Hadoop & EcoSystem 설치 설명서

문서 History

Version	일시(수정/생성)	수정내용
V1.0	2013-01-16	Hadoop 설치설명서 작성 By 이광복

담당자 지정

부서	담당자	역할 및 업무담당	연락처(개인)
프로프로팀	이광복	Project Manager	koangbok@hotmail.com

- 개인 연락처: 회사용 Email 이 아닌 개인 Email, 또는 핸드폰 번호를 기입함. 차후 퇴사 등의 이유로 연락이 필요한 경유를 대비.

문서의 Detail

- 보관기간 : 영구보존
- 문서보안 : 공개문서

목차

1 설치전 준비작업

1.1 Hadoop Cluster 설치 & Ecosystems 설치 개요

설치순서	프로젝트명	버전	설치경로	주요기능
1	JDK	1.7.0_10	/usr/local/jdk	가상머신 사용하기 위함.
2	hadoop	1.1.1	/usr/local/hadoop	Hadoop 프레임워크
3	zookeeper	3.4.5	/usr/local/zookeeper	서버(Node)상태등 관리
4	Hbase	0.94.4	/usr/local/hbase	Nosql 저장

1.2 소스 다운로드

Linux설치시 프로젝트 버전마다 때로는 설치가 에러날수 있습니다. 매뉴얼을 다시 한번 확인해주시고 버전을 바꾸신 후 설치를 다시 시도해보세요..

- JDK : http://www.oracle.com/technetwork/java/javase/downloads/index.html
- Hadoop : http://apache.mirror.cdnetworks.com/hadoop/core/
- Zookeeper : http://apache.mirror.cdnetworks.com/zookeeper
- Hbase : http://apache.mirror.cdnetworks.com/hbase

1.3 Hadoop JVM 경로(가상머신에 사용한 공간)

프로젝트명	버전	설정경로	주요기능
Hadoop	JVM(가상머신)	/home/hadoop/hdfs	가상머신 데이터 사용 공간지정
Zookeeper		/home/hadoop/zk_data	쥬키퍼 통신 데이터 저장

1.4 서버사양(연구목적 최소사양)

제 개인적인 권장 최소 사양은 5대가 필요하다고 생각합니다. 본 문서는 서버 3대로 설치하며, 이유는 제 개인적으로 운영하는 서버가 3대밖에 없어서 그렇습니다. 하지만 서버 5대일 때 구성을 어떻게 할지 구성도를 공개하오니 잘 맞춰서 세팅 해주시기 바랍니다.

서버수 : 총 3대
OS : CentOS Linux 64bit (32bit용은 추천하지 않습니다.)
CPU : 듀얼 코어
메모리 : 1G
HDD : 최소 30G이상 (500G 이상)

차후 분석할 파일용량이 1 Terrybite 일 경우 메모리 및 HDD가 더 필요함.

2 설치 구성도 및 Pre Setting & 주의사항

2.1 설치순서 및 주의사항

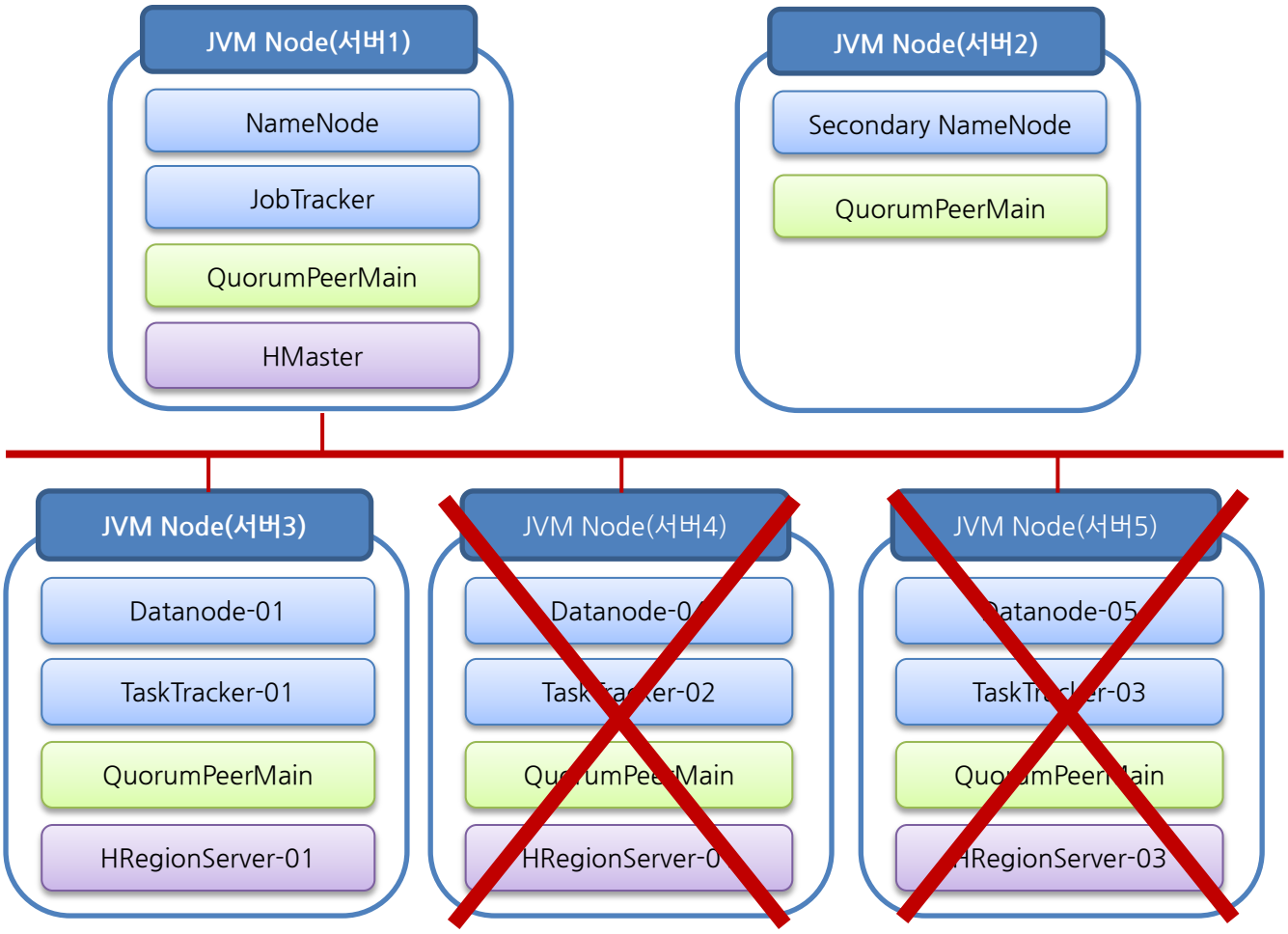
서버 **3대로 설치(node가 3개)**하며 서버4, 서버5는 혹시 서버 5대 이상이시면 설치 가능합니다. 설치문서를 잘 보셔야 하는 부분은 설치 시 전체 서버에 세팅을 하셔야 하는지 서버1에만 설치하셔야 하는지 잘 봐주세요. **네임노드, 세커너리네임노드, 데이터 노드 입니다. 데이터노드는 확장 가능**합니다.

2.2 Hadoop cluster 관리표

구분	IP주소	SSH (필수)	JDK (필수)	Hadoop(Node별구성)			주키퍼	HBase
				NameNode	Secondary	DataNode		
서버1	27.102.206.203	Y	Y	Y			Y	Y
서버2	27.102.207.171	Y	Y		Y		Y	
서버3	27.102.203.39	Y	Y			Y	Y	

2.3 Hadoop 구성도(이미지)

참고사항:**본 문서는 3대(서버1,서버2,서버3)**에 대한 내용을 포함합니다. 서버 5대일경우 서버4, 서버5 형식으로 추가하시면 됩니다. 본 문서는 4,5번은 세팅에 사용 안 하지만 참고해주세요.



2.4 Hosts 지정(3대모두)

```
# 2.2 의 서버 구성도와 비교해서 보세요. 기존 2 줄은 그대로, 마지막 3 줄을 추가했습니다.
# 리눅스는 대/소문자를 확실히 구분합니다. 주의해서 세팅해주세요.
# 해당 IP 는 실제 세팅하시는 분의 IP 와 같이 잘 매칭 시키신 후 진행하세요..
# IP 추가시 아래와 같이 tab 을 활용해서 간격을 조정해주세요.
#27.102.206.203 --탭(tab)-- ns1 --탭(tab)-- zookeeper1 --탭(tab)-- NameNode

[root@ns1 local]# su -
root@ns1's password : *********

[root@ns1 local]# vi /etc/hosts
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1          localhost localhost.localdomain localhost6 localhost6.localdomain6
27.102.206.203          ns1       zookeeper1       NameNode   ← 이줄 추가
27.102.207.171          ns2       zookeeper2       SecondaryNode ← 이줄 추가
27.102.206.39           ns3       zookeeper3       DataNode01 ← 이줄 추가

#잘 되어있나 확인.. 서버 3 대를 모두 똑같이 세팅해야함.
[root@ns1 local]# more /etc/hosts
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1          localhost localhost.localdomain localhost6 localhost6.localdomain6
27.102.206.203          ns1       zookeeper1       NameNode
27.102.207.171          ns2       zookeeper2       SecondaryNode
27.102.206.39           ns3       zookeeper3       DataNode01
```

꼭~~~ 서버 3대에 각 각 똑같이 해야함..

2.5 Path 지정 및 폴더변경(3대모두)

```
[root@ns1 install]# vi /etc/profile
# /etc/profile

# System wide environment and startup programs, for login setup
# Functions and aliases go in /etc/bashrc
----------------------------중간생략-----------------------
unset i
unset pathmunge

# 하단부분에 아래와 같이 세팅
export JAVA_HOME=/usr/local/jdk
export HADOOP_HOME=/usr/local/hadoop
export ZOOKEEPER_HOME=/usr/local/zookeeper
export PATH=$PATH: $JAVA_HOME/bin:$HADOOP_HOME/bin:$ZOOKEEPER_HOME/bin

[root@ns1 install]# source /etc/profile
```

3 SSH 암호화키 설정

3.1 SSH 암호화키 필요성

서버가 여러대(여러node)로 구성되있을때에는 서버의 환경설정들을 다른 서버에 복재합니다. 서버가 100대라면 서버 100대를 일일이 세팅하는것보다 1대에 세팅후 나머지 99대에 설정된 것을 복제하는방법을 사용합니다.이때 ssh(scp)를 통해서 원본파일을 다른 Node에 복재(copy)하게되는데, 이때 ssh로 다른 서버에 접근시에는 지속적으로 암호를 물어본다. ssh 요청시 암호를 물어보게되면 문제가 발생하고 서버와 서버통신시 암호화키를 공유하여 ssh요청시 암호를 안 물어보게 설정하는 Section입니다.

