설정 18](#_Toc488069804)

[**2)** Ip 수정 18](#_Toc488069805)

[**3)** 호스트 수정 18](#_Toc488069806)

[**8.** SSH 설정 19](#_Toc488069807)

[**9.** 하둡(Hadoop) 설치 및 설정 19](#_Toc488069808)

[**1)** 하둡 다운 19](#_Toc488069809)

[**2)** 압축 풀기 및 경로 설정 19](#_Toc488069810)

[**3)** 하둡 설정파일 수정 19](#_Toc488069811)

[**4)** 하둡 Path 설정 19](#_Toc488069812)

[**5)** 하둡 실행 19](#_Toc488069813)

[**6)** WordCount 예제 실행 19](#_Toc488069814)

[**III.** 이클립스 연동 20](#_Toc488069815)

1. **하둡 살펴보기**
   1. **하둡(Hadoop)이란?**

* 하둡은 대용량 데이터를 분산 처리할 수 있는 자바 기반의 오픈 소스 프레임워크이다.
* 하둡은 분산 파일 시스템인 HDFS(Hadoop Distributed File System)에 데이터를 저장하고, 분산 처리 시스템인 맵리듀스(MapReduce)를 이용해 데이터를 처리 한다.
* 하둡은 여러 대의 서버에 데이터를 저장하고, 데이터가 저장된 각 서버에서 동시에 데이터를 처리하는 방식이다.
  1. **하둡 프로젝트 설치 프로그램**

|  |  |  |  |
| --- | --- | --- | --- |
| **프로그램** | **종류** | **버전** | **사이트** |
| **Virtual Machine** |  | 12.0 | <https://www.vmware.com/products/workstation/workstation-evaluation.html> |
| **Linux** |  | 7.3(1611) | <http://isoredirect.centos.org/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1611.iso> |
| **JAVA** | java에 대한 이미지 검색결과 | 1.8.0\_131 | <http://www.oracle.com/technetwork/java/javase/downloads/index-jsp-138363.html> |
| **Hadoop** |  | 1.2.1 | <http://hadoop.apache.org/releases.html> |
| **Eclipse** |  | Oxygen | <https://www.eclipse.org/downloads/?> |
| **Maven** |  | 3.5.0 | <https://maven.apache.org/download.cgi> |

1. **하둡 프로젝트 개발 준비**
   1. **실행 모드 설정  
      - 가상 분산(Pseudo-distributed) 모드**하나의 장비에 모든 하둡 환경설정을 하고, 하둡 서비스도 이 장비에서만 제공하는 방식, HDFS와 맵리듀스와 관련된 데몬을 하나의 장비에서만 실행하게 된다. 주로 하둡을 처음 공부하는 분들이 이와 같은 방식으로 테스트 환경을 구성한다.  
      **- 완전 분산(Fully distributed) 모드**  
      여러 대의 장비에 하둡이 설치된 경우, 하둡으로 라이브 서비스를 하게 될 경우 이와 같은 방식으로 구성한다.
   2. **VMware 설정**
      1. **OS(Linux) 설치**

|  |
| --- |
| **#OS 설치하기 – “New Virtual Machine”** |
| C:\Users\kslbs\Desktop\vm 설정\1.png |

|  |
| --- |
| **#설치 방식 선택 – “Typical”** |
| C:\Users\kslbs\Desktop\vm 설정\2.png |

|  |
| --- |
| **#OS Install 방식 선택 – “I will install the operating system later”** |
| C:\Users\kslbs\Desktop\vm 설정\3.png |

|  |
| --- |
| **#OS 종류 선택 – “CentOS 64-bit”** |
| C:\Users\kslbs\Desktop\vm 설정\4.png |

|  |
| --- |
| **#경로 선택 – “C\hadoop” (폴더 생성)** |
| C:\Users\kslbs\Desktop\vm 설정\5.png |
| **#경로 선택 –** “**C:\hadoop\**wibibooks01” |
| C:\Users\kslbs\Desktop\vm 설정\6.png |

|  |
| --- |
| **#Disk Capacity 설정 – “100GB”** |
| C:\Users\kslbs\Desktop\vm 설정\7.png |

|  |
| --- |
| **#Setting확인 – “Finish”** |
| C:\Users\kslbs\Desktop\vm 설정\8.png |

|  |
| --- |
| **#가상 OS사양 설정 - ”Edit virtual Machine Settings”** |
| C:\Users\kslbs\Desktop\vm 설정\9.png |
| **#Memory – “2GB(2048)”** |
| C:\Users\kslbs\Desktop\vm 설정\10.png |

