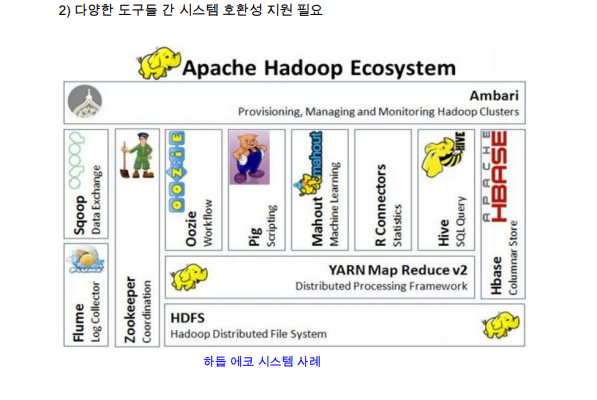
0907



Apache : 대표적 오픈소스 ~~

HDFS : 분산 저장

Map reduce : 분산돼 있는 데이터를 하나로 모아서 (매핑 : 키와 값을 쌍 지어줌)…~~ 하둡의 핵심기술!

Ambari : 모니터해주는 프로그램?

바탕화면-내폴더-다운로드 에 다운 잘 됐는지 확인

프로그램-설정-지역및언어-

한글(hangul) 하나만 추가.- 설정 – 단어단위로 입력 체크 – 토글 키에 hangul 제거

한자키도 f9만 남기고 제거 -적용, 확인

터미널로 와서 shift + 스페이스바 누르면 한글과 영어 변환가능

cd 다운로드

ll

mv haddop(tab키) ../hadoop2.tar.gz //파일 옮기기

tar -zxvf hadoop2.(tab) //압축풀기

ll //확인

rm hadoop2.tar.gz //압축 파일 삭제

y

mv hadoop-2.10.0 hadoop2

gedit /etc/profile

<텍스트 편집기 뜨면 맨 아래에 내용 추가>

export HADOOP\_HOME=/root/hadoop2

export PATH=$PATH:$HADOOP\_HOME/bin:$HADOOP\_HOME/sbin

export HDFS\_NAMENODE\_USER="root"  
export HDFS\_DATANODE\_USER="root"  
export HDFS\_SECONDARYNAMENODE\_USER="root"  
export YARN\_RESOURCEMANAGER\_USER="root"  
export YARN\_NODEMANAGER\_USER="root"

리눅스 껐다 키고 다시 터미널

hadoop version // Hadoop 2.10.0 뜨면 성공

cd hadoop2/share/hadoop

ll

cd mapreduce

ll // .jar로 끝나는 파일들 실행시킬 수 있다.

cd~

hadoop jar $HADOOP\_HOME/share/hadoop/mapreduce/Hadoop-mapreduce--examples-2.10.0.jar wordcount $HADOOP\_HOME/etc/Hadoop/Hadoop-env.sh wordcount\_output //한 줄로 모두 입력 후 실행

gedit ~/wordcount\_output/part-r-00000

cd wordcount\_output

ssh localhost

yes , 1111

껐다가 터미널 다시 키기

ssh localhost

1111

ssh localhost

1111

exit

exit

ls -al // .ssh 폴더 확인

cd .ssh

ssh-keygen -t rsa // 엔터엔터엔터

ll

cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys

ssh localhost // 다시 접속 테스트,비밀번호를 안물어봐야 정상

exit

gedit $HADOOP\_HOME/etc/hadoop/hadoop-env.sh

export JAVA\_HOME=/usr/local/jdk11 // 추가하고 저장 닫기

gedit $HADOOP\_HOME/etc/hadoop/hadoop/core-site.xml

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration> // 추가하고 저장 닫기

gedit $HADOOP\_HOME/etc/hadoop/hdfs-site.xml

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

</configuration> // 추가,저장 닫기

hdfs namenode -format

start-dfs.sh

yes

jps // Jps

NameNode

DataNode

SecondaryNameNode 네개 뜨는 지 확인

stop-dfs.sh