본 문서는 ssh키의 사용자를 root로 하였으며 root로 하였을시 보안상의 문제가 있다고 합니다. 일반 계정(예:hadoop)등으로 scp를 세팅을 시도해보았으나 계속 암호를 물어보는 현상이 발생해 실패하였습니다.

본 암호화키 생성 및 공유의 기본 사용자는 root로 진행합니다.

3.2 암호화키 생성 (서버1번-27.102.206.203)

```
[root@ns1 ~]# cd /root/
[root@ns1 ~]# mkdir .ssh
[root@ns1 ~]# cd .ssh
[root@ns1 .ssh]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa): 그냥 엔터
Enter passphrase (empty for no passphrase): 그냥 엔터
Enter same passphrase again: 그냥 엔터
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
cd:c8:8f:f2:9a:df:ea:c6:ec:9f:a7:53:37:10:69:ce root@ns1
The key's randomart image is:
+--[ RSA 2048]----+
|                 .     |
|                +.     |
|               + .     |
|          . +   E      |
|           S o   .     |
|             o   . o   |
|         .o. .. . .    |
|          ++.....      |
|         o**o=+        |
+---------------------+

[root@ns1 .ssh]# ls -al
total 16
drwxr-xr-x   2 root root 4096 Jan 15 09:58 .
dr-xr-x---. 4 root root 4096 Jan 15 09:57 ..
```

```
-rw-------  1 root root 1675 Jan 15 09:58 id_rsa ←이 파일이 존재해야함
-rw-r--r--  1 root root  390 Jan 15 09:58 id_rsa.pub ←이 파일이 존재해야함
```

3.3 암호화키 생성 (서버2번-27.102.207.171)

```
[root@ns2 ~]# cd /root/
[root@ns2 ~]# mkdir .ssh
[root@ns2 ~]# cd .ssh
[root@ns2 .ssh]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa): 그냥 엔터
Enter passphrase (empty for no passphrase): 그냥 엔터
Enter same passphrase again: 그냥 엔터
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
3a:79:42:a5:be:51:bc:e4:d8:42:68:b3:72:fd:b0:22 root@ns2

[root@ns2 .ssh]# ls -al
total 20
drwxr-xr-x 2 root root 4096 Jan 15 09:57 .
drwxr-x--- 6 root root 4096 Jan 15 09:57 ..
-rw------- 1 root root 1675 Jan 15 09:57 id_rsa  <-이 파일이 존재해야함
-rw-r--r-- 1 root root  390 Jan 15 09:57 id_rsa.pub <-이 파일이 존재해야함
```

3.4 암호화키 생성 (서버3번-27.102.206.39)

```
[root@ns3 ~]# cd /root/
[root@ns3 ~]# mkdir .ssh
[root@ns3 ~]# cd .ssh
[root@ns3 .ssh]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa): 그냥 엔터
Enter passphrase (empty for no passphrase): 그냥 엔터
Enter same passphrase again: 그냥 엔터
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
28:4b:f6:b7:2b:4f:f4:3f:ae:76:60:bd:66:15:d9:e5 root@ns3
[root@ns3 .ssh]# ls -al
total 20
drwxr-xr-x 2 root root 4096 Jan 15 10:05 .
drwxr-x--- 6 root root 4096 Jan 15 10:04 ..
-rw------- 1 root root 1675 Jan 15 10:05 id_rsa <-이 파일이 존재해야함
-rw-r--r-- 1 root root  390 Jan 15 10:05 id_rsa.pub <-이 파일이 존재해야함
```

3.5 암호화키 조합 (서버1번-27.102.206.203)

```
[root@ns1 .ssh]# pwd
/root/.ssh
[root@ns1 .ssh]# cat id_rsa.pub >> authorized_keys
[root@ns1 .ssh]# ls -al
total 20
drwxr-xr-x  2 root root 4096 Jan 15 10:06 .
dr-xr-x---. 4 root root 4096 Jan 15 09:57 ..
-rw-r--r--  1 root root  390 Jan 15 10:06 authorized_keys <-이 파일이 생겨남
-rw-------  1 root root 1675 Jan 15 09:58 id_rsa
-rw-r--r--  1 root root  390 Jan 15 09:58 id_rsa.pub
[root@ns1 .ssh]# more authorized_keys
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAQEAs4lNMLMVav0TPxDSTcTTG5Fve/8b5wkmhHBSXTnJ5qJ8U+kKD
nBO0y5fQ3ggwBj/zlwnoDRFqr030I5UYrjxFUIla0NSli/u611DVcjbXlbv3GWfdILNlu/St3jjxd6v3RerBeWGx
vvRtX4zsIaJ16XGtlFpUcYk+DdnAKIra3HbSA7G46wGM/Uw15DgMEmVzUkswt4ywJ3RKPsRSezxEwrPF
oYtUUZUJK2CJnPma1LMjOug5qdbTuUfbLPNunCgFWAdxjod9Qg67qXbOm2sOIwmU5vdtDgSHJj6AKH
5HlT4vr4bjt9KCrzhCvxYOfj0YzqJxVd1k6rmeBREbfO60w== root@ns1   <-암호화키 1 개 가지고있네요.

[root@ns1 .ssh]# ssh root@SecondaryNode cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
The authenticity of host 'secondarynode (27.102.207.171)' can't be established.
RSA key fingerprint is 41:6a:b2:c4:73:1b:b1:59:37:7a:2b:60:ea:bf:5b:c0.
Are you sure you want to continue connecting (yes/no)? yes
root@secondarynode's password: **********
```

```
[root@ns1 .ssh]# ssh root@DataNode01 cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
The authenticity of host 'datanode01 (27.102.206.39)' can't be established.
RSA key fingerprint is 41:6a:b2:c4:73:1b:b1:59:37:7a:2b:60:ea:bf:5b:c0.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'datanode01,27.102.206.39' (RSA) to the list of known hosts.
root@datanode01's password: ***********

[root@ns1 .ssh]# more authorized_keys
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAQEAs4lNMLMVav0TPxDSTcTTG5Fve/8b5wkmhHBSXTnJ5qJ8U+kKD
nBO0y5fQ3ggwBj/zlwnoDRFqr030l5UYrjxFUlla0NSli/u611DVcjbXlbv3GWfdlLNlu/St3jjxd6v3RerBeWGx
vvRtX4zslaJ16XGtlFpUcYk+DdnAKIra3HbSA7G46wGM/Uw15DgMEmVzUkswt4ywJ3RKPsRSezxEwrPF
oYtUUZUJK2CJnPma1LMjOug5qdbTuUfbLPNunCgFWAdxjod9Qg67qXbOm2sOlwmU5vdtDgSHJj6AKH
5HlT4vr4bjt9KCrzhCvxYOfj0YzqJxVd1k6rmeBREbfO60w== root@ns1 <- 1 번서버암호화키.
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAQEA2OLijAc+LdbT+PxmO8Ufvmq6r7Itdb6rwSlXyPZ0CQyR0bp4h2a
UuxvMbOV8Dn/GQbd7CJa/QyU7PqJuemTenvEC8llxElNt/lcpF4TBXsNNJ7zrNZD1OkffE1uWjent8v9ru+S
YIRRGY6SzQVk1nah1WcL0jPwyhO6ZU1fJZgQG2E8HIL+yrm0fTY4fqtD4jvKlz6/njDWEk1uvXDcbJ5hu2/
rbZ0BQmXHlAbBOLc2Rd8+CpPnFN8GpEO7vYLcg++BDaYO+KeBrwZ/yEhilzVwTNGGpLwGEHo+AXKjP
11HGSGvqAu5sDQiVnkt3GBGY3V3GrUoa4l3K1kAXZSR9dQ== root@ns2 <- 2 번서버암호화키.
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAQEArpLziguVOCg2scGQ1ON0oM5xc9kHF7ZYtpQIK7BxwwaGFa46v
FS2cPm5+Cf1c+mn9mosP1ioNrma+FXqjl0BjpZ5PkFplkvMEvpA4T4FUYmjfWYE44MEzwhEHcW3q5njr
C9cTRvo2zZ0nwvkiVjR+JT/Drki34l1GBtlU7oiK69/rs1MgzRzmHUHDCWW0se5W5J7wpE62xJTpRRrgCP
oLrvLJF5DlaxJSbhimTg8J4YNFqC8W0zFYZiXROWAFCd2Gvp7nFCenaPF0kJaku5WBrvwoc8wz+Qww2C
FjVKXNQpXnG+C+CiZ6iuLNjxa7aDf2YrBQTzRrZXi3Mbc8fcFcQ== root@ns3 <- 3 번서버암호화키

# 위와 같이 서버 1 에 다른서버 2 개의 암호화키를 통합했습니다.
# 이제 할일을 3 개 통합된 암호화키를 서버 2 와 서버 3 에도 나눠 주는 것.
```

3.6 암호화키 배포 (서버1번-27.102.206.39)

```
# 서버 1 -> 서버 2,3 에 암호화 키 배포하기.
[root@ns1 .ssh]# scp -rp authorized_keys root@SecondaryNode:~/.ssh/authorized_keys
root@secondarynode's password: ***********
authorized_keys                                   100% 1170     1.1KB/s   00:00

[root@ns1 .ssh]# scp -rp authorized_keys root@DataNode01:~/.ssh/authorized_keys
root@datanode01's password: ***********
authorized_keys                                   100% 1170     1.1KB/s   00:00
```

