

클러스터 실습 환경

순천향대학교 컴퓨터공학과
이 상 정

클러스터 실습 환경

학습 내용

1. 빅데이터 실습 환경
2. VirtualBox 클러스터 서버
3. 윈도우 클라이언트

1. 빅데이터 실습 환경

클러스터 실습 환경

PC 실습 환경

□ PC 에서 클라이언트와 클러스터 서버 실습

□ 클러스터 서버

- VirtualBox 가상머신
 - 2 개의 리눅스 서버
 - 1개의 NAT 게이트웨이
- 하둡, 스파크, 카프카(Kafka), 제플린(Zeppelin) 서버



□ 클라이언트

- ssh 접속 프로그램 한글 Putty
- Notepad++ 코드 편집기



컴퓨터 네트워크 복습: IP 주소 - CIDR

□ CIDR (Classless InterDomain Routing)

- IP 주소의 서브넷 부분이 임의의 길이를 가짐
- 주소 형식: **a.b.c.d/x**, x는 주소의 서브넷 부분의 비트 수



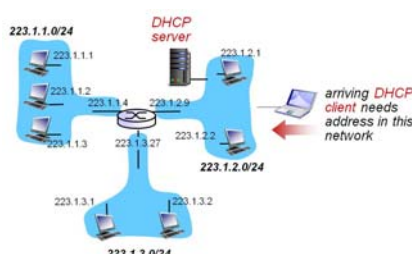
컴퓨터 네트워크 복습: DHCP (Dynamic Host Configuration Protocol)

□ DHCP 목표

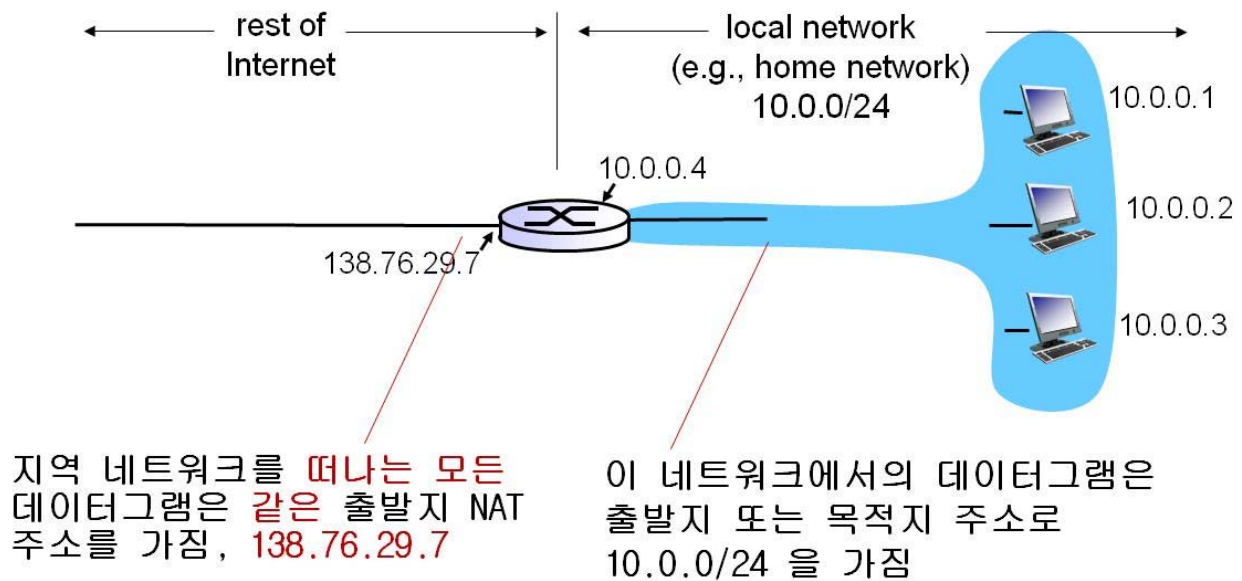
- 호스트가 네트워크에 접속할 때 서버로부터 IP 주소를 **동적으로 획득**
- 네트워크에 연결되었을 때만 주소를 가지므로 **주소의 재사용** 가능
- 짧은 시간 동안에 네트워크에 연결되는 모바일 사용자 지원

□ DHCP 동작 개요

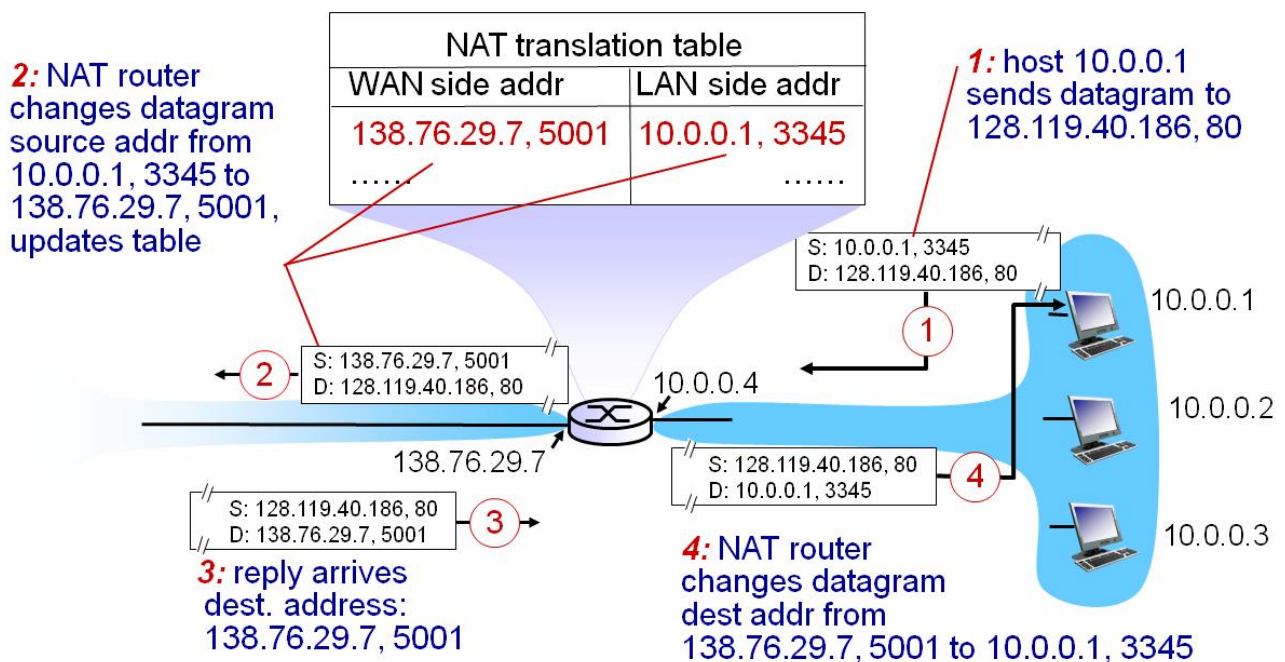
- 호스트는 **DHCP 서버 발견(DHCP discover)** 메시지 발송
- DHCP 서버는 **DHCP 제공(DHCP offer)** 메시지 응답
- 호스트는 **DHCP 요청(DHCP request)** 메시지로 IP 주소 요청
- DHCP 서버는 **DHCP ACK(DHCP ack)** 메시지로 IP 주소 전송



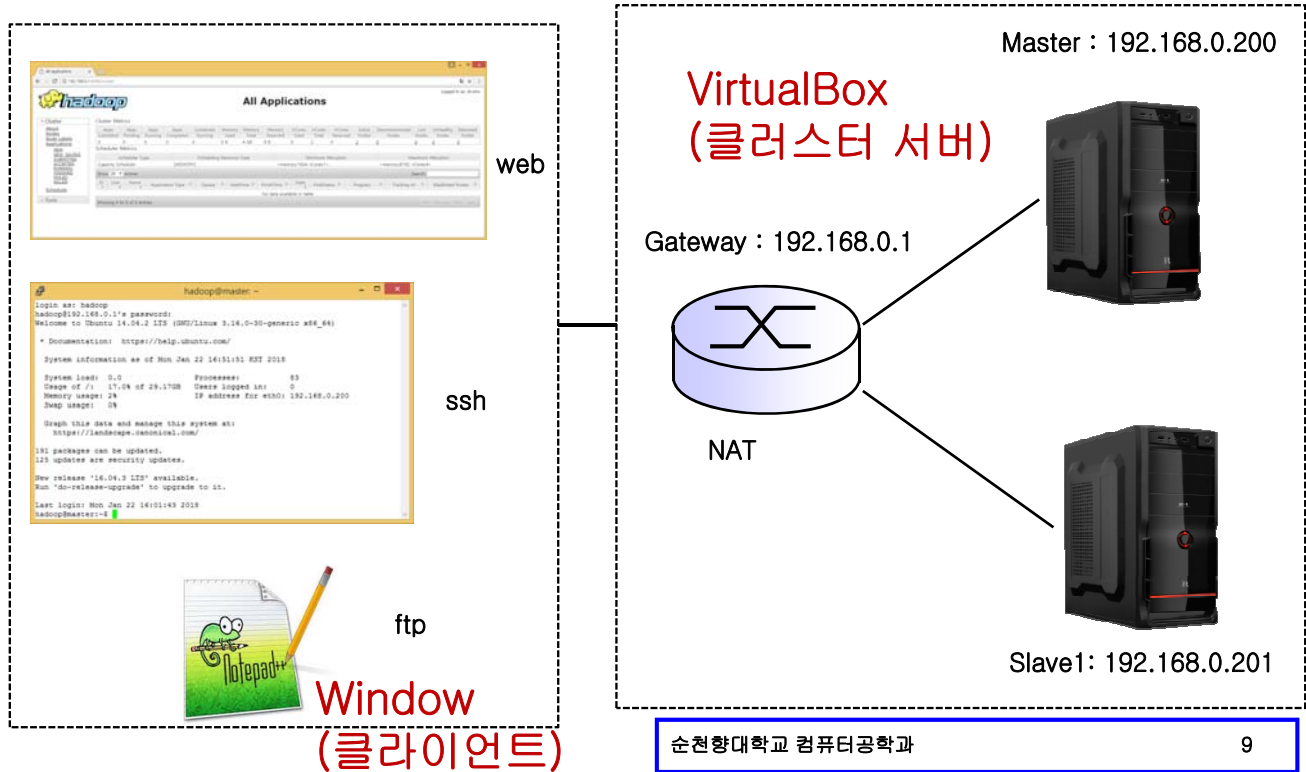
컴퓨터 네트워크 복습: 네트워크 주소 변환, NAT (Network Address Translation)