|  |
| --- |
| **#Processors 변경 – “2 Core”** |
| C:\Users\kslbs\Desktop\vm 설정\11.png |

|  |
| --- |
| **#ISO 이미지 찾기 – “Browse…”** |
| C:\Users\kslbs\Desktop\vm 설정\12.png |

|  |
| --- |
| **#ISO 이미지 선택 – “CentOS-7-x86\_64-DVD-1611.iso”** |
| C:\Users\kslbs\Desktop\vm 설정\13.png |

|  |
| --- |
| **#ISO 이미지 넣기 끝** |
| C:\Users\kslbs\Desktop\vm 설정\14.png |

|  |
| --- |
| **#VMware에서 OS 화면 크기조절 – “1024 x 768”** |
| C:\Users\kslbs\Desktop\하둡 설치\vm 설정\15.png |

|  |
| --- |
| **#VMware 설정 끝 – “Power on this virtual machine”** |
| C:\Users\kslbs\Desktop\vm 설정\16.png |

|  |
| --- |
| **#CentOS 설치 – “Install CentOS Linux 7”** |
| C:\Users\kslbs\Desktop\vm 설정\17.png |

|  |
| --- |
| **#CentOS 설치 시작 – “한국어”** |
| C:\Users\kslbs\Desktop\vm 설정\18.png |

|  |
| --- |
| **#설치 요약 메뉴 시작 – “네트워크 및 호스트명”** |
| C:\Users\kslbs\Desktop\vm 설정\25.png |

|  |
| --- |
| **#이더넷 연결 – “끔 -> 켬”** |
| C:\Users\kslbs\Desktop\vm 설정\21.pngC:\Users\kslbs\Desktop\vm 설정\22.png |

|  |
| --- |
| **#설치 요약 – “소프트웨어 선택”** |
| C:\Users\kslbs\Desktop\vm 설정\20.png |
| **#OS 소프트웨어 선택 – “서버 – GUI 사용(DNS, SMTP, FTP)”** |
| C:\Users\kslbs\Desktop\vm 설정\23.png |

|  |
| --- |
| **#서버 – GUI사용(“개발용 도구”)** |
| C:\Users\kslbs\Desktop\vm 설정\24.png |

|  |
| --- |
| **#CentOS 설정 끝 – “설치 시작”** |
| C:\Users\kslbs\Desktop\vm 설정\25.png |

|  |
| --- |
| **#설치 중 – “사용자 설정 메뉴 설정하기”** |
| C:\Users\kslbs\Desktop\vm 설정\28.png |

|  |
| --- |
| **#사용자 설정 메뉴 – “Root 암호 설정”** |
| C:\Users\kslbs\Desktop\vm 설정\26.png |

|  |
| --- |
| **#사용자 설정 메뉴 – “Hadoop 로컬 계정 생성”** |
| C:\Users\kslbs\Desktop\vm 설정\27.png |

|  |
| --- |
| **#CentOS 설정 끝 – “재부팅”** |
| C:\Users\kslbs\Desktop\하둡 설치\vm 설정\29.png |

|  |
| --- |
| **#초기 설정 메뉴** |
| C:\Users\kslbs\AppData\Local\Microsoft\Windows\INetCache\Content.Word\30.png |

|  |
| --- |
| **#라이센스 정보 - ”약관에 동의합니다”** |
| C:\Users\kslbs\Desktop\하둡 설치\vm 설정\31.png |

|  |
| --- |
| **#설정 완료 - ”설치 완료”** |
| C:\Users\kslbs\Desktop\하둡 설치\vm 설정\32.png |

|  |
| --- |
| **#CentOS 7 설치 끝** |
| C:\Users\kslbs\Desktop\하둡 설치\vm 설정\33.png |

* 1. 하둡 실행 계정 생성
  2. **IP 설정**

|  |
| --- |
| **#최종 목표** |
| C:\Users\kslbs\Desktop\하둡 설치\설치 목표\1.png  마스터(네임노드) 하나에 데이터 노드 3개로 완전 분산 모드 테스트 서비스를 구현한다. |

|  |
| --- |
| **#우측 상단 – “Wired Settings”** |
| **C:\Users\kslbs\Desktop\IP 설정\1.png** |

|  |
| --- |
| **#현 네트워크 상태 IPv4를 확인 – “우측 하단 설정 버튼”** |
| **C:\Users\kslbs\Desktop\IP 설정\2.png** |