3.7 서버2에 암호화키 파일 확인(서버2번-27.102.207.171)

```
[root@ns2 .ssh]# cd /root/.ssh
[root@ns2 .ssh]# ls -al
total 24
drwxr-xr-x 2 root root 4096 Jan 15 10:25 .
drwxr-x--- 6 root root 4096 Jan 15 09:57 ..
-rw-r--r-- 1 root root 1170 Jan 15 10:08 authorized_keys <- 이 파일이 생겨났어야함.
-rw------- 1 root root 1675 Jan 15 09:57 id_rsa
-rw-r--r-- 1 root root  390 Jan 15 09:57 id_rsa.pub100% 1170     1.1KB/s   00:00
```

3.8 암호화키 정상 확인(정상인경우)

```
[root@ns1 .ssh]# ssh DataNode01 date
Tue Jan 15 10:42:26 KST 2013100% 1170     1.1KB/s   00:00

[root@ns1 .ssh]# ssh SecondaryNode date
Tue Jan 15 10:38:27 KST 2013

# 유의사항 ssh 로 2 번,3 번 서버의 date 날짜를 요청했을 때 root 비밀번호를 물어보면 암호화키 설정이 잘됫된것입니다. 처음부터 다시한번 확인해주세요.
```

3.9 암호화키 잘못 세팅확인(잘못 세팅된 경우)

```
[root@ns1 .ssh]# ssh SecondaryNode date
root@secondarynode's password: *********  <- 이렇게 비밀번호를 넣으라고하면 잘못세팅된것임
Tue Jan 15 10:42:26 KST 2013100% 1170     1.1KB/s   00:00
```

3.10 암호화 키가 안되는경우 대처법

1. .ssh 폴더의 권한은 700바꾼다. (chmod 700 /root/.ssh)
2. authorized_keys의 권한은 644로 바꾼다. (chmod 700 /root/.ssh/authorized_keys)

4 Sun JDK설치 (서버 3대 모두)

4.1 주의사항

JDK는 OpenJDK와 Sun JDK가 있는데 OpenJDK가 설치 되 있다면 삭제 후 Sun JDK를 설치요망.

4.2 다운로드

http://www.oracle.com/technetwork/java/javase/downloads/index.html

4.3 압축풀기

```
[root@ns1 install]# tar xvfzp jdk-7u10-linux-x64.tar.gz
jdk1.7.0_10/
jdk1.7.0_10/THIRDPARTYLICENSEREADME-JAVAFX.txt
jdk1.7.0_10/THIRDPARTYLICENSEREADME.txt
jdk1.7.0_10/lib/
jdk1.7.0_10/lib/jexec
jdk1.7.0_10/lib/visualvm/visualvm/modules/
~~~~~~~~~~~~~~~~~~~~~~~~~~이하 생략~~~~~~~~~~~~~~~~~~~~~~~~~~
jdk1.7.0_10/man/ja_JP.UTF-8/man1/javac.1
jdk1.7.0_10/man/ja_JP.UTF-8/man1/tnameserv.1
jdk1.7.0_10/man/ja_JP.UTF-8/man1/pack200.1
jdk1.7.0_10/man/ja_JP.UTF-8/man1/jcmd.1
jdk1.7.0_10/man/ja_JP.UTF-8/man1/jstat.1

[root@ns1 install]# mv jdk1.7.0_10 /usr/local/jdk
```

4.4 JDK 설치 확인

```
# 서버 1,2,3 모두 아래 Command 를 날리시면 똑같이 모두 확인 해야합니다.
[root@ns1 install]# java -version
java version "1.7.0_10"
Java(TM) SE Runtime Environment (build 1.7.0_10-b18)
Java HotSpot(TM) 64-Bit Server VM (build 23.6-b04, mixed mode)
```

5 Hadoop 설치(서버 3대 모두)

5.1 주의사항(*필독*)

- Hadoop설치시 본 문서를 copy-paste를 하지마시고 꼭~ 직접 입력하세요. 빨간색으로 보이는 부분이 서버에 직접 command 날리는 명령어 입니다.
- Hadoop설치시 각각 서버에 접속해서 해야되는부분이 있습니다. (Section 5.2와 5.3)
- Hadoop의 설치는 서버 모두 똑 같은 세팅이 필요하며, 본 문서는 서버1(27.102.206.203)에 설치후에 Copy하는 방법으로 작성되었습니다.

  Ex) 서버1 세팅완료 -> 서버2,3에 복재 -> 구동

5.2 Hadoop JVM에 사용할 폴더 생성(서버3대 모두)

```
[root@ns1 conf]# mkdir -p /home/hadoop/hdfs/data
[root@ns1 conf]# mkdir -p /home/hadoop/hdfs/temp
[root@ns1 conf]# mkdir -p /home/hadoop/hdfs/name
```

5.3 Hadoop 다운받기(서버1에서만 다운:27.102.206.203)

http://apache.mirror.cdnetworks.com/hadoop/core/

5.4 Hadoop 설치(서버1에서만 설치:27.102.206.203)

```
[root@ns1 install]# wget http://apache.mirror.cdnetworks.com/hadoop/core/hadoop-1.1.1/hadoop-1.1.1.tar.gz
Resolving apache.mirror.cdnetworks.com... 61.110.198.174
Connecting to apache.mirror.cdnetworks.com|61.110.198.174|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 61522578 (59M) [application/x-gzip]
Saving to: “hadoop-1.1.1.tar.gz”
100%[======================================>] 61,522,578  18.1M/s   in 3.2s
2013-01-14 17:03:50 (18.1 MB/s) - “hadoop-1.1.1.tar.gz” saved [61522578/61522578]
[root@ns1 install]# tar xvfzp hadoop-1.1.1.tar.gz
hadoop-1.1.1/
hadoop-1.1.1/.eclipse.templates/
hadoop-1.1.1/.eclipse.templates/.externalToolBuilders/
hadoop-1.1.1/.eclipse.templates/.launches/
hadoop-1.1.1/bin/
~~~~~~~~~~~~~~~~~~~~~~~~이하 생략~~~~~~~~~~~~~~~~~~~~~~~~
hadoop-1.1.1/src/contrib/ec2/bin/launch-hadoop-master
hadoop-1.1.1/src/contrib/ec2/bin/launch-hadoop-slaves
hadoop-1.1.1/src/contrib/ec2/bin/list-hadoop-clusters
hadoop-1.1.1/src/contrib/ec2/bin/terminate-hadoop-cluster
```

```
[root@ns1 local]# mv hadoop-1.1.1 /usr/local/hadoop
[root@ns1 conf]# cd /usr/local/hadoop/conf
[root@ns1 conf]# ls -al
total 84
drwxr-xr-x  2 root root 4096 Jan 15 11:59 .
drwxr-xr-x 15 root root 4096 Nov 19 19:50 ..
-rw-rw-r--  1 root root 7457 Nov 19 19:50 capacity-scheduler.xml
-rw-rw-r--  1 root root  535 Nov 19 19:50 configuration.xsl
-rw-rw-r--  1 root root  178 Nov 19 19:50 core-site.xml  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root  327 Nov 19 19:50 fair-scheduler.xml
-rw-rw-r--  1 root root 2237 Nov 19 19:50 hadoop-env.sh  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root 1488 Nov 19 19:50 hadoop-metrics2.properties
-rw-rw-r--  1 root root 4644 Nov 19 19:50 hadoop-policy.xml
-rw-rw-r--  1 root root  178 Nov 19 19:50 hdfs-site.xml  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root 4441 Nov 19 19:50 log4j.properties
-rw-rw-r--  1 root root 2033 Nov 19 19:50 mapred-queue-acls.xml
-rw-rw-r--  1 root root  178 Nov 19 19:50 mapred-site.xml  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root   10 Nov 19 19:50 masters  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root   10 Nov 19 19:50 slaves  <- 이 파일 설정할 대상
-rw-rw-r--  1 root root 1243 Nov 19 19:50 ssl-client.xml.example
-rw-rw-r--  1 root root 1195 Nov 19 19:50 ssl-server.xml.example
-rw-rw-r--  1 root root  382 Nov 19 19:50 taskcontroller.cfg
```

5.5 Hadoop 6개 파일 설정(서버1에서만 설정:27.102.206.203)

파일명	conf 경로	파일 설명
core-site.xml	/usr/local/hadoop/conf	Hadoop core 관련 설정
hdfs-site.xml	/usr/local/hadoop/conf	HDFS 관련 설정파일
mapred-site.xml	/usr/local/hadoop/conf	MapReduce 관련 설정파일
masters	/usr/local/hadoop/conf	Master 서버 설정(NameNode)
slaves	/usr/local/hadoop/conf	Slave 서버 설정(DataNode)
hadoop-env.sh	/usr/local/hadoop/conf	Hadoop의 환경변수

5.6 core-site.xml 설정(서버1에서만 설정:27.102.206.203)

```
[root@ns1 conf]# vi /usr/local/hadoop/conf/core-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
        <property>
                <name>fs.default.name</name>
                <value>hdfs://NameNode:9000</value>
```

```
        </property>
        <property>
                <name>hadoop.tmp.dir</name>
                <value>/home/hadoop/hdfs/temp</value>
        </property>
</configuration>
```

5.7 hdfs-site.xml 설정(서버1에서만 설정:27.102.206.203)

```
[root@ns1 conf]# vi /usr/local/hadoop/conf/hdfs-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
        <property>
                <name>dfs.replication</name>
                <value>1</value>
        </property>
        <property>
                <name>dfs.name.dir</name>
                <value>/home/hadoop/hdfs/name</value>
        </property>
        <property>
                <name>dfs.data.dir</name>
                <value>/home/hadoop/hdfs/data</value>
        </property>
</configuration>
```

5.8 mapred-site.xml 설정(서버1에서만 설정:27.102.206.203)

```
[root@ns1 conf]# vi /usr/local/hadoop/conf/mapred-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
        <property>
                <name>mapred.job.tracker</name>
                <value>NameNode:9001</value>
          </property>
          <property>
                <name>mapred.local.dir</name>
                <value>/home/hadoop/hdfs/mapred</value>
          </property>
          <property>
                <name>mapred.system.dir</name>
               <value>/home/hadoop/hdfs/mapred</value>
          </property>
</configuration>
```

