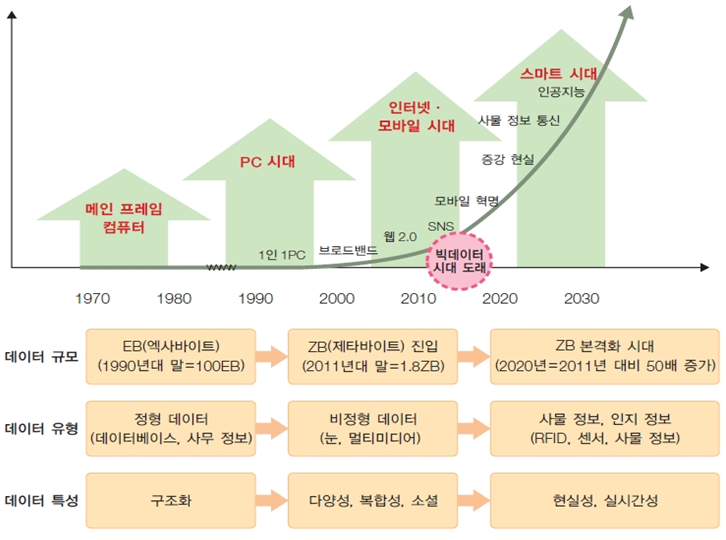
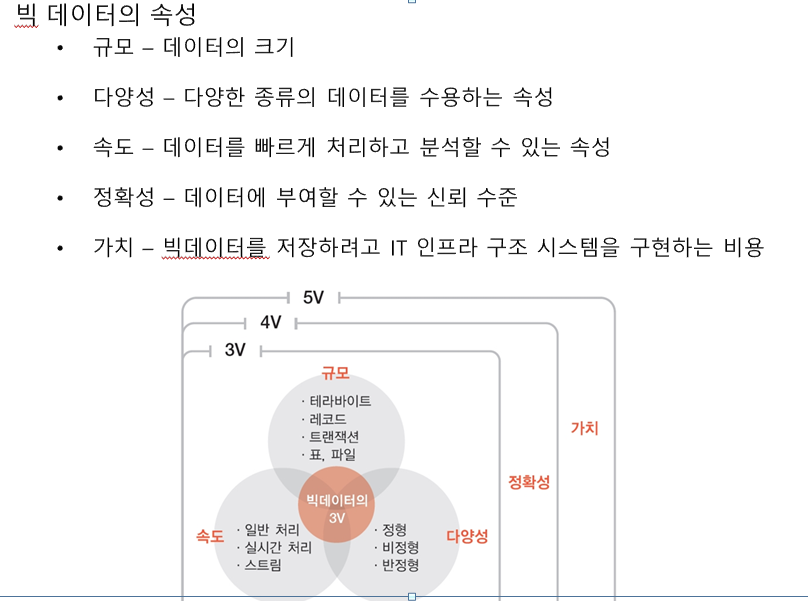
****

**1.volume : 데이터 생성 주기 빠르다**

**2.velocity : 저장 , 처리 속도 빠르다**

**3.variety : sns (각종형태 – 비정형 데이터)**

**정형 데이터 : 타입(길이) primary key**





**하둡 클러스터링(multi node 구성)**

**= 빅데이터 저장(분산) 처리 시스템**

**1. 1개 컴퓨터(vm) : single node 운영**

**2. 4개 컴퓨터 multi clustering node 운영**

**1> vm / 리눅스 설치**

**-yum 업데이트 방지 설정**

**mv /etc/yum.repos.d/CentOS-Base.repo /etc/yum.repos.d/CentOS-Base.repo.bak**

**cd /etc/yum.repos.d**

**wget** [**http://download.hanbit.co.kr/centos/7/CentOS-Base.repo**](http://download.hanbit.co.kr/centos/7/CentOS-Base.repo)