컴퓨터 네트워크 복습: NAT 사용 예



실습 네트워크 환경



2. VirtualBox 클러스터 서버

VirtualBox 가상 머신 - 다운로드

Virtual Box 다운로드 설치 (설치는 참고자료 참조)

- <https://www.virtualbox.org/wiki/Downloads>
- 최신 버전은 6.1.4 (2020년 3월 기준)
- **!! 주의:** 오라클 라이선스 문제로 확장팩은 사용하지 않음
 - 교내에서 확장팩 사용 시 학과가 라이선스 비용 부담
- 아래는 6.0.8 기준 설치 예

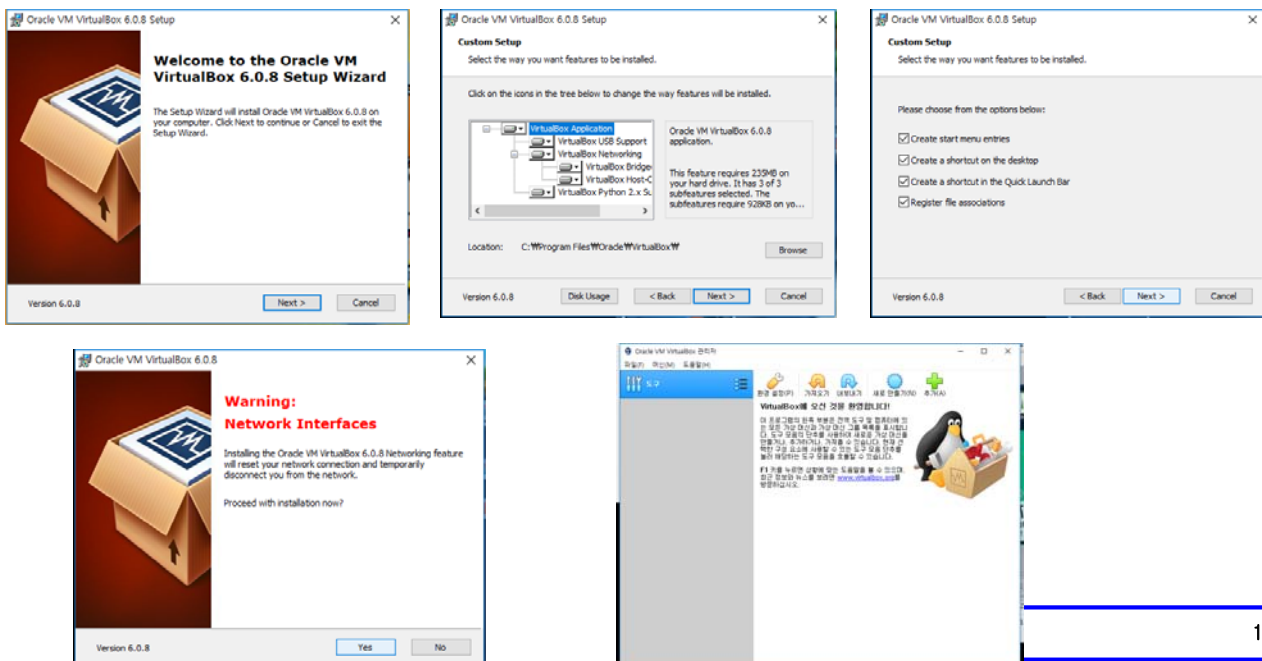


11

VirtualBox 가상 머신 - 설치

다운로드 받은 파일 실행하여 설치

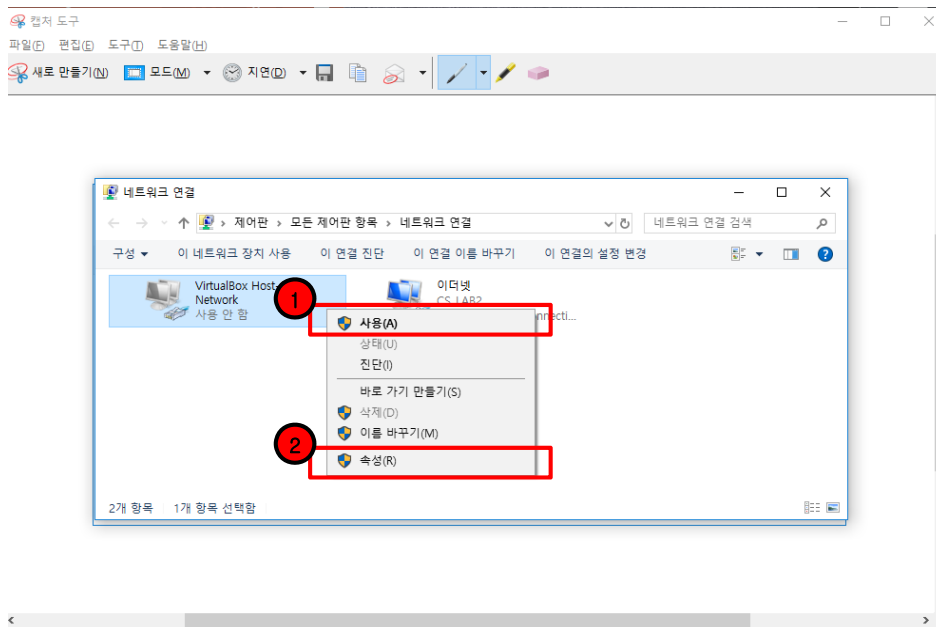
- 디폴트로 Next 버튼 눌러 설치



12

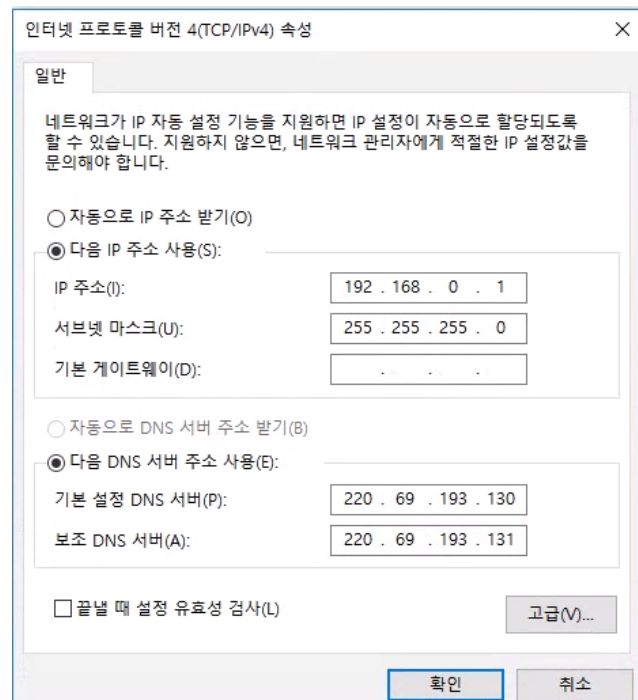
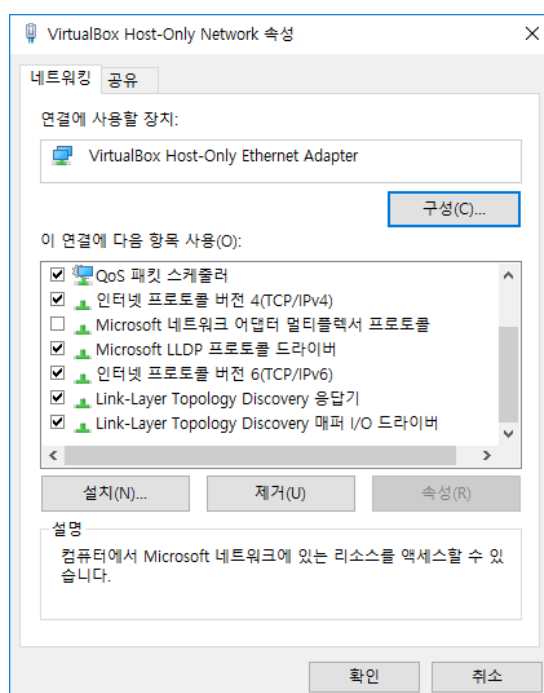
VirtualBox 가상 머신 – 게이트웨이 설정 (1)