5.9 master 설정(서버1에서만 설정:27.102.206.203)

```
# 주의 :SecondaryNameNode 의 정보를 입력한다. NameNode 의 정보 입력하지 말아주세요.
# 주의 : 아래 내용은 /etc/hosts 에 등록되어있는 정보를 넣어야 합니다. 대/소문자 구분해서~~
[root@ns1 conf]# vi /usr/local/hadoop/conf/master
SecondaryNode
```

5.10 slave 설정(서버1에서만 설정:27.102.206.203)

```
# 주의 : 아래 내용은 /etc/hosts 에 등록되어있는 정보를 넣어야 합니다. 대/소문자 구분해서~~
[root@ns1 conf]# vi /usr/local/hadoop/conf/slave
DateNode01
```

5.11 hadoop-env.sh(서버1에서만 설정:27.102.206.203)

```
[root@ns1 conf]# vi /usr/local/hadoop/conf/hadoop-env.sh
# Set Hadoop-specific environment variables here.
# The only required environment variable is JAVA_HOME.  All others are
-------------------------중간 생략 ---------------------------------------
# The scheduling priority for daemon processes.  See 'man nice'.
# export HADOOP_NICENESS=10

#맨 마지막줄에 아래 한줄 추가
export HADOOP_HOME_WARN_SUPPRESS=1
```

5.12 hadoop 설치파일 배포 (서버1에서만 설정:27.102.206.203)

```
[root@ns1 local]# cd /usr/local/

# 서버 1 에 세팅된 환경을 압축합니다. (다른서버 배포목적)
[root@ns1 local]# tar cvfpz hadoop.tar.gz ./hadoop
./hadoop/
./hadoop/hadoop-examples-1.1.1.jar
./hadoop/src/
./hadoop/src/core/
./hadoop/src/core/overview.html
./hadoop/src/core/org/
-------------------------중간 생략 ---------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml

# 서버 1 -> 서버 2 로 복재
[root@ns1 local]# scp -rp hadoop.tar.gz SecondaryNode:/usr/local
hadoop.tar.gz                                 100%   59MB  29.4MB/s   00:02

# 서버 1 -> 서버 3 로 복재
[root@ns1 local]# scp -rp hadoop.tar.gz DataNode01:/usr/local
hadoop.tar. gz                                100%   59MB  14.7MB/s   00:04
```

5.13 서버2에 Hadoop환경 구축(서버2에서만 설정:27.102.207.171)

```
[root@ns2 conf]# cd /usr/local
[root@ns2 conf]# ls -al
total 60440
drwxr-xr-x 14 root   root       4096 Aug 14 18:22 apache
drwxr-xr-x  6 root   root       4096 Jul 27 15:36 apr
drwxr-xr-x  5 root   root       4096 Jul 27 15:37 apr-util
drwxr-xr-x  2 root   root       4096 Aug 10 13:48 bin
```

```
drwxr-xr-x   2 root    root         4096 May 11   2011 etc
drwxr-xr-x   2 root    root         4096 May 11   2011 games
-rw-r--r--   1 root    root     61699246 Jan 15   2013 hadoop.tar.gz <- 이 파일 있어야함.
drwxr-xr-x   2 root    root         4096 May 11   2011 include
drwxr-xr-x   8 thorpe thorpe       4096 Nov 28 21:10 jdk
drwxr-xr-x   2 root    root         4096 May 11   2011 lib
drwxr-xr-x   2 root    root         4096 May 11   2011 lib64
drwxr-xr-x   2 root    root         4096 May 11   2011 libexec
drwxr-xr-x   4 root    root         4096 Sep 10 11:33 mongoDB
drwxr-xr-x   7 root    root         4096 Aug 14 18:44 php
drwxr-xr-x   2 root    root         4096 May 11   2011 sbin
drwxr-xr-x   4 root    root         4096 Oct   5   2011 share
drwxr-xr-x   2 root    root         4096 Jul 20 11:06 src
[root@ns2 local]# tar xvfzp hadoop.tar.gz
./hadoop/
./hadoop/hadoop-examples-1.1.1.jar
./hadoop/src/
./hadoop/src/core/
./hadoop/src/core/overview.html
./hadoop/src/core/org/
--------------------------중간 생략 ---------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml
```

5.14 서버3에 Hadoop환경 구축 (서버3에서만 설정:27.102.206.39)

```
[root@ns3 conf]# cd /usr/local
[root@ns3 conf]# ls -al
total 60440
drwxr-xr-x 14 root    root         4096 Aug 14 18:22 apache
drwxr-xr-x   6 root    root         4096 Jul 27 15:36 apr
drwxr-xr-x   5 root    root         4096 Jul 27 15:37 apr-util
drwxr-xr-x   2 root    root         4096 Aug 10 13:48 bin
drwxr-xr-x   2 root    root         4096 May 11   2011 etc
drwxr-xr-x   2 root    root         4096 May 11   2011 games
-rw-r--r--   1 root    root     61699246 Jan 15   2013 hadoop.tar.gz <- 이 파일 있어야함.
drwxr-xr-x   2 root    root         4096 May 11   2011 include
drwxr-xr-x   8 thorpe thorpe       4096 Nov 28 21:10 jdk
drwxr-xr-x   2 root    root         4096 May 11   2011 lib
drwxr-xr-x   2 root    root         4096 May 11   2011 lib64
drwxr-xr-x   2 root    root         4096 May 11   2011 libexec
drwxr-xr-x   4 root    root         4096 Sep 10 11:33 mongoDB
drwxr-xr-x   7 root    root         4096 Aug 14 18:44 php
drwxr-xr-x   2 root    root         4096 May 11   2011 sbin
```

```
drwxr-xr-x   4 root    root        4096 Oct   5   2011 share
drwxr-xr-x   2 root    root        4096 Jul 20 11:06 src

[root@ns3 local]# tar xvfzp hadoop.tar.gz
./hadoop/
./hadoop/hadoop-examples-1.1.1.jar
./hadoop/src/
./hadoop/src/core/
./hadoop/src/core/overview.html
./hadoop/src/core/org/
-------------------------중간 생략 ---------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml
```

5.15 Hadoop Path 지정 (서버3대모두)

```
[root@ns1 install]# vi /etc/profile
# /etc/profile

# System wide environment and startup programs, for login setup
# Functions and aliases go in /etc/bashrc
----------------------------중간생략-----------------------
unset i
unset pathmunge

# 하단부분에 아래와 같이 세팅
export JAVA_HOME=/usr/local/jdk
export HADOOP_HOME=/usr/local/Hadoop  <-  이부분 한줄 삽입
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin <-  이부분 부분 추가
[root@ns1 install]# source /etc/profile
```

5.16 Hadoop Format하기(서버1이세만 실행)

컴퓨터를 하드 디스크를 사시거나 USB를 사시면, 그걸 이용하기위해서는 먼저 format을 합니다. 하디디스크 초기화라고 생각하셔도 되는데 Hadoop또한 구성시 처음 하실일은 format하는 일입니다.

```
[root@ns1 hdfs]# hadoop namenode -format
13/01/15 15:29:27 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:    host = NameNode/27.102.206.203
STARTUP_MSG:    args = [-format]
STARTUP_MSG:    version = 1.1.1
STARTUP_MSG:    build = https://svn.apache.org/repos/asf/hadoop/common/branches/branch-1.1 -r
```

```
1411108; compiled by 'hortonfo' on Mon Nov 19 10:48:11 UTC 2012
************************************************************/
Re-format filesystem in /home/hadoop/hdfs/name ? (Y or N) Y  <-  여기 Y를 대문자로 입력해주세요.
13/01/15 15:29:29 INFO util.GSet: VM type       = 64-bit
13/01/15 15:29:29 INFO util.GSet: 2% max memory = 19.33375 MB
13/01/15 15:29:29 INFO util.GSet: capacity      = 2^21 = 2097152 entries
13/01/15 15:29:29 INFO util.GSet: recommended=2097152, actual=2097152
13/01/15 15:29:30 INFO namenode.FSNamesystem: fsOwner=root
13/01/15 15:29:30 INFO namenode.FSNamesystem: supergroup=supergroup
13/01/15 15:29:30 INFO namenode.FSNamesystem: isPermissionEnabled=true
13/01/15 15:29:30 INFO namenode.FSNamesystem: dfs.block.invalidate.limit=100
13/01/15 15:29:30 INFO namenode.FSNamesystem: isAccessTokenEnabled=false
accessKeyUpdateInterval=0 min(s), accessTokenLifetime=0 min(s)
13/01/15 15:29:30 INFO namenode.NameNode: Caching file names occuring more than 10 times
13/01/15 15:29:31 INFO common.Storage: Image file of size 110 saved in 0 seconds.
13/01/15 15:29:31 INFO namenode.FSEditLog: closing edit log: position=4,
editlog=/home/hadoop/hdfs/name/current/edits
13/01/15 15:29:31 INFO namenode.FSEditLog: close success: truncate to 4,
editlog=/home/hadoop/hdfs/name/current/edits
13/01/15 15:29:31 INFO common.Storage: Storage directory /home/hadoop/hdfs/name has been
successfully formatted.
13/01/15 15:29:31 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at ns1/27.102.206.203
************************************************************/
[root@ns1 hdfs]#
```

만약 hadoop namenode -format가 실행명령어가 없다고 error이 뜨면 /etc/profile 파일의 path경로를 확인해주세요.