**rm –f \*.repo~**

**yum clean all**

**-selinux 방지 설정**

**gedit /etc/sysconfig/selinux**

**-네트워크**

**HWADDR="00:0C:29:1C:BE:DD"**

**TYPE="Ethernet"**

**BOOTPROTO=none**

**DEFROUTE="yes"**

**PEERDNS="yes"**

**PEERROUTES="yes"**

**IPV4\_FAILURE\_FATAL="no"**

**IPV6INIT="yes"**

**IPV6\_AUTOCONF="yes"**

**IPV6\_DEFROUTE="yes"**

**IPV6\_PEERDNS="yes"**

**IPV6\_PEERROUTES="yes"**

**IPV6\_FAILURE\_FATAL="no"**

**NAME="eno16777736"**

**UUID="a65735bc-8032-4db0-85dd-a8902cecadcf"**

**ONBOOT="yes"**

**IPADDR=192.168.111.110**

**NETMASK=255.255.255.0**

**GATEWAY=192.168.111.2**

**DNS1=192.168.111.2**

**3> 방화벽 해제(root)**

**systemctl status firewalld**

**- openjdk1.7 삭제**

**2> jdk 설치 (root)**

**gedit /etc/profile**

**export JAVA\_HOME=/usr/local/jdk1.8**

**export PATH=$JAVA\_HOME/bin:$PATH**

**source /etc/profile**

**reboot**

**4> hadoop 설치(hadoop)**

**/root/다운로드/\***

**/home/사용자명디렉토리/다운로드/\***

**gedit /etc/sysconfig/network-scripts/ifcfg-eno16777736**

**hostnamectl set-hostname slave2**

**systemctl resrart network**

master vm

192.168.111.110 /master

jdk / hadoop

slave3 vm

192.168.111.150

slave2 vm

192.168.111.140

slave1 vm

192.168.111.130

**<Master:hadoop>**

**ssh-keygen –t rsa**

**master**

**export HADOOP\_HOME=/home/hadoop/hadoop-1.2.1**

**export HADOOP\_HOME\_WARN\_SURPRESS="TRUE"**

**PATH=$HADOOP\_HOME/bin:$PATH**

**2g /1 / 20gb / NAT / ISO**

**언어**

**키보드**

**설치방식**

**소프트웨어파티션**

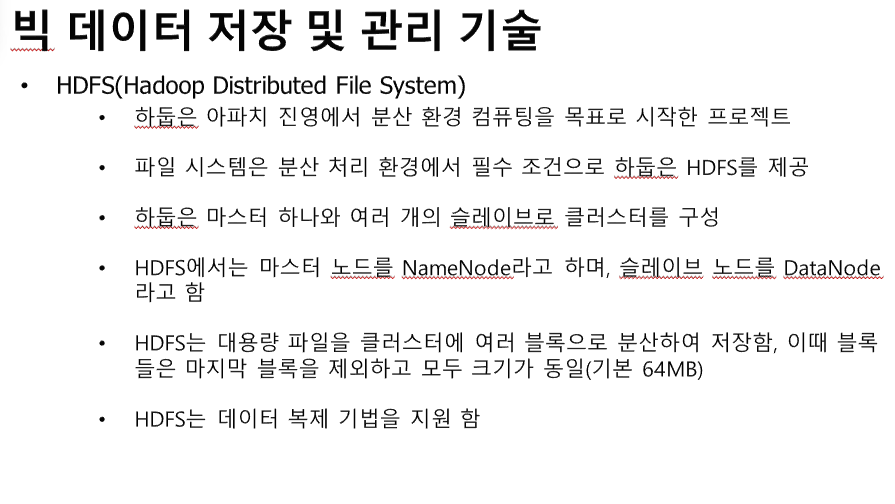
**네트워크**

**1장 빅데이터**

**2장 하둡 설치**

**HDFS ; HADOOP DISTRIBUTED FILE SYSTEM**

**- 빅데이터 : 1장**



**- 하둡 설치 : 2장**

**- 하둡 내부 구조 이해와 하둡 명령어: 3장**

**- 하둡 core 시스템 map/reduce : 4,5,6장**

**- 하둡 에코 시스템 hive : 17장**