- ❑ 윈도우 네트워크 연결->이더넷 설정 -> VirtualBox Host-Only 선택



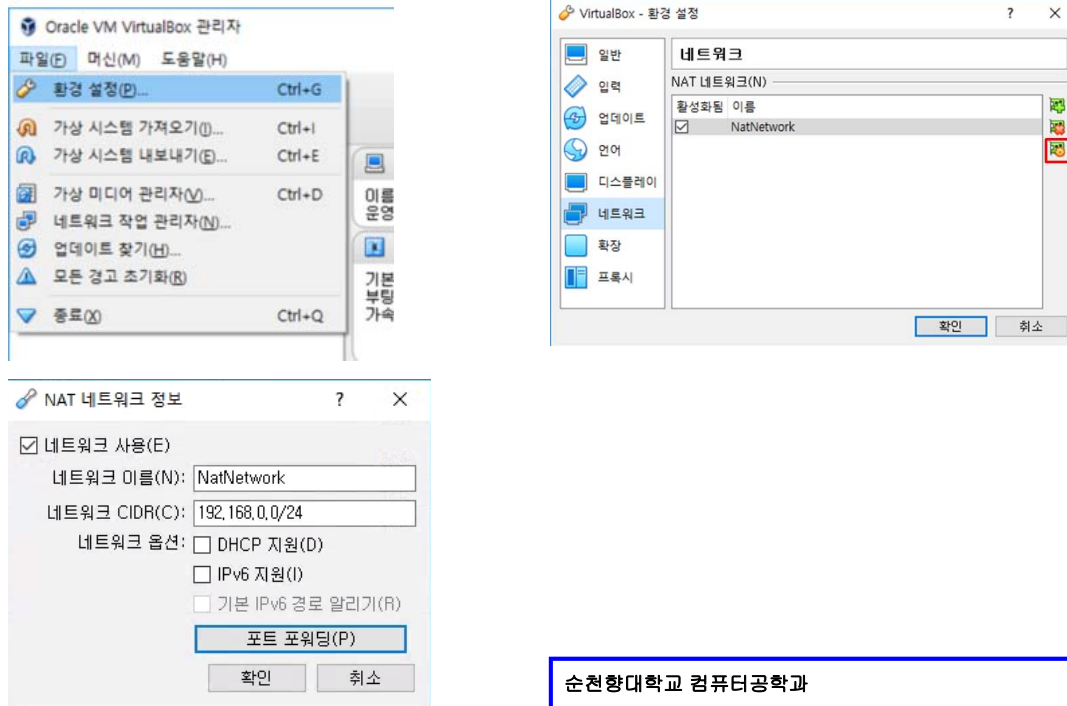
13

VirtualBox 가상 머신 – 게이트웨이 설정 (2)



VirtualBox 가상 머신 - 게이트웨이 NAT 설정 (1)

□ NAT(port forwarding) 테이블 설정



NAT 네트워크 정보

☒ 네트워크 사용(E)
 네트워크 이름(N): NatNetwork
 네트워크 CIDR(C): 192.168.0.0/24
 네트워크 옵션:
☐ DHCP 지원(D)
☐ IPv6 지원(I)
☐ 기본 IPv6 경로 알리기(R)
☒ 포트 포워딩(P)
 확인 취소

VirtualBox - 환경 설정

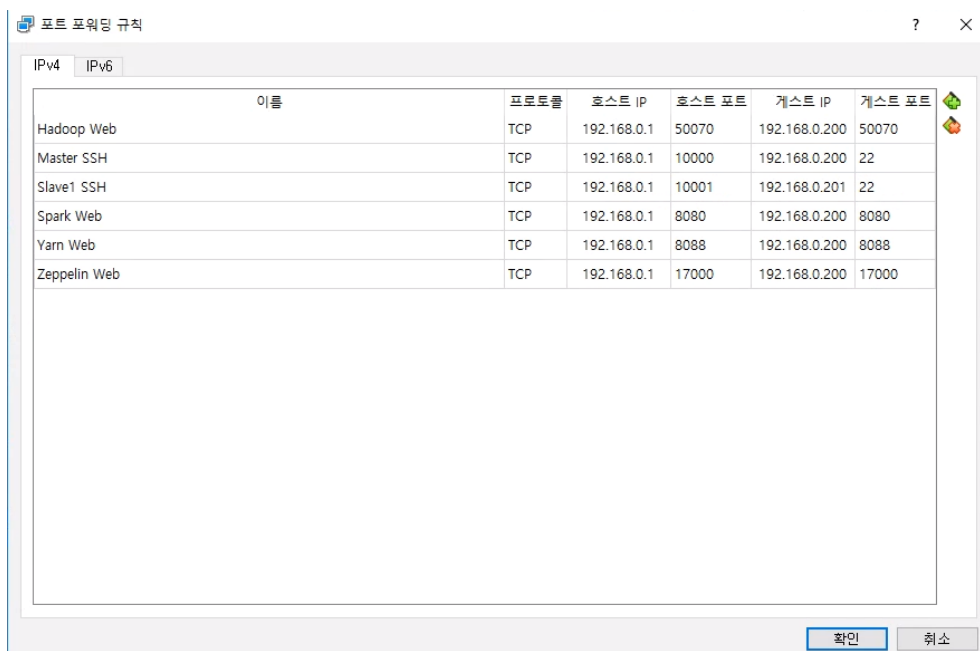
네트워크
 NAT 네트워크(N)
 활성화된 이름
☒ NatNetwork
 확인 취소

순천대학교 컴퓨터공학과 15

VirtualBox 가상 머신 게이트웨이 - NAT 설정 (2)

□ 포트 포워딩(port forwarding) 규칙 설정

- SSH, FTP, Spark, Yarn, Zeppelin



포트 포워딩 규칙

IPv4 IPv6

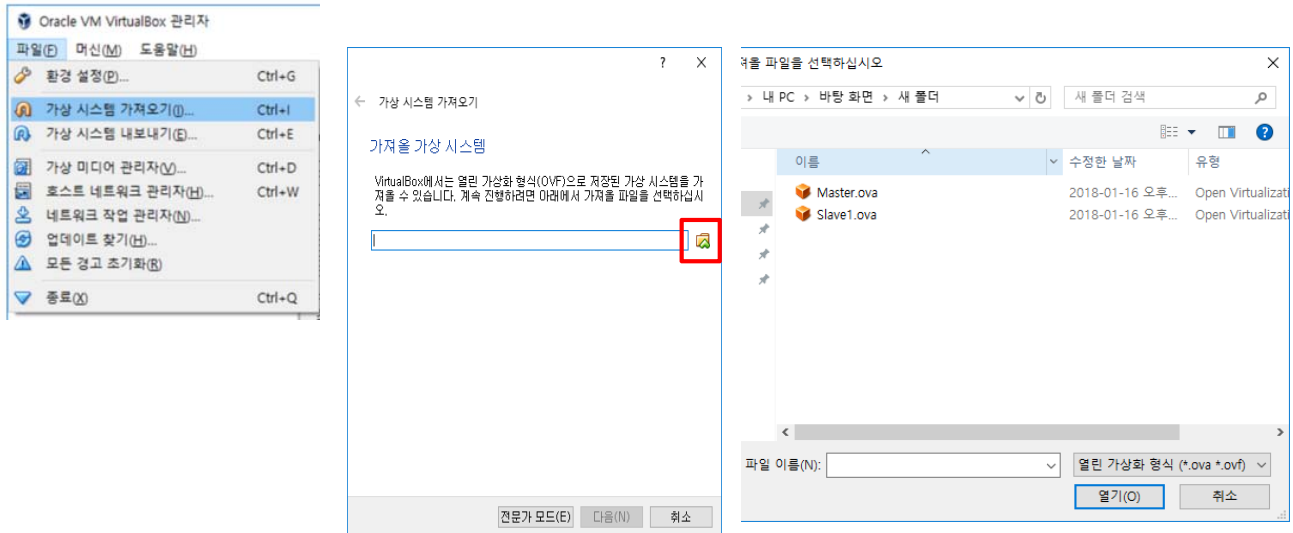
이름	프로토콜	호스트 IP	호스트 포트	게스트 IP	게스트 포트
Hadoop Web	TCP	192.168.0.1	50070	192.168.0.200	50070
Master SSH	TCP	192.168.0.1	10000	192.168.0.200	22
Slave1 SSH	TCP	192.168.0.1	10001	192.168.0.201	22
Spark Web	TCP	192.168.0.1	8080	192.168.0.200	8080
Yarn Web	TCP	192.168.0.1	8088	192.168.0.200	8088
Zeppelin Web	TCP	192.168.0.1	17000	192.168.0.200	17000

확인 취소

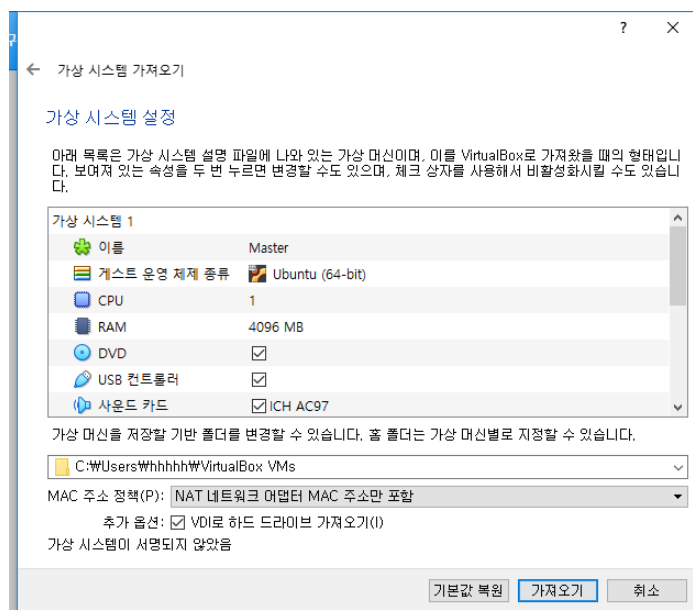
16

VirtualBox 가상 머신 - 가상 머신 가져오기 (1)

- 외부 파일로 저장된 가상 머신의 이미지 가져오기 (import)
 - Master, Slave1 노드의 이미지 가져오기



VirtualBox 가상 머신 - 가상 머신 가져오기 (2)



VirtualBox 가상 머신 - 가상 머신 가져오기 (3)