5.17 Hadoop 데몬띄우기(서버1이서만 실행)

```
[root@ns1 bin]# start-all.sh
starting namenode, logging to /usr/local/hadoop/libexec/../logs/hadoop-root-namenode-ns1.out
DataNode01: starting datanode, logging to /usr/local/hadoop/./logs/hadoop-root-datanode-ns3.out
localhost: starting secondarynamenode, logging to /usr/local/hadoop/libexec/../logs/hadoop-root-
secondarynamenode-ns1.out
starting jobtracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-root-jobtracker-ns1.out
DataNode01: starting tasktracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-root-
tasktracker-ns3.out
```

5.18 Hadoop 데몬 확인(서버1)

```
[root@ns1 conf]# jps
30324 Jps
30053 NameNode <-  이게 떠있어야함
30212 JobTracker <-  이게 떠있어야함
```

5.19 Hadoop 데몬 확인(서버2)

```
[root@ns2 conf]# jps
18029 SecondaryNameNode <-  이게 떠있어야함
18076 Jps
```

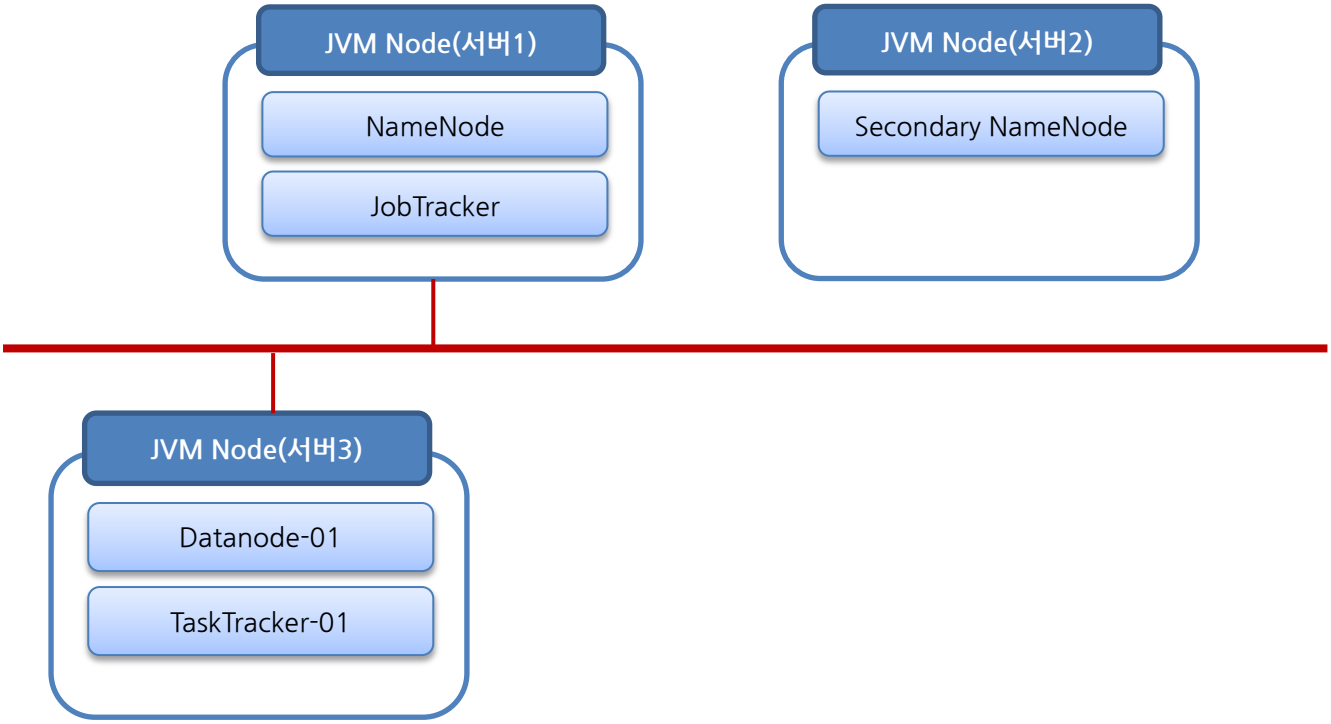
5.20 Hadoop 데몬 확인(서버3)

```
[root@ns3 conf]# jps
28739 Jps
28655 TaskTracker <-  이게 떠있어야함
28558 DataNode <-  이게 떠있어야함
```

5.21 아래 그램처림 세팅되었습니다.

참고로 QuorumPeerMain과 HMaster은 주키퍼와 Hbase가 설치되면 나오는 데몬입니다.

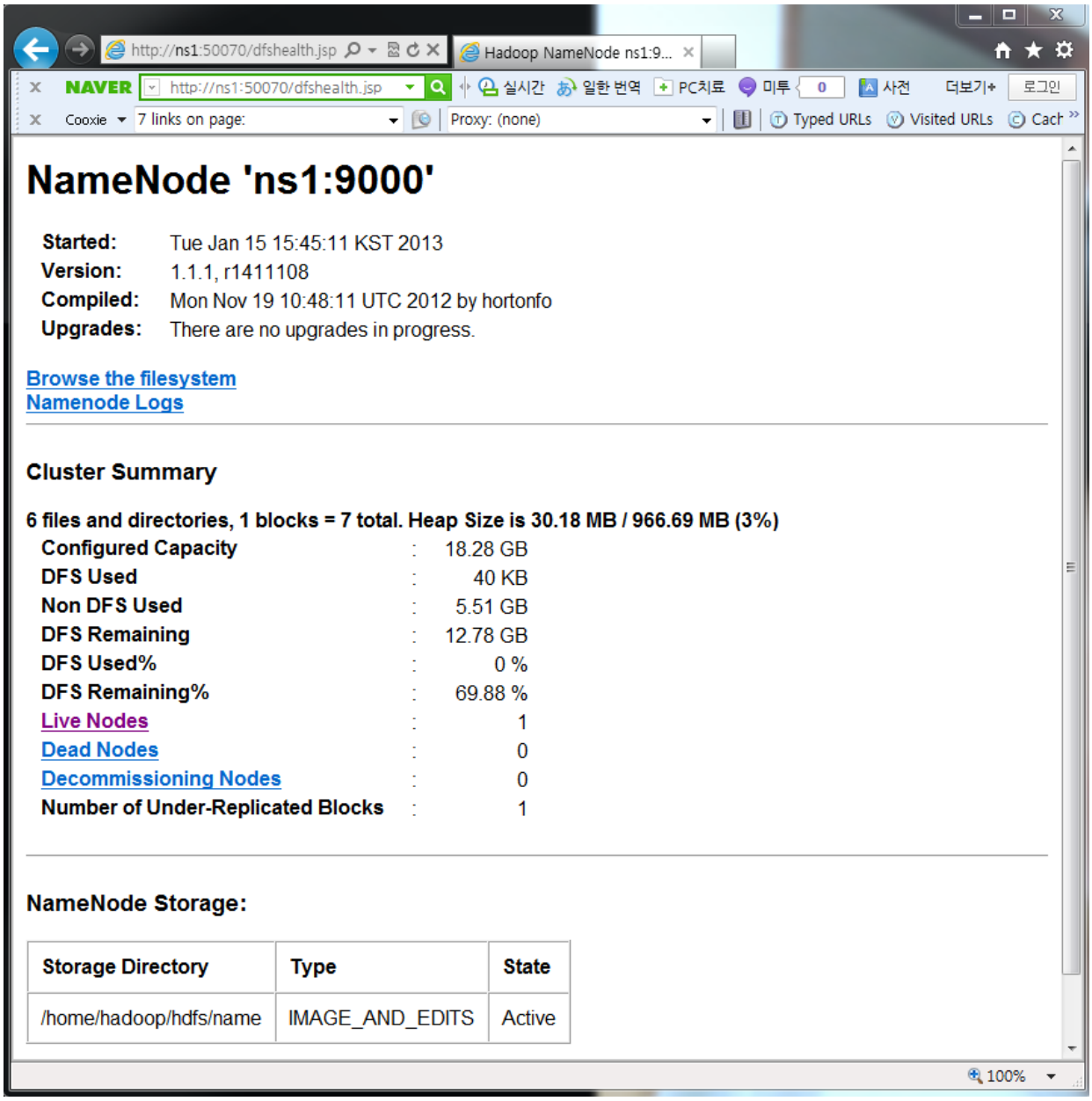
아직 Zoopeeker과 Hbase가 설치 안되어서 없네요..~~



5.22 Hadoop 웹으로 확인

HDFS확인 URL : http://IP:50070

MapReduce 확인 URL 은 http://IP 주소:50030



위에 이미지를 보시면 Live Nodes 가 1 대로 되어있는데 DataNode 의 숫자입니다. DataNode 가 5 대이면 LiveNode(살아있는 서버) + Dead Nodes(죽은서버) = 5 가 되야합니다.

이렇게 해서 서버 3대로 Hadoop Cluters를 구축하였습니다. HDFS 사용법 및 MapReduce 사용법은 저희 카페에서 찾아보세요. (http://cafe.naver.com/hadoopkr)

6 Zooekkeper설치 (서버 3대 모두)

6.1 주의사항(*필독*)

- HBase전에 Zookeeper이 설치완료되있어야 합니다.

6.2 Hosts 확인

```
[root@ns1 local]# more /etc/hosts
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1          localhost localhost.localdomain localhost6 localhost6.localdomain6
27.102.206.203          ns1       zookeeper1       NameNode
27.102.207.171          ns2       zookeeper2       SecondaryNode
27.102.206.39           ns3       zookeeper3       DataNode01
```

6.3 Zookeeper 다운받기(서버1에서만 다운:27.102.206.203)

http://apache.mirror.cdnetworks.com/zookeeper/

6.4 Zookeeper 설치(서버1에서만 설치:27.102.206.203)

```
[root@ns1 install]# wget http://apache.mirror.cdnetworks.com/zookeeper/zookeeper-3.4.5/zookeeper-3.4.5.tar.gz
--2013-01-15 16:15:11--  http://apache.mirror.cdnetworks.com/zookeeper/zookeeper-3.4.5/zookeeper-3.4.5.tar.gz
Resolving apache.mirror.cdnetworks.com... 61.110.198.174
Connecting to apache.mirror.cdnetworks.com|61.110.198.174|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 16402010 (16M) [application/x-gzip]
Saving to: “zookeeper-3.4.5.tar.gz”

100%[======================================>] 16,402,010  19.8M/s   in 0.8s

2013-01-15 16:15:12 (19.8 MB/s) - “zookeeper-3.4.5.tar.gz” saved [16402010/16402010]
```