|  |
| --- |
| **#현 네트워크 설정 상태 – “IPv4 탭”** |
| **C:\Users\kslbs\Desktop\IP 설정\3.png** |

|  |
| --- |
| **#Addresses – “Manual”** |
| **C:\Users\kslbs\Desktop\IP 설정\4.png** |

|  |
| --- |
| **#Addresses – “IP, S/M, G/W, DNS” 수정 -> Apply** |
| **C:\Users\kslbs\Desktop\IP 설정\5.png** |

|  |
| --- |
| **#네트워크 – “OFF -> ON”** |
| **C:\Users\kslbs\Desktop\IP 설정\6.png** |

|  |
| --- |
| **#IP 확인 – “ifconfig”** |
| **C:\Users\kslbs\Desktop\IP 설정\7.png** |

* 1. **호스트 설정**

|  |
| --- |
| **#하둡 테스트 환경 구성** |
| |  |  | | --- | --- | | **호스트명** | **하둡 설치 내용** | | wikibooks01 | 네임노드(NameNode) | | wikibooks02 | 보조네임노드(SecondaryNameNode), 데이터노드(DataNode) | | wikibooks03 | 데이터노드 | | wikibooks04 | 데이터노드 | |

|  |
| --- |
| **#호스트 설정하기** |
| |  | | --- | | **#호스트명 변경** | | [hadoop@localhost ~]$ hostnamectl set-hostname wikibooks01  새 콘솔 열어 변경확인 |   C:\Users\kslbs\Desktop\IP 설정\8.png   |  | | --- | | **#호스트명 정보 파일 변경** | | [hadoop@localhost ~]$ vi /etc/hosts |  |  | | --- | | **#호스트명 정보 파일 변경** | | 192.168.42.120 wikibooks01  192.168.42.121 wikibooks02  192.168.42.122 wikibooks03  192.168.42.123 wikibooks04 | | :wq |   C:\Users\kslbs\Desktop\IP 설정\9.png |

|  |
| --- |
| **#호스트 적용 확인 – “새 콘솔창 켜서 확인”** |
|  |

* 1. 인코딩(Encoding)설정

|  |
| --- |
| **#현 인코딩 상태 확인** |
| [hadoop@wikibooks01 ~]$ echo $LANG  en\_US.utf8 |

|  |
| --- |
| **#인코딩 방식을 변경** |
| [hadoop@wikibooks01 ~]$ su  Password:  [root@wikibooks01 hadoop]# vi /etc/sysconfig/i18n  LANG="ko\_KR.UTF-8"  SUPPORTED="en\_US.UTF-8:en\_US:ko\_KR.eucKR:ko\_KR:ko"  SYSFONT="lat0-sun16"  SYSFONTACM="8859-15" |

|  |
| --- |
| **#인코딩 적용 후 확인** |
| [root@wikibooks01 hadoop]# source /etc/sysconfig/i18n  [root@wikibooks01 hadoop]# echo $LANG  ko\_KR.UTF-8 |
| **#언어관련 파라미터 UTF-8인지 확인 명령어 – “locale”** |
| [root@wikibooks01 hadoop]# locale  LANG=ko\_KR.UTF-8  LC\_CTYPE="ko\_KR.UTF-8"  LC\_NUMERIC=ko\_KR.UTF-8  LC\_TIME=ko\_KR.UTF-8  LC\_COLLATE="ko\_KR.UTF-8"  LC\_MONETARY=ko\_KR.UTF-8  LC\_MESSAGES="ko\_KR.UTF-8"  LC\_PAPER=ko\_KR.UTF-8  LC\_NAME="ko\_KR.UTF-8"  LC\_ADDRESS="ko\_KR.UTF-8"  LC\_TELEPHONE="ko\_KR.UTF-8"  LC\_MEASUREMENT=ko\_KR.UTF-8  LC\_IDENTIFICATION="ko\_KR.UTF-8"  LC\_ALL= |

* 1. 자바 설치
     1. 자바 다운
     2. 자바 설치
     3. 자바 Path 설정
  2. VMware PC 4대 만들기
     1. Mac Address 설정
     2. Ip 수정
     3. 호스트 수정
  3. SSH 설정
  4. 하둡(Hadoop) 설치 및 설정
     1. 하둡 다운

|  |
| --- |
| Wget <http://ftp.daum.net/apache/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz> |

* + 1. 압축 풀기 및 경로 설정
    2. 하둡 설정파일 수정
    3. 하둡 Path 설정
    4. 하둡 실행
    5. WordCount 예제 실행

1. 이클립스 연동