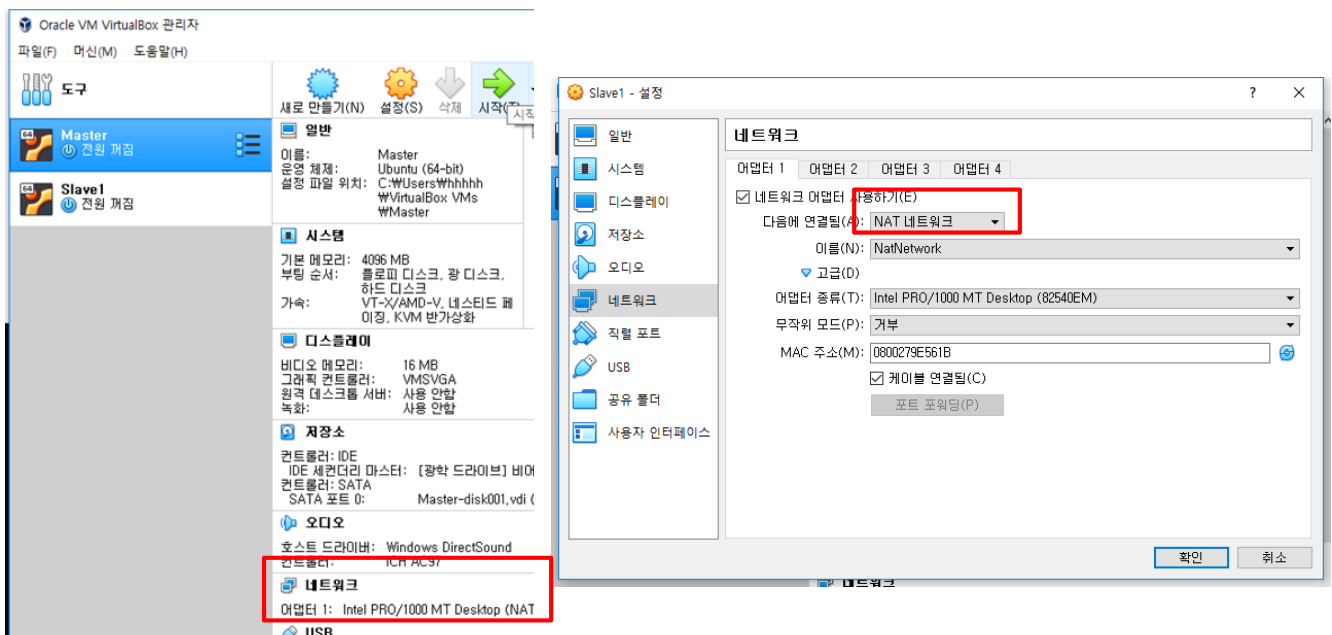
가상 머신 가져오기 결과 확인



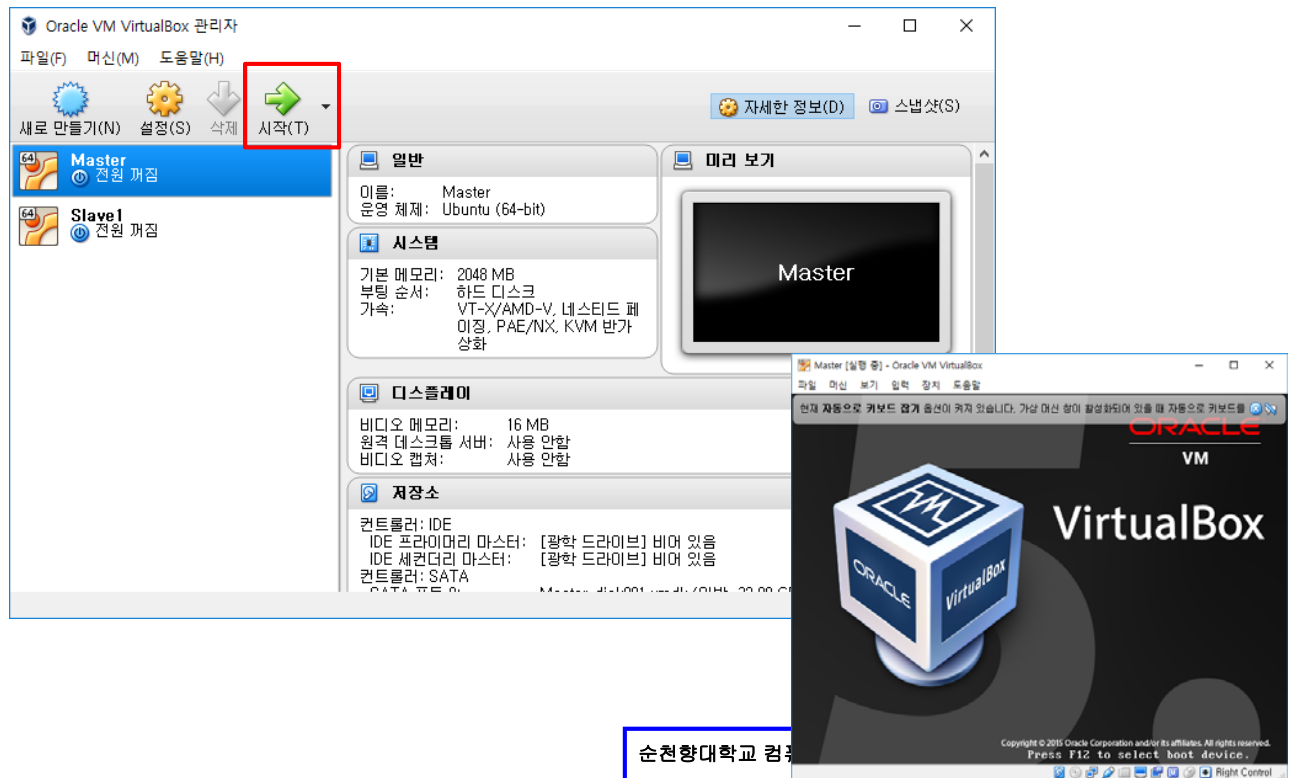
19

VirtualBox 가상 머신 - 네트워크 설정

Master, Slave1 네트워크를 NAT 네트워크로 설정



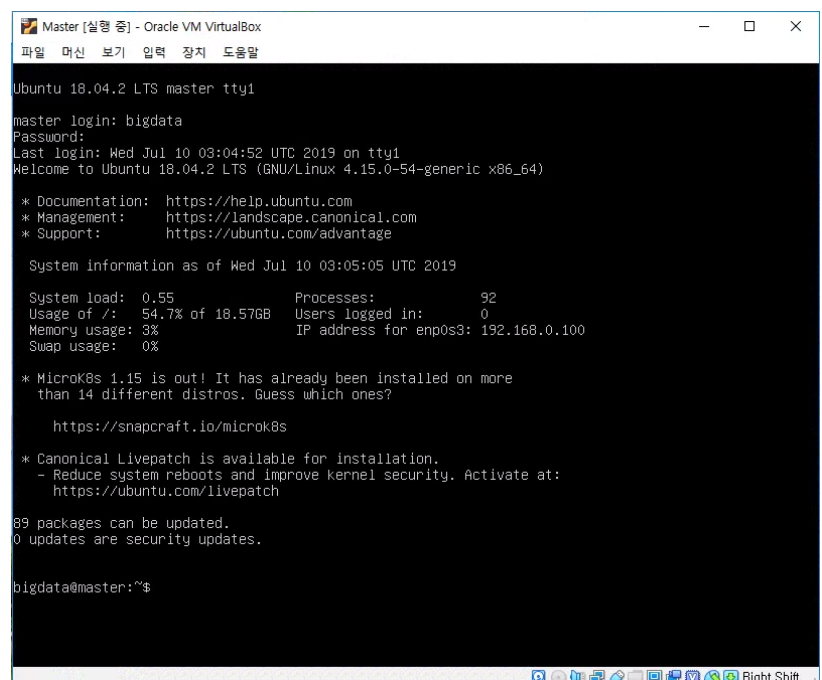
VirtualBox 가상 머신 - 우분투 서버 실행 (1)



VirtualBox 가상 머신 - 우분투 서버 실행 (2)

계정 정보

- 사용자 계정 : **bigdata**
- 비번 : **lecture**



Master 서버 - IP 주소 설정 (1)

□ IP 주소 확인

- \$ ip addr

```
bigdata@master:~$ ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: enp0s3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 08:00:27:83:0e:85 brd ff:ff:ff:ff:ff:ff
    inet 192.168.0.100/24 brd 192.168.0.255 scope global enp0s3
        valid_lft forever preferred_lft forever
    inet6 fe80::a00:27ff:fe83:e85/64 scope link
        valid_lft forever preferred_lft forever
bigdata@master:~$
```

Master 서버 - IP 주소 설정 (2)