6.5 Zookeeper 압축풀기(서버1에서만 설치:27.102.206.203)

```
[root@ns1 local]# tar xvfzp zookeeper-3.4.5.tar.gz
-------------------------중간 생략 ---------------------------------------
zookeeper-3.4.5/bin/zkEnv.cmd
zookeeper-3.4.5/bin/zkEnv.sh
zookeeper-3.4.5/bin/zkServer.cmd
zookeeper-3.4.5/bin/zkServer.sh
```

```
[root@ns1 local]# mv zookeeper-3.4.5 /usr/local/zookeeper
[root@ns1 local]# cd /usr/local/zookeeper/conf
[root@ns1 conf]# ls -al
total 12
-rw-r--r-- 1 501 games  535 Oct  1 02:53 configuration.xsl
-rw-r--r-- 1 501 games 2161 Oct  1 02:53 log4j.properties
-rw-r--r-- 1 501 games  808 Oct  1 02:53 zoo_sample.cfg
[root@ns1 conf]# cp -rp zoo_sample.cfg zoo.cfg
```

6.6 Zookeeper 설정파일 수정

```
[root@ns1 local]# cd /usr/local/zookeeper/conf
[root@ns1 conf]# vi zoo.cfg
# The number of milliseconds of each tick
tickTime=2000
# The number of ticks that the initial
# synchronization phase can take
initLimit=10
# The number of ticks that can pass between
# sending a request and getting an acknowledgement
syncLimit=5
# the directory where the snapshot is stored.
# 아래 Data 경로 수정
dataDir=/home/hadoop/zk_data
# the port at which the clients will connect
clientPort=2181

#Zookeeper Servers
# 아래 3 줄 추가
server.1=27.102.206.203:2888:3888
server.2=27.102.207.171:2888:3888
server.3=27.102.206.39:2888:3888
```

6.7 Zookeeper 설치파일 배포 (서버1에서만 설정:27.102.206.203)

```
[root@ns1 local]# cd /usr/local/
[root@ns1 local]# tar cvfpz zookeeper.tar.gz ./zookeeper
zookeeper/
zookeeper/README_packaging.txt
zookeeper/zookeeper-3.4.5.jar.sha1
zookeeper/zookeeper-3.4.5.jar.asc
zookeeper/src/
zookeeper/src/c/
zookeeper/src/c/config.h
--------------------------중간 생략 ----------------------------------------
```

```
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml

# 서버 1 -> 서버 2 로 복재
[root@ns1 local]# scp -rp zookeeper.tar.gz SecondaryNode:/usr/local
zookeeper.tar.gz                                      100%   16MB  29.4MB/s   00:02

# 서버 1 -> 서버 3 로 복재
[root@ns1 local]# scp -rp zookeeper.tar.gz DataNode01:/usr/local
zookeeper.tar. gz                                     100%   16MB  14.7MB/s   00:04
```

6.8 hadoop-env.sh(서버2에서만 설정:27.102.207.171)

```
[root@ns2 conf]# cd /usr/local
[root@ns2 conf]# ls -al
total 60440
drwxr-xr-x 14 root    root         4096 Aug 14 18:22 apache
drwxr-xr-x  6 root    root         4096 Jul 27 15:36 apr
drwxr-xr-x  5 root    root         4096 Jul 27 15:37 apr-util
drwxr-xr-x  2 root    root         4096 Aug 10 13:48 bin
drwxr-xr-x  2 root    root         4096 May 11   2011 etc
drwxr-xr-x  2 root    root         4096 May 11   2011 games
drwxr-xr-x 16 root    root         4096 Jan 15 15:40 hadoop
drwxr-xr-x  2 root    root         4096 May 11   2011 include
drwxr-xr-x  8 thorpe thorpe        4096 Nov 28 21:10 jdk
drwxr-xr-x  2 root    root         4096 May 11   2011 lib
drwxr-xr-x  2 root    root         4096 May 11   2011 lib64
drwxr-xr-x  2 root    root         4096 May 11   2011 libexec
drwxr-xr-x  7 root    root         4096 Aug 14 18:44 php
drwxr-xr-x  2 root    root         4096 May 11   2011 sbin
drwxr-xr-x  4 root    root         4096 Oct   5   2011 share
drwxr-xr-x  2 root    root         4096 Jul 20 11:06 src
-rw-r--r--  1 root    root     16402407 Jan 15 16:30 zookeeper.tar.gz <- 이 파일 있어야함.
[root@ns2 local]# tar xvfzp hadoop.tar.gz
./hadoop/
./hadoop/hadoop-examples-1.1.1.jar
./hadoop/src/
./hadoop/src/core/
./hadoop/src/core/overview.html
./hadoop/src/core/org/
-------------------------중간 생략 -------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
```

```
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml
```

6.9 hadoop-env.sh(서버3에서만 설정:27.102.206.39)

```
[root@ns3 conf]# cd /usr/local
[root@ns3 conf]# ls -al
total 60440
drwxr-xr-x 14 root   root        4096 Aug 14 18:22 apache
drwxr-xr-x  6 root   root        4096 Jul 27 15:36 apr
drwxr-xr-x  5 root   root        4096 Jul 27 15:37 apr-util
drwxr-xr-x  2 root   root        4096 Aug 10 13:48 bin
drwxr-xr-x  2 root   root        4096 May 11   2011 etc
drwxr-xr-x  2 root   root        4096 May 11   2011 games
drwxr-xr-x 16 root   root        4096 Jan 15 15:40 hadoop
drwxr-xr-x  2 root   root        4096 May 11   2011 include
drwxr-xr-x  8 thorpe thorpe       4096 Nov 28 21:10 jdk
drwxr-xr-x  2 root   root        4096 May 11   2011 lib
drwxr-xr-x  2 root   root        4096 May 11   2011 lib64
drwxr-xr-x  2 root   root        4096 May 11   2011 libexec
drwxr-xr-x  7 root   root        4096 Aug 14 18:44 php
drwxr-xr-x  2 root   root        4096 May 11   2011 sbin
drwxr-xr-x  4 root   root        4096 Oct   5   2011 share
drwxr-xr-x  2 root   root        4096 Jul 20 11:06 src
-rw-r--r--  1 root   root    16402407 Jan 15 16:30 zookeeper.tar.gz <- 이 파일 있어야함.
[root@ns3 local]# tar xvfzp hadoop.tar.gz
./hadoop/
./hadoop/hadoop-examples-1.1.1.jar
./hadoop/src/
./hadoop/src/core/
./hadoop/src/core/overview.html
./hadoop/src/core/org/
--------------------------중간 생략 ----------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml
```

6.10 Hadoop Path 지정 (서버3대모두)

```
[root@ns1 install]# vi /etc/profile
# /etc/profile

# System wide environment and startup programs, for login setup
# Functions and aliases go in /etc/bashrc
----------------------------중간생략-----------------------
unset i
unset pathmunge

# 하단부분에 아래와 같이 세팅
export JAVA_HOME=/usr/local/jdk
export HADOOP_HOME=/usr/local/Hadoop
export ZOOKEEPER_HOME=/usr/local/zoopeeker  <-  이부분 한줄 삽입
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$ZOOKEEPER_HOME/bin <- 부분추가
[root@ns1 install]# source /etc/profile
```

6.11 Zookeeper 설정(서버1에서만 설정:27.102.206.203)

```
[root@ns1 local]# mkdir /home/hadoop/zk_data
[root@ns1 local]# more /usr/local/zookeeper/conf/zoo.cfg
# The number of milliseconds of each tick
tickTime=2000
----------------------------중간생략-----------------------
#Zookeeper Servers
server.1=27.102.206.203:2888:3888  <-  이부분에 보시면 server.1 이라고 적혀있는데 myid 에 연관됨.
server.2=27.102.207.171:2888:3888
server.3=27.102.206.39:2888:3888
[root@ns1 local]# cd /home/hadoop/zk_data
[root@ns1 local]# vi myid
1  <-  이부분 넣기
```

6.12 Zookeeper 설정(서버2에서만 설정:27.102.207.171)

```
[root@ns2 local]# mkdir /home/hadoop/zk_data
[root@ns2 local]# more /usr/local/zookeeper/conf/zoo.cfg
# The number of milliseconds of each tick
tickTime=2000
----------------------------중간생략-----------------------
#Zookeeper Servers
server.1=27.102.206.203:2888:3888
server.2=27.102.207.171:2888:3888 <-  이부분에 보시면 server.2 이라고 적혀있는데 myid 에 연관됨.
server.3=27.102.206.39:2888:3888
[root@ns1 local]# cd /home/hadoop/zk_data
[root@ns1 local]# vi myid
2  <-  이부분 넣기
```

6.13 Zookeeper 설정(서버3에서만 설정:27.102.206.39)

```
[root@ns3 local]# mkdir /home/hadoop/zk_data
[root@ns3 local]# more /usr/local/zookeeper/conf/zoo.cfg
# The number of milliseconds of each tick
tickTime=2000
----------------------------중간생략-----------------------
#Zookeeper Servers
server.1=27.102.206.203:2888:3888
server.2=27.102.207.171:2888:3888
server.3=27.102.206.39:2888:3888 <-  이부분에 보시면 server.3 이라고 적혀있는데 myid 에 연관됨.
[root@ns3 local]# cd /home/hadoop/zk_data
[root@ns3 local]# vi myid
3  <-  이부분 넣기
```

6.14 Zookeeper 실행 및 확인(서버1에만)

```
[root@ns1 zookeeper]# zkServer.sh start
JMX enabled by default
Using config: /usr/local/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED

#서버 1 번에 잘 떠있는지 확인
[root@ns1 zookeeper]# jps
30871 Jps
30853 QuorumPeerMain  <- 이데몬이 보이면 성공..
30053 NameNode
30212 JobTracker
[root@ns1 zookeeper]#
```