□ Master 서버의 IP 주소와 게이트웨이, 네임 서버(DNS 서버)를 설정

- Ubuntu 18.04 LTS 부터 IP 설정 시 **netplan** 사용
 - 설정파일에서 주소 수정
 - 경로 : **/etc/netplan/50-cloud-init.yaml**
- \$ **sudo nano /etc/netplan/50-cloud-init.yaml**

```
Master [실행 중] - Oracle VM VirtualBox
파일  머신  보기  입력  장치  도움말
GNU nano 2.9.3 /etc/netplan/50-cloud-init.yaml

# This file is generated from information provided by
# the datasource.  Changes to it will not persist across an instance.
# To disable cloud-init's network configuration capabilities, write a file
# /etc/cloud/cloud.cfg.d/99-disable-network-config.cfg with the following:
# network: {config: disabled}
network:
  ethernets:
    enp0s3:
      dhcp4: no
      addresses: [192.168.0.200/24]
      gateway4: 192.168.0.1
      nameservers:
        addresses: [220.69.193.130,220.69.193.131]

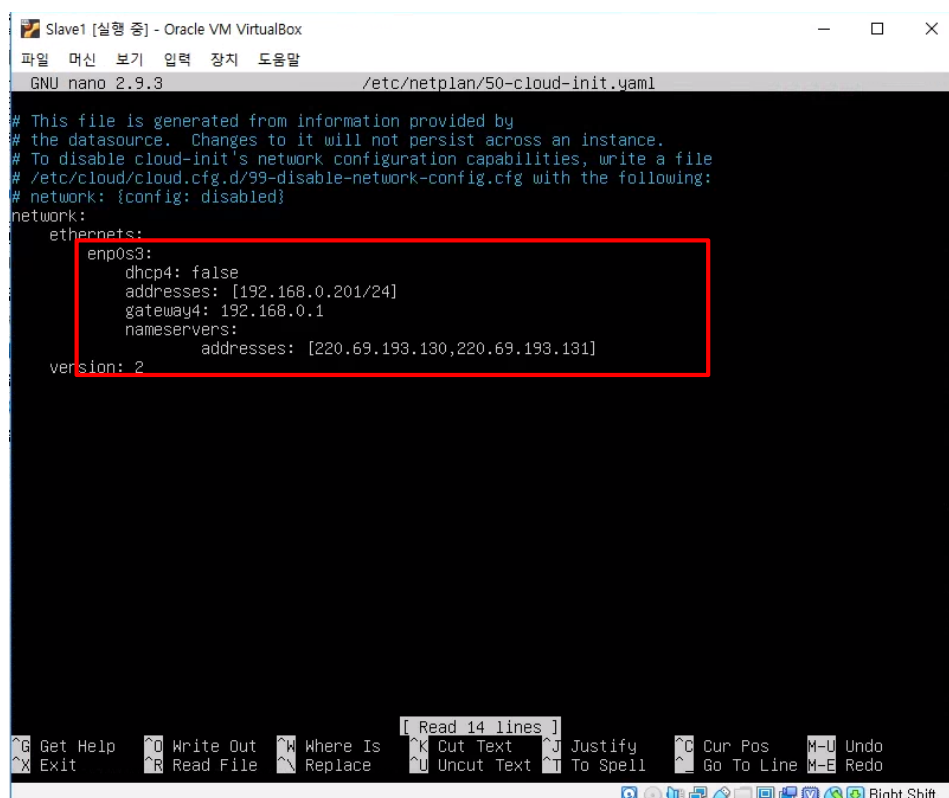
version: 2
```

Master 서버 - IP 주소 설정 (3)

- 수정된 설정 파일 적용
\$ **sudo netplan apply**
- 수정된 IP 주소 확인
\$ **ip addr**

```
bigdata@master:~$ ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: enp0s3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 08:00:27:89:0e:85 brd ff:ff:ff:ff:ff:ff
    inet 192.168.0.200/24 brd 192.168.0.255 scope global enp0s3
        valid_lft forever preferred_lft forever
    inet6 fe80::a00:27ff:fe83:e85/64 scope link
        valid_lft forever preferred_lft forever
bigdata@master:~$
```

Slave1 서버 - IP 주소 설정



```
Slave1 [실행 중] - Oracle VM VirtualBox
파일  머신  보기  입력  장치  도움말
GNU nano 2.9.3 /etc/netplan/50-cloud-init.yaml

# This file is generated from information provided by
# the datasource.  Changes to it will not persist across an instance.
# To disable cloud-init's network configuration capabilities, write a file
# /etc/cloud/cloud.cfg.d/99-disable-network-config.cfg with the following:
# network: {config: disabled}
network:
  ethernets:
    enp0s3:
      dhcp4: false
      addresses: [192.168.0.201/24]
      gateway4: 192.168.0.1
      nameservers:
        addresses: [220.69.193.130,220.69.193.131]
  version: 2

[ Read 14 lines ]
Get Help  Write Out  Where Is  Cut Text  Justify  Cur Pos  M-U  Undo
Exit      Read File  Replace  Uncut Text  To Spell  Go To Line  M-E  Redo
Right Shift
```

3. 윈도우 클라이언트

클러스터 실습 환경

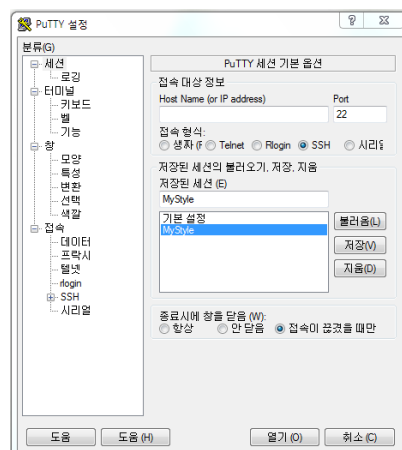
iPuTTY

□ iPuTTY

- 가상 머신 서버를 ssh로 접속하는 클라이언트 프로그램
- <https://github.com/iPuTTY/iPuTTY/releases>
- 압축 파일을 다운로드하고 압축 해제하여 사용



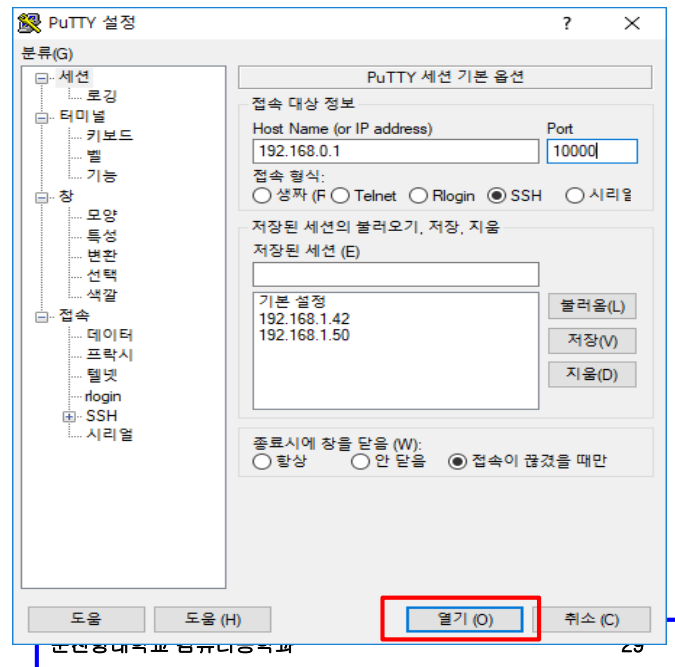
PuTTY



PuTTY - 서버 접속 (1)

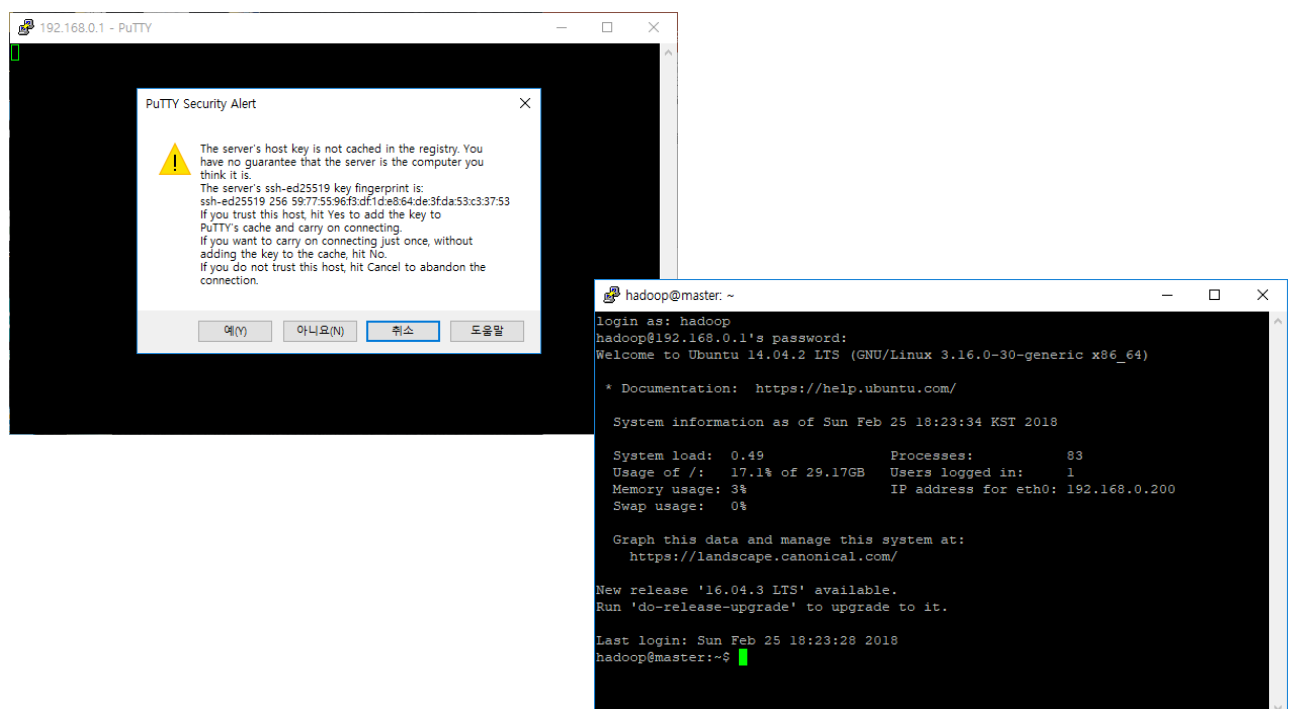
□ PuTTY 실행하여 Master 서버를 ssh로 접속

- 포트포워딩 규칙에 따라 IP, Port 입력
- Host Name
 - 192.168.0.1
- port
 - 10000
- 세션
 - Master
- 저장 후 열기 선택하여 실행



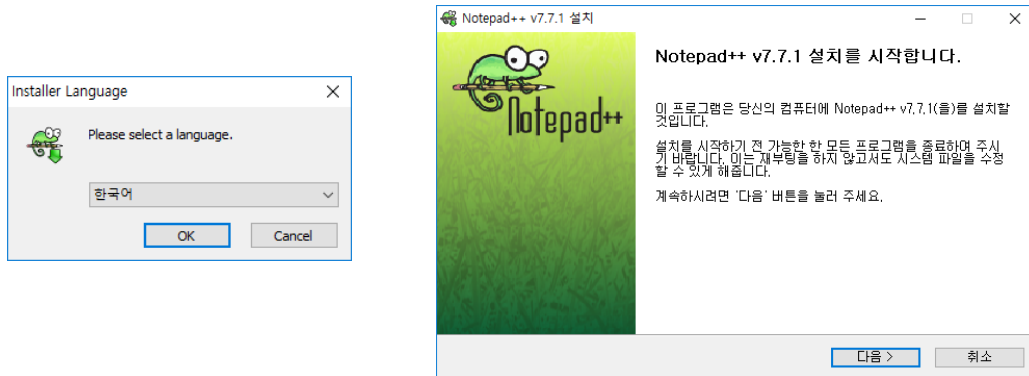
PuTTY - 서버 접속 (2)

□ 최초 접속시 보안경고 창

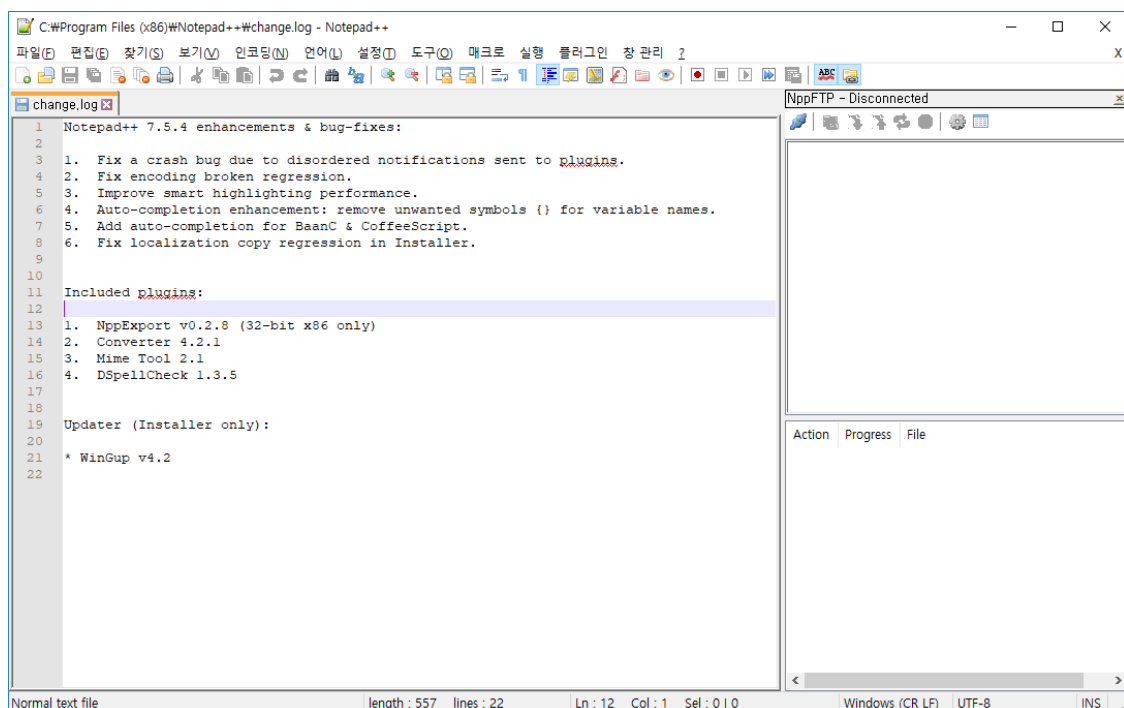


Notepad++ - 설치

- ❑ 오픈 소스 문서 및 소스 코드 편집기(editor)
 - FTP를 사용한 원격 서버의 파일 편집 기능
- ❑ 다운로드 및 설치
 - <https://notepad-plus-plus.org/>



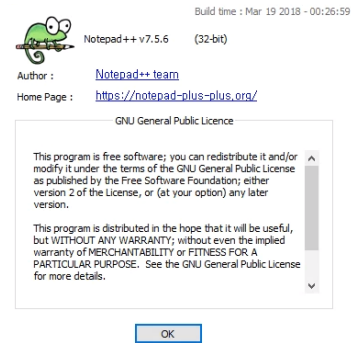
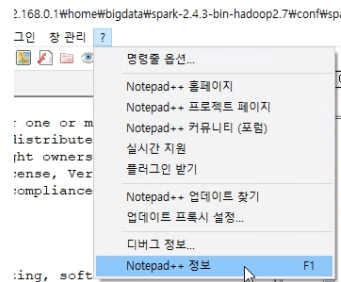
Notepad++ - 실행



Notepad++ – 플러그인 관리자 설치 (1)

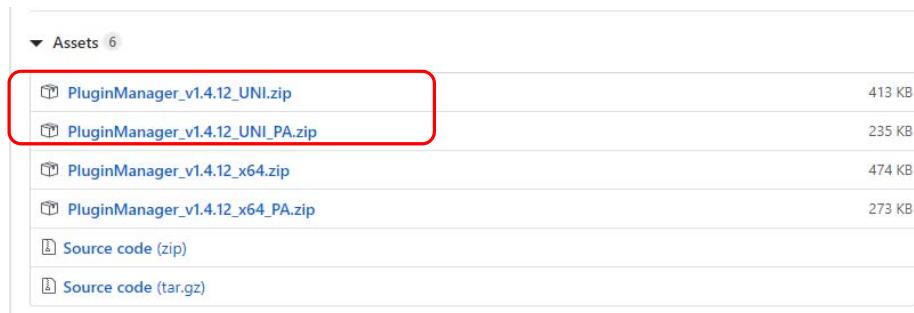
Notepad++ 플러그인 관리자 설치

- 버전 확인



- 다운로드

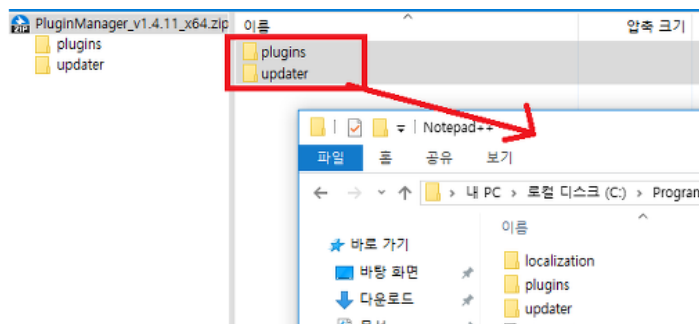
<https://github.com/bruderstein/nppPluginManager/releases>



33

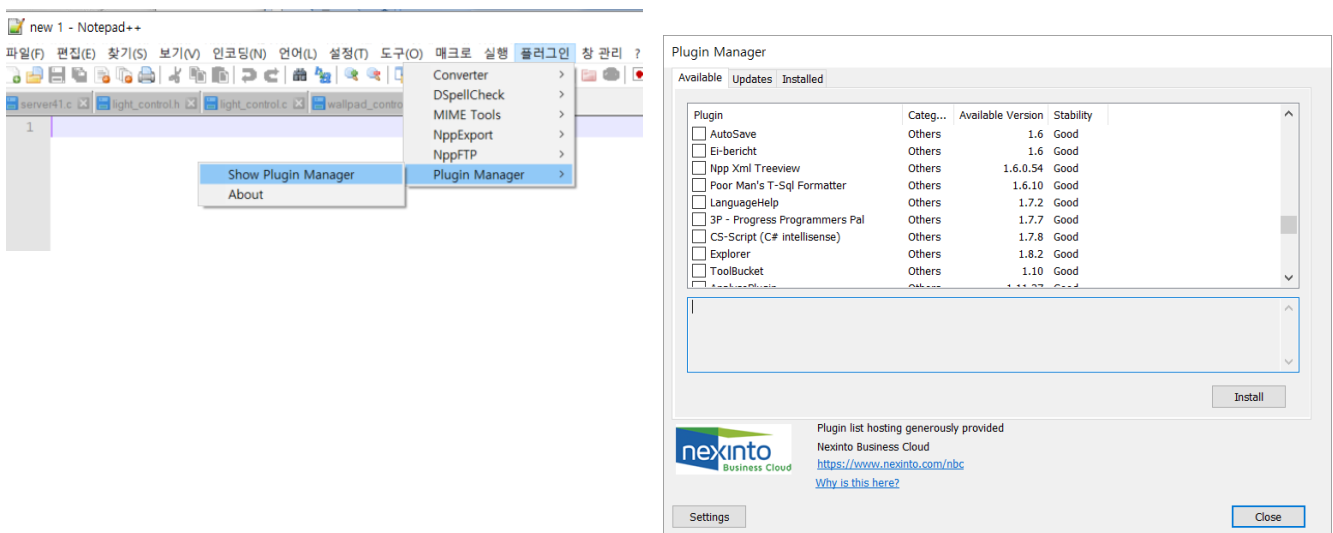
Notepad++ – 플러그인 관리자 설치 (2)

- 압축 풀고 생성된 **plugins**, **updater** 폴더를 Notepad++ 설치 경로인 C:\Program Files (x86)\Notepad++ 에 복사