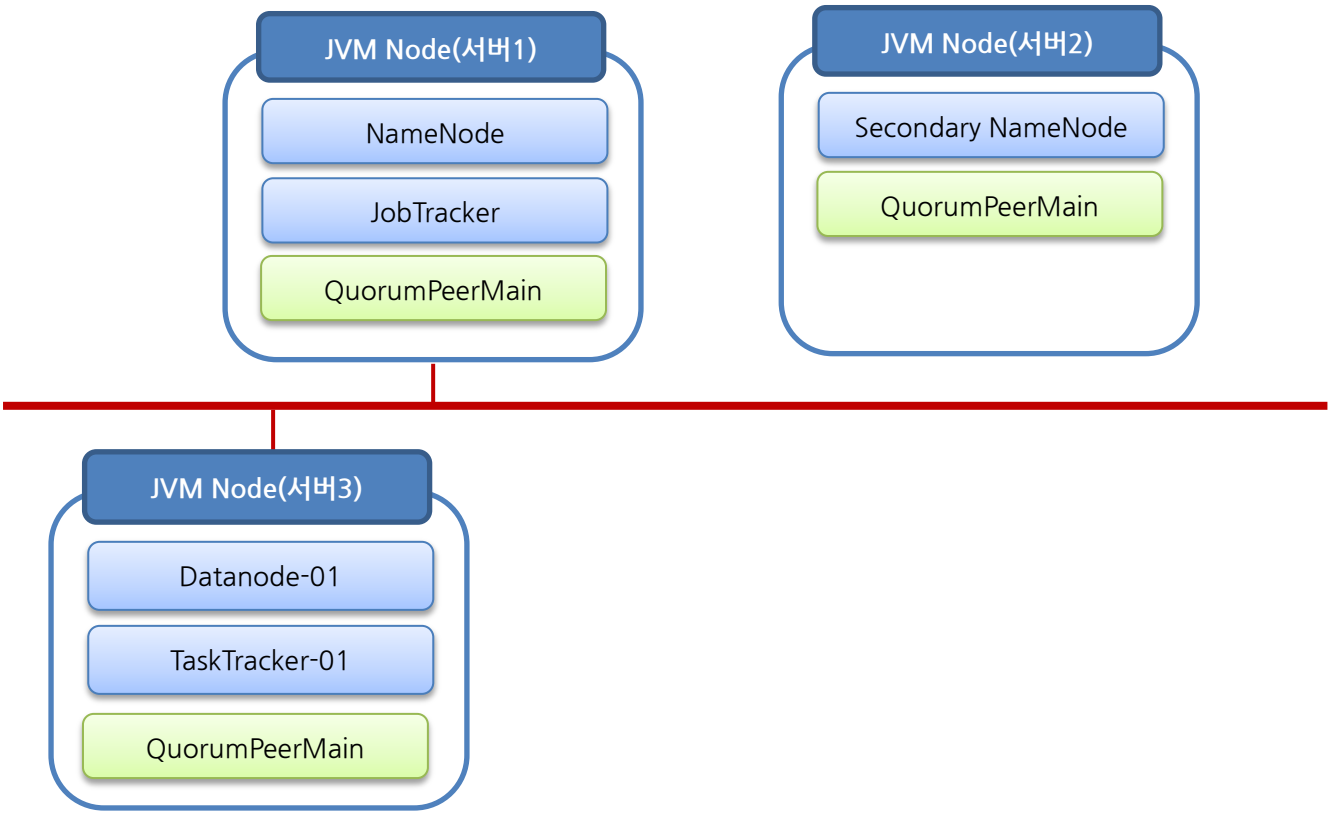
6.15 Zookeeper 데몬 확인(서버2)

```
[root@ns2 zookeeper]# zkServer.sh start
JMX enabled by default
Using config: /usr/local/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
[root@ns2 zookeeper]# jps
18029 SecondaryNameNode
19627 QuorumPeerMain   <- 이데몬이 보이면 성공..
18076 Jps
```

6.16 Zookeeper 데몬 확인(서버3)

```
[root@ns3 zookeeper]# zkServer.sh start
JMX enabled by default
Using config: /usr/local/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
[root@ns3 conf]# jps
28739 Jps
29913 QuorumPeerMain   <- 이데몬이 보이면 성공..
28655 TaskTracker
28558 DataNode
```

6.17 아래 그램처림 세팅되었습니다.



6.18 Zookeeper 접속 및 Simple 테스트

```
[root@ns1 bin]# zkCli.sh
Connecting to localhost:2181
2013-01-16 11:30:57,471 [myid:] - INFO  [main:Environment@100] - Client environment:zookeeper.version=3.4.5-
1392090, built on 09/30/2012 17:52 GMT
2013-01-16 11:30:57,501 [myid:] - INFO  [main:Environment@100] - Client environment:host.name=ns1
2013-01-16 11:30:57,502 [myid:] - INFO  [main:Environment@100] - Client environment:java.version=1.7.0_10
2013-01-16 11:30:57,504 [myid:] - INFO  [main:Environment@100] - Client environment:java.vendor=Oracle Corporation
2013-01-16 11:30:57,505 [myid:] - INFO  [main:Environment@100] - Client environment:java.home=/usr/local/jdk/jre
2013-01-16 11:30:57,507 [myid:] - INFO  [main:Environment@100] - Client
environment:java.class.path=/usr/local/zookeeper/bin/../build/classes:/usr/local/zookeeper/bin/../build/lib/*.jar:/usr/local
/zookeeper/bin/../lib/slf4j-log4j12-1.6.1.jar:/usr/local/zookeeper/bin/../lib/slf4j-api-1.6.1.jar:/usr/local/zookeeper/bin/.
./lib/netty-3.2.2.Final.jar:/usr/local/zookeeper/bin/../lib/log4j-1.2.15.jar:/usr/local/zookeeper/bin/../lib/jline-
0.9.94.jar:/usr/local/zookeeper/bin/../
zookeeper-3.4.5.jar:/usr/local/zookeeper/bin/../src/java/lib/*.jar:/usr/local/zookeeper/bin/../conf:
2013-01-16 11:30:57,509 [myid:] - INFO  [main:Environment@100] - Client
environment:java.library.path=/usr/java/packages/lib/amd64:/usr/lib64:/lib64:/lib:/usr/lib
2013-01-16 11:30:57,511 [myid:] - INFO  [main:Environment@100] - Client environment:java.io.tmpdir=/tmp
2013-01-16 11:30:57,512 [myid:] - INFO  [main:Environment@100] - Client environment:java.compiler=<NA>
2013-01-16 11:30:57,513 [myid:] - INFO  [main:Environment@100] - Client environment:os.name=Linux
2013-01-16 11:30:57,514 [myid:] - INFO  [main:Environment@100] - Client environment:os.arch=amd64
2013-01-16 11:30:57,515 [myid:] - INFO  [main:Environment@100] - Client environment:os.version=2.6.32-
279.14.1.el6.x86_64
2013-01-16 11:30:57,516 [myid:] - INFO  [main:Environment@100] - Client environment:user.name=root
2013-01-16 11:30:57,518 [myid:] - INFO  [main:Environment@100] - Client environment:user.home=/root
2013-01-16 11:30:57,519 [myid:] - INFO  [main:Environment@100] - Client environment:user.dir=/usr/local/zookeeper/bin
2013-01-16 11:30:57,525 [myid:] - INFO  [main:ZooKeeper@438] - Initiating client connection, connectString=localhost:
2181 sessionTimeout=30000 watcher=org.apache.zookeeper.ZooKeeperMain$MyWatcher@30097f5f
Welcome to ZooKeeper!
2013-01-16 11:30:57,633 [myid:] - INFO  [main-SendThread(localhost:2181):ClientCnxn$SendThread@966] -
Opening socket connection to server localhost/0:0:0:0:0:0:0:1:2181. Will not attempt to authenticate using SASL (unknown
error)
2013-01-16 11:30:57,659 [myid:] - INFO  [main-SendThread(localhost:2181):ClientCnxn$SendThread@849] -
Socket connection established to localhost/0:0:0:0:0:0:0:1:2181, initiating session
JLine support is enabled
2013-01-16 11:30:57,819 [myid:] - INFO  [main-SendThread(localhost:2181):ClientCnxn$SendThread@1207] -
Session establishment complete on server localhost/0:0:0:0:0:0:0:1:2181, sessionid = 0x13c411c53740000,
negotiated timeout = 30000

WATCHER::

WatchedEvent state:SyncConnected type:None path:null

[zk: localhost:2181(CONNECTED) 0]    <- 이것까지 뜨면 성공
[zk: localhost:2181(CONNECTED) 0] ls <- 테스트 command
ZooKeeper -server host:port cmd args
        connect host:port
        get path [watch]
        ls path [watch]
        set path data [version]
        rmr path
        delquota [-n|-b] path
        quit
```

```
        printwatches on|off
        create [-s] [-e] path data acl
        stat path [watch]
        close
        ls2 path [watch]
        history
        listquota path
        setAcl path acl
        getAcl path
        sync path
        redo cmdno
        addauth scheme auth
        delete path [version]
        setquota -n|-b val path
[zk: localhost:2181(CONNECTED) 1] get /zookeeper <- 테스트 command
cZxid = 0x0
ctime = Thu Jan 01 09:00:00 KST 1970
mZxid = 0x0
mtime = Thu Jan 01 09:00:00 KST 1970
pZxid = 0x0
cversion = -1
dataVersion = 0
aclVersion = 0
ephemeralOwner = 0x0
dataLength = 0
numChildren = 1
```

7 Hbase 설치 (서버 3대 모두)

7.1 주의사항(*필독*)

- HBase전에 Zookeeper이 설치완료되있어야 합니다.