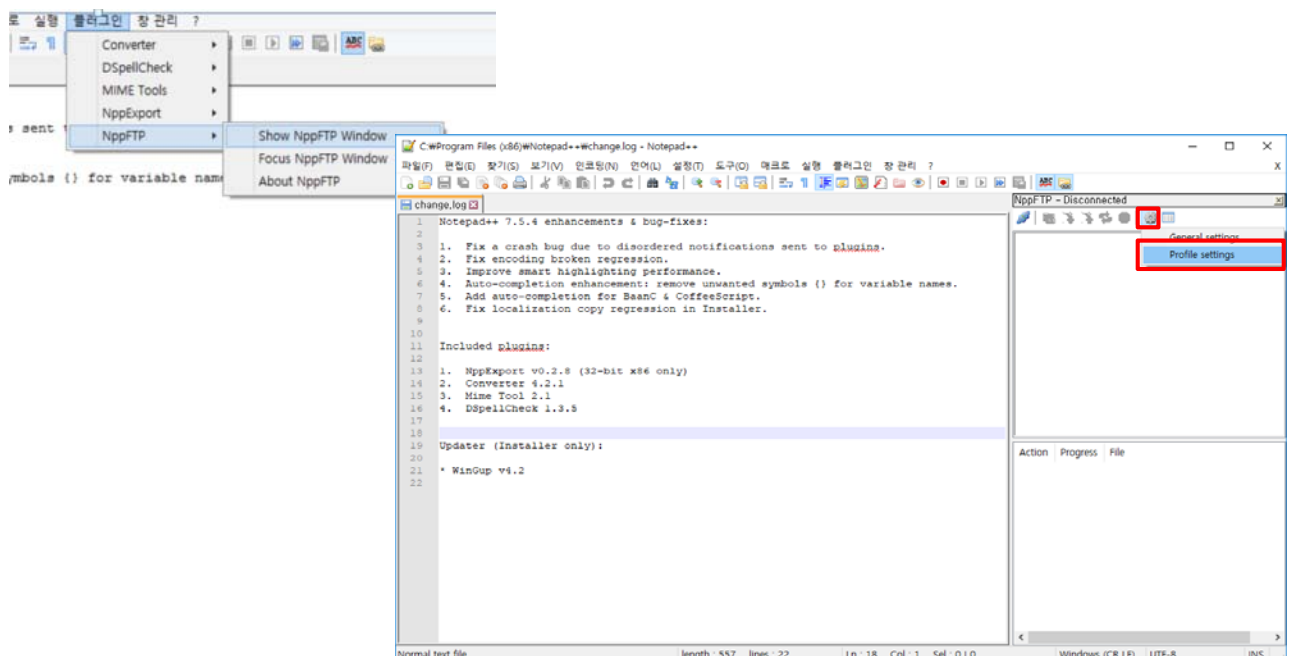
Notepad++ - FTP 플러그인 설치

- 원격 서버 연결 편집을 위해 FTP 플러그인(NppFTP) 설치
 - 플러그인 메뉴 -> Plugin Manager -> Show Plugin Manager
 - NppFTP 선택 후 Install

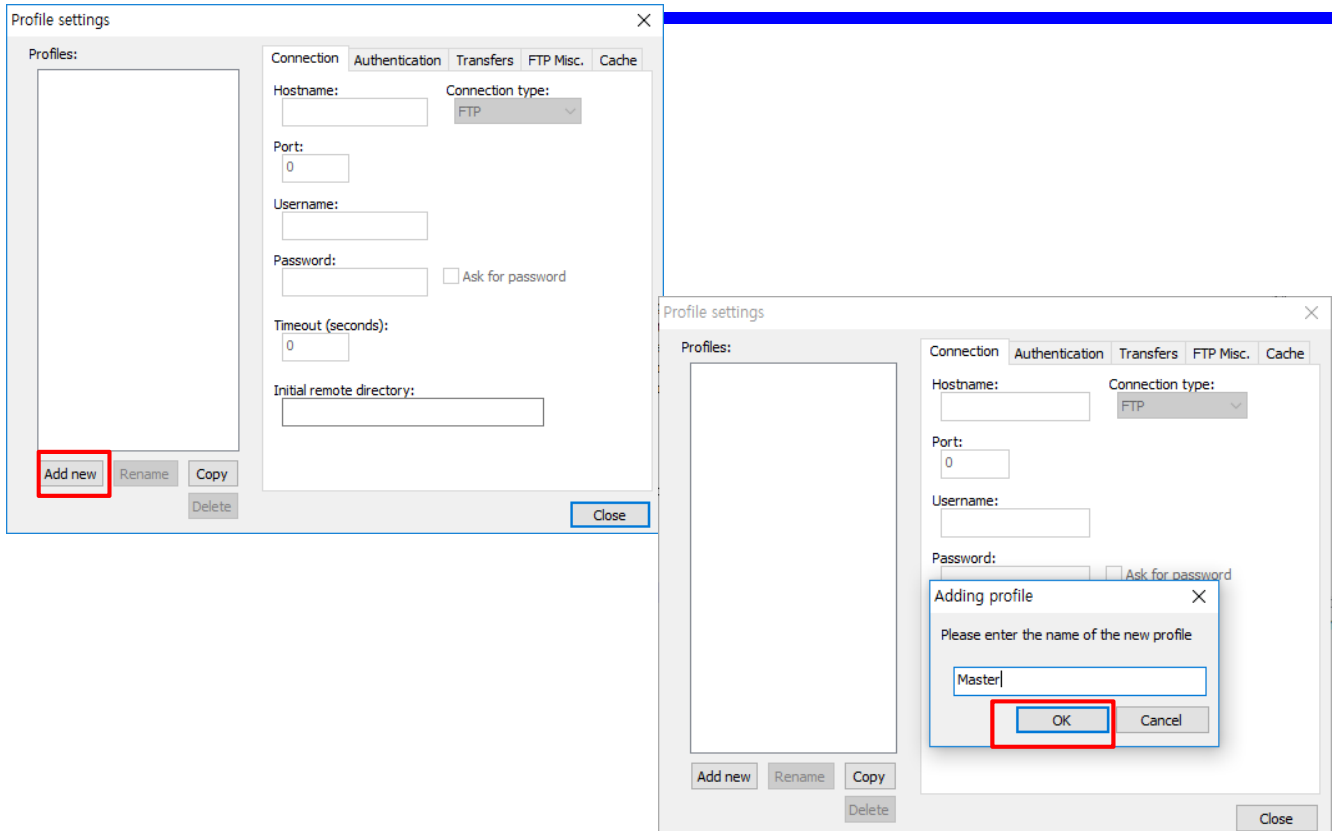


Notepad++ - FTP 플러그인 설정 (1)

- Notepad++ 플러그인 확인
 - 플러그인 -> NppFTP -> Show NppFTP Window



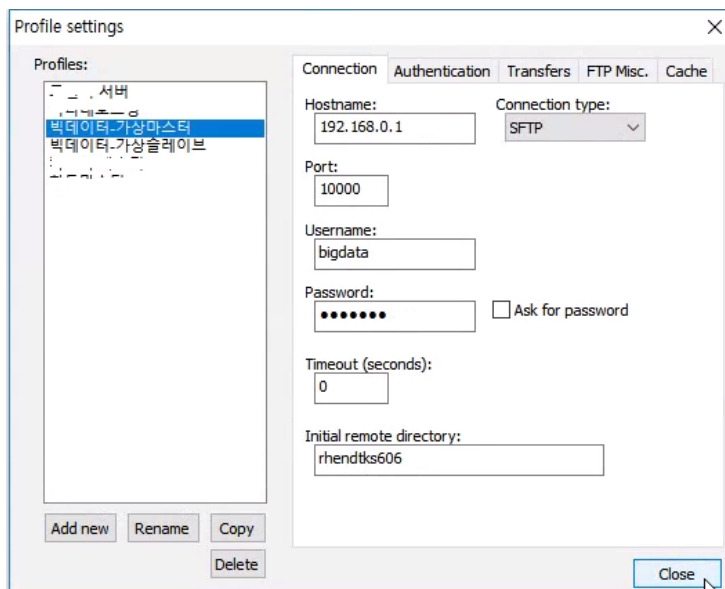
Notepad++ - FTP 플러그인 설정 (3)



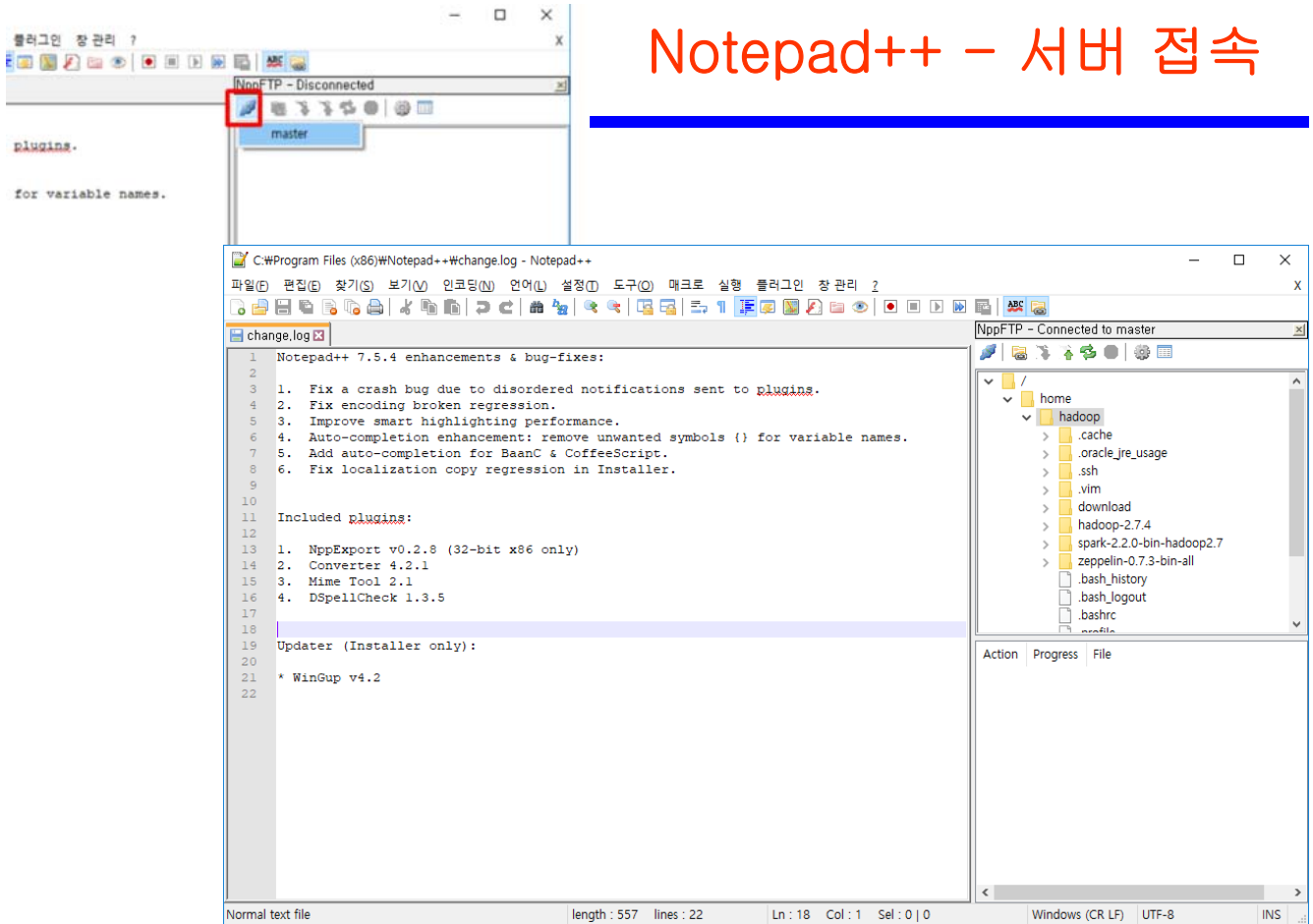
Notepad++ - FTP 플러그인 설정 (4)

□ FTP 설정

- sFTP : SSH 포트 사용

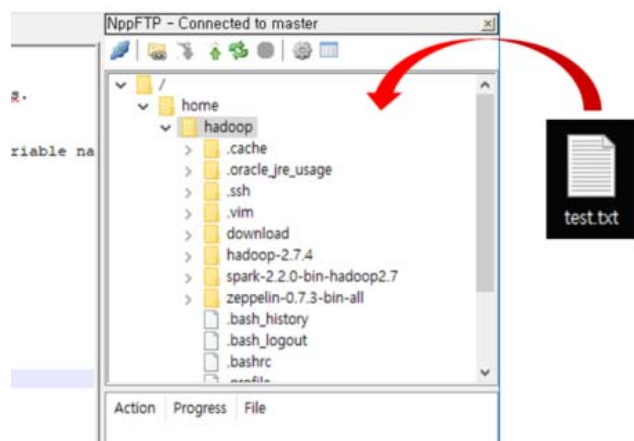
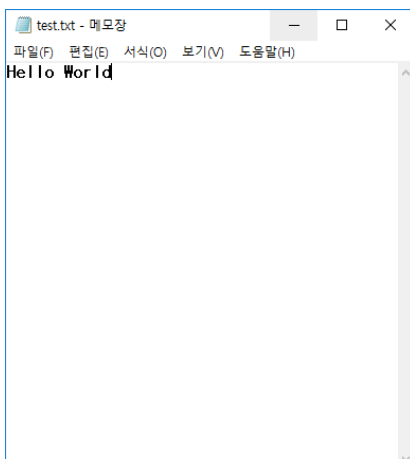


Notepad++ - 서버 접속

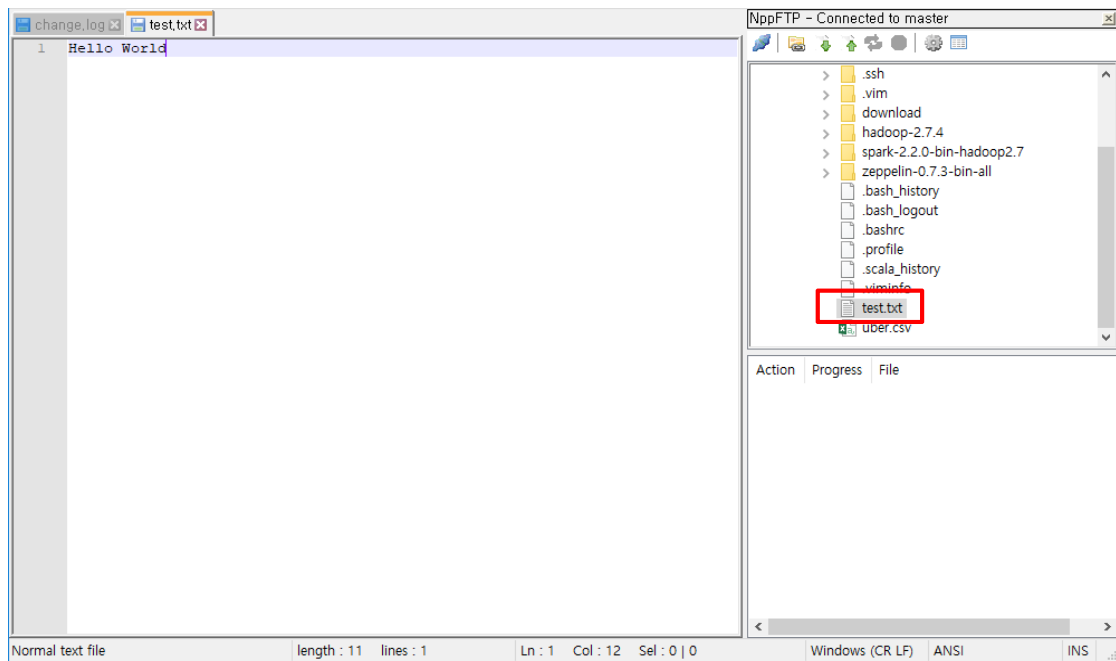


Notepad++ - 파일 생성 (1)

- ❑ 클라이언트에서 파일 생성 후 서버에 전송
- ❑ test.txt 예
 - 메모장에서 파일 생성
 - 생성된 파일 NppFTP 윈도우로 드래그

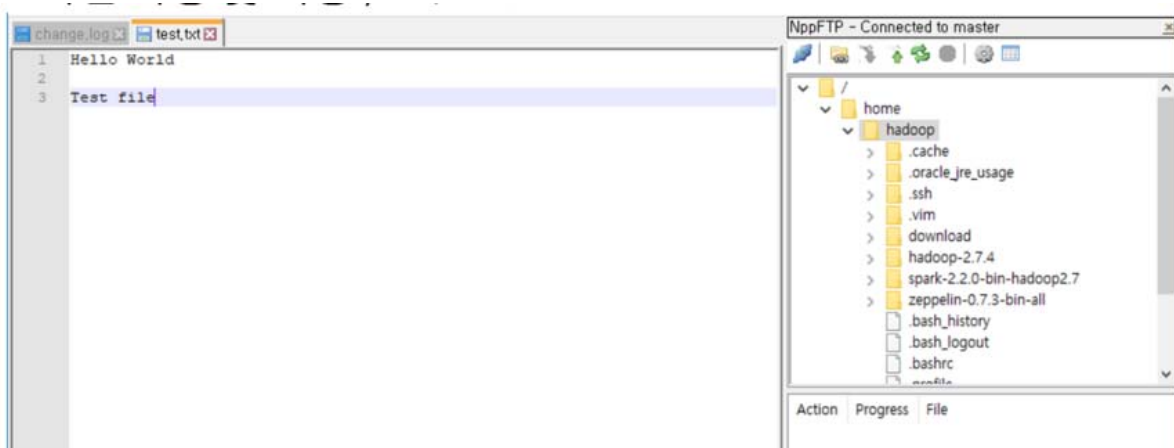


Notepad++ - 파일 생성 (2)



Notepad++ - 원격 파일 수정

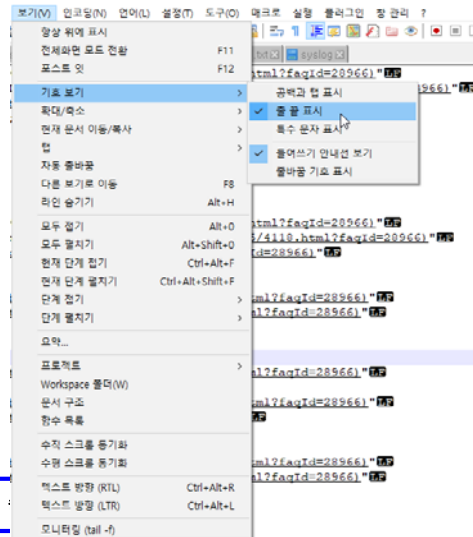
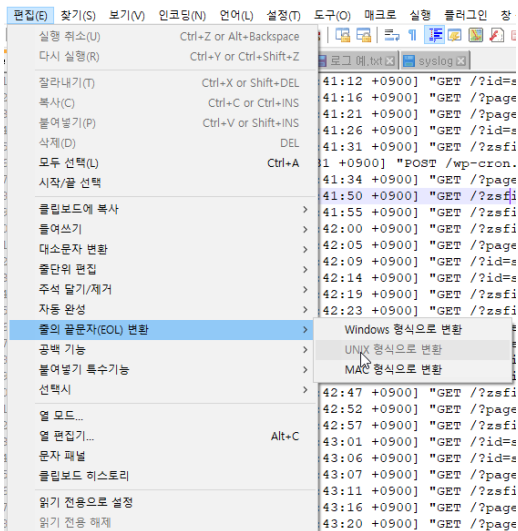
□ 원격 파일 수정 후 저장 예



□ 윈도우와 리눅스는 EOL(End Of Line) 표현이 다름

- 윈도우: CR LF
- 리눅스: LF

□ Notepad++ 저장 시 EOL 변환 필요



43

□ 가상 머신의 Master, Slave1 서버에 임의 프로그램을 작성하고 실행 하여라.

- 언어는 C, C++, Java, Python 등으로 작성된 임의의 프로그램
 - Master, Slave1 서버의 프로그램과 언어는 서로 달라야 함
 - 필요하면 컴파일러, 인터프리터 등 설치
- Notepad++로 소스 코드 작성
- 실행은 PuTTY접속 후 (컴파일 후) 실행하여 결과 확인