7.2 Hosts 확인

```
[root@ns1 local]# more /etc/hosts
127.0.0.1     localhost localhost.localdomain localhost4 localhost4.localdomain4
::1           localhost localhost.localdomain localhost6 localhost6.localdomain6
27.102.206.203          ns1       zookeeper1       NameNode
27.102.207.171          ns2       zookeeper2       SecondaryNode
27.102.206.39           ns3       zookeeper3       DataNode01
```

7.3 Hbase 다운받기(서버1에서만 다운:27.102.206.203)

http://apache.mirror.cdnetworks.com/hbase

7.4 Hbase 설치(서버1에서만 설치:27.102.206.203)

```
[root@ns1 install]# wget http://apache.mirror.cdnetworks.com/hbase/hbase-0.94.4/hbase-0.94.4.tar.gz
--2013-01-16 17:26:50--  http://apache.mirror.cdnetworks.com/hbase/hbase-0.94.4/hbase-0.94.4.tar.gz
Resolving apache.mirror.cdnetworks.com... 61.110.198.174
Connecting to apache.mirror.cdnetworks.com|61.110.198.174|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 48845007 (47M) [application/x-gzip]
Saving to: “hbase-0.94.4.tar.gz”

100%[======================================>] 48,845,007   19.4M/s   in 2.4s

2013-01-16 17:26:52 (19.4 MB/s) - “hbase-0.94.4.tar.gz” saved [48845007/48845007
```

7.5 Hbase 압축풀기(서버1에서만 설치:27.102.206.203)

```
[root@ns1 local]# tar xvfzp hbase-0.94.4.tar.gz
hbase-0.94.4/
hbase-0.94.4/src/
hbase-0.94.4/src/packages/
hbase-0.94.4/src/packages/conf-pseudo/
hbase-0.94.4/src/packages/conf-pseudo/hbase-site.xml
hbase-0.94.4/src/packages/build.xml
-------------------------중간 생략 -------------------------------------
hbase-0.94.4/hbase-0.94.4-tests.jar
hbase-0.94.4/NOTICE.txt
hbase-0.94.4/hbase-0.94.4.jar
hbase-0.94.4/sbin/
hbase-0.94.4/sbin/update-hbase-env.sh
[root@ns1 local]# mv hbase-0.94.4 /usr/local/hbase
[root@ns1 local]# cd /usr/local/hbase/conf
[root@ns1 conf]# ls -al
total 36
drwxr-xr-x  2 1002 1002 4096 Jan  3 16:12 .
drwxr-xr-x 10 1002 1002 4096 Jan  3 16:13 ..
-rw-r--r--  1 1002 1002 2971 Jan  3 16:12 hadoop-metrics.properties
-rw-r--r--  1 1002 1002 5395 Jan  3 16:12 hbase-env.sh <- 이 파일 수정
-rw-r--r--  1 1002 1002 2250 Jan  3 16:12 hbase-policy.xml
-rw-r--r--  1 1002 1002  983 Jan  3 16:12 hbase-site.xml <- 이 파일 수정
-rw-r--r--  1 1002 1002 2711 Jan  3 16:12 log4j.properties
-rw-r--r--  1 1002 1002   10 Jan  3 16:12 regionservers
```

7.6 Hbase 설정파일 수정(hbase-site.xml)

```
[root@ns1 local]# vi /usr/local/hbase/conf/hbase-site.xml
-------------------------위부분 생략생략 ---------------------------------------
<configuration>

        <property>
                <name>hbase.rootdir</name>
                <value>hdfs://27.102.206.203:9000/hbase</value>
        </property>

        <property>
                <name>hbase.master</name>
                <value>27.102.206.203:60000</value>
        </property>

        <property>
                <name>hbase.zookeeper.quorum</name>
                <value>27.102.206.203,27.102.207.171,27.102.206.39</value>
        </property>

        <property>
                <name>hbase.zookeeper.property.dataDir</name>
                <value>/home/hadoop/zk_data</value>
        </property>

        <property>
                <name>hbase.cluster.distributed</name>
                <value>true</value>
        </property>

        <property>
                <name>dfs.support.append</name>
                <value>true</value>
        </property>

        <property>
                <name>dfs.datanode.max.xcievers</name>
                <value>4096</value>
        </property>

</configuration>
```

7.7 Hbase_env.sh 설정 (서버1에서만 설정:27.102.206.203)

```
[root@ns1 local]# vi /usr/local/hbase/conf/hbase-env.sh
-------------------------위부분 생략생략 ---------------------------------------
# Tell HBase whether it should manage it's own instance of Zookeeper or not.
```

```
# export HBASE_MANAGES_ZK=true

export JAVA_HOME=/usr/local/jdk                    <- 이줄 추가
export HBASE_CLASSPATH=/usr/local/hadoop/conf      <- 이줄 추가
export HBASE_MANAGES_ZK=true                       <- 이줄 추가


# 서버 1 -> 서버 2 로 복재
[root@ns1 local]# scp -rp hbase.tar.gz SecondaryNode:/usr/local
zookeeper.tar.gz                                   100%   16MB  29.4MB/s   00:02

# 서버 1 -> 서버 3 로 복재
[root@ns1 local]# scp -rp hbase.tar.gz DataNode01:/usr/local
zookeeper.tar. gz                                  100%   16MB  14.7MB/s   00:04
```

7.8 regionserver (서버1에서만 설정:27.102.206.203)

```
#주의사항 : Datanode 정보만 입력한다. Namenode,SecondNamenode 는 미입력
[root@ns1 local]# vi /usr/local/hbase/conf/regionserver
27.102.206.39    <- 이줄 추가
```

7.9 Hbase 설치파일 배포 (서버1에서만 설정:27.102.206.203)

```
[root@ns1 local]# cd /usr/local/
[root@ns1 local]# tar cvfpz hbase.tar.gz ./hbase
hbase/
./hbase/src/
./hbase/src/assembly/
./hbase/src/assembly/all.xml
--------------------------중간 생략 ---------------------------------------
./hadoop/ivy/libraries.properties
./hadoop/ivy/hadoop-streaming-pom-template.xml
./hadoop/ivy/hadoop-client-pom-template.xml
./hadoop/ivy/hadoop-examples-pom-template.xml

# 서버 1 -> 서버 2 로 복재
[root@ns1 local]# scp -rp hbase.tar.gz root@SecondaryNode:/usr/local/
hbase.tar.gz                                       100%   47MB  23.3MB/s   00:02

# 서버 1 -> 서버 3 로 복재
[root@ns1 local]# scp -rp hbase.tar.gz root@DataNode01:/usr/local/
hbase.tar.gz                                       100%   47MB  23.3MB/s   00:02
[root@ns1 local]#
```

7.10 서버2에 Hbase환경 구축(서버2에서만 설정:27.102.207.171)

```
[root@ns2 conf]# cd /usr/local
[root@ns2 conf]# ls -al
drwxr-xr-x.  2 root    root         4096 Jan   9 15:21 bin
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 etc
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 games
drwxr-xr-x  16 root    root         4096 Jan 15 15:33 hadoop
-rw-r--r--   1 root    root     48823075 Jan 16 17:45 hbase.tar.gz <- 이 파일 있어야함.
drwxr-xr-x.  7 root    root         4096 Jan   9 15:21 include
drwxr-xr-x   8 thorpe thorpe       4096 Nov 28 21:10 jdk
drwxr-xr-x.  3 root    root         4096 Jan   9 15:21 lib
drwxr-xr-x.  2 root    root         4096 Jul 10   2012 lib64
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 libexec
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 sbin
drwxr-xr-x.  6 root    root         4096 Jan   9 15:21 share
drwxr-xr-x.  2 root    root         4096 Oct 10 11:30 src
drwxr-xr-x  10      501 games        4096 Jan 16 10:44 zookeeper
 [root@ns2 local]# tar xvfzp hbase.tar.gz
./hbase/
-------------------------중간 생략 -------------------------------------
hbase/docs/images/icon_success_sml.gif
hbase/docs/images/architecture.gif
hbase/docs/index.html
hbase/LICENSE.txt
```

7.11 서버3에 Hbase환경 구축 (서버3에서만 설정:27.102.206.39)

```
[root@ns3 conf]# cd /usr/local
[root@ns3 conf]# ls -al
drwxr-xr-x.  2 root    root         4096 Jan   9 15:21 bin
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 etc
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 games
drwxr-xr-x  16 root    root         4096 Jan 15 15:33 hadoop
-rw-r--r--   1 root    root     48823075 Jan 16 17:45 hbase.tar.gz <- 이 파일 있어야함.
drwxr-xr-x.  7 root    root         4096 Jan   9 15:21 include
drwxr-xr-x   8 thorpe thorpe       4096 Nov 28 21:10 jdk
drwxr-xr-x.  3 root    root         4096 Jan   9 15:21 lib
drwxr-xr-x.  2 root    root         4096 Jul 10   2012 lib64
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 libexec
drwxr-xr-x.  2 root    root         4096 Sep 23   2011 sbin
drwxr-xr-x.  6 root    root         4096 Jan   9 15:21 share
drwxr-xr-x.  2 root    root         4096 Oct 10 11:30 src
drwxr-xr-x  10      501 games        4096 Jan 16 10:44 zookeeper
 [root@ns2 local]# tar xvfzp hbase.tar.gz
./hbase/
-------------------------중간 생략 -------------------------------------
hbase/docs/images/icon_success_sml.gif
```

```
hbase/docs/images/architecture.gif
hbase/docs/index.html
hbase/LICENSE.txt
```

7.12 Hbase 실행 및 확인(서버1)

```
[root@ns1 local]# cd /usr/local/hbase/bin
[root@ns1 bin]# ./start-hbase.sh
The authenticity of host '27.102.206.203 (27.102.206.203)' can't be established.
RSA key fingerprint is c1:4b:6d:9a:b9:da:91:e5:57:71:6e:27:db:a6:01:71.
27.102.206.39: starting zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-root-zookeeper-ns3.out
27.102.207.171: starting zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-root-zookeeper-ns2.out
27.102.206.203: Host key verification failed.
starting master, logging to /usr/local/hbase/bin/../logs/hbase-root-master-ns1.out
27.102.206.39: starting regionserver, logging to /usr/local/hbase/bin/../logs/hbase-root-regionserver-ns3.out
[root@ns1 bin]# jps
31694 Jps
31149 QuorumPeerMain
31565 HMaster  <- 이 이게 뜨면 정상
30053 NameNode
30212 JobTracker
```

7.13 Hbase 데몬 확인(서버2)

```
#주의 : 2 번째 서버는 hbase 가 구동되지 않는다.
[root@ns2 conf]# jps
28739 Jps
28655 TaskTracker
28558 DataNode
```

7.14 Hbase 데몬 확인(서버3)

```
[root@ns3 ~]# jps
30417 HRegionServer  <- 이 이게 뜨면 정상
28655 TaskTracker
29913 QuorumPeerMain
30633 Jps
28558 DataNode
```

7.15 Hbase 접속및 Simple 테스트

```
[root@ns1 local]# cd /usr/local/hbase/bin
[root@ns1 bin]# ./hbase shell
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 0.94.4, r1428173, Thu Jan  3 06:29:56 UTC 2013

hbase(main):001:0> <- 이런 화면이 보이면 정상
hbase(main):002:0> status 'simple'
1 live servers
    ns3:60020 1358339850385
        requestsPerSecond=0, numberOfOnlineRegions=2, usedHeapMB=27, maxHeapMB=983
0 dead servers
Aggregate load: 0, regions: 2
```