

Lecture 02 - 명령행 인터페이스

개요

- 아마존 웹서비스(AWS) 이해하기 (20m)
- 아마존 웹서비스에 데이터 분석용 원격 컴퓨터(EC2) 만들고 접속하기 (30m)
- 휴식 (10m)
- 명령행 인터페이스 개념 이해하기 (50m)

왜 원격 컴퓨터가 필요한가

- 윈도우는 일반 사용자에게 적합. 데이터 분석, 프로그래밍 등에는 맥 또는 리눅스가 적합
- 모든 수강생이 최대한 동일한 환경에서 실습을 할 수 있음
- 왜 하필 아마존? 실습용 장난감 환경이 아니라 실제로 쓸 수 있는 제대로 된 환경. 실제 업무에서 가장 널리 쓰임

아마존 웹서비스 가입하기 (20m)

AWS 개요

- Amazon Web Services의 약자. 서버, 데이터베이스, 저장소, 네트워크, 방화벽 등 각 요소를 잘게 분해하여 부품 형태로 제공. 사용한만큼만 과금
- 세계 여러 지역의 데이터센터 중 원하는 곳을 선택할 수 있음 (한국 지역도 서비스 예정. 일부 기업에서는 이미 사용 중)

가입 절차

1. <https://aws.amazon.com> 접속 후 “Create an AWS Account” 클릭
2. 이메일 주소 입력 → “I am a new user” 선택 → “Sign in using our secure server” 클릭
3. 이름, 이메일 주소 (재입력), 패스워드 입력 후 “Create account” 클릭
4. 주소 등 각종 정보 입력 후 “Create Account and Continue” 클릭 (**전화번호는 반드시 “010-1234-5678” 형식으로 입력**)

5. Country Code가 올바른지 확인한 후 전화 인증 버튼 클릭

Country Code	Phone Number	Ext
Korea, Republic of (+82) ▾	010-2598-7416	
<div>Call Me Now</div>		

6. 전화가 오면 PIN 번호를 입력하고 “Continue to select your Support Plan” 클릭

2. Call in progress

Please follow the instructions on the telephone and key in the following Personal Identification Number (PIN) on your telephone when prompted.

PIN: 3895

If you have not yet received a call at the number indicated above please wait. This page will automatically update with what you need to do next.

7. “Basic (Free)” 선택 후 “Continue”
8. “Launch Management Console”을 선택하면 관리 화면이 나옴

다음부터는...

1. <https://aws.amazon.com> 접속 후 이메일/패스워드를 입력하고 로그인
2. 끝

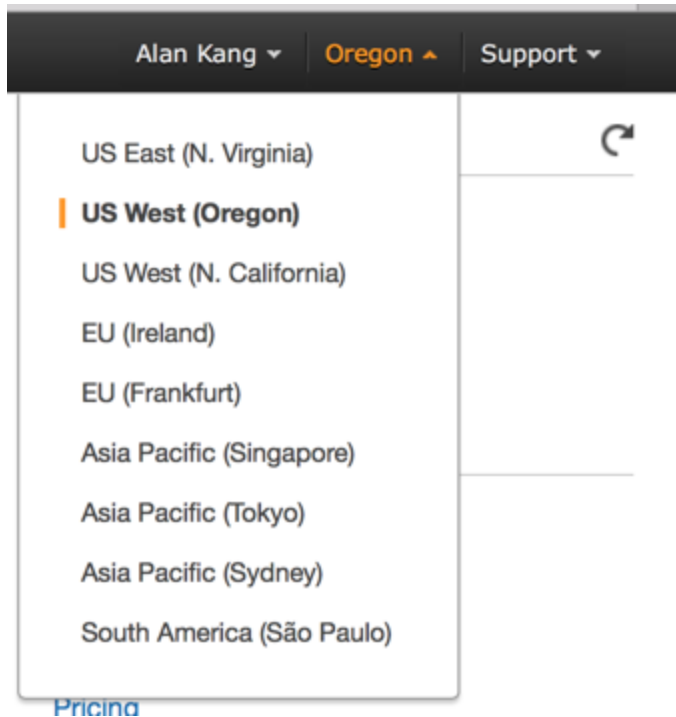
EC2 원격 서버 만들기 (30m)

EC2 개요

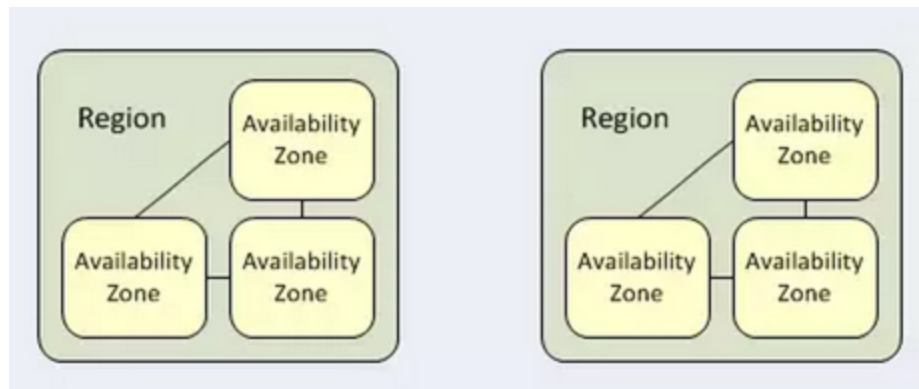
- 우리 수업에서는 AWS 서비스 중 EC2 (Elastic Computing Cloud)를 사용할 예정
- EC2란?
 - 원격에서 접속할 수 있는 컴퓨터
 - CPU, 메모리, 운영체제, 하드디스크, 네트워크 카드 등 필요한 부품을 조립해서 가상의 컴퓨터를 만드는 것
 - 실제 컴퓨터 한 대를 쓰는 것은 아니고 가상화(Virtualization)를 이용

EC2 인스턴스 만들기

1. 대시보드에서 “EC2” 클릭
2. 화면 우측 상단에서 “Asia Pacific (Singapore)” 또는 “Asia Pacific (Tokyo)” 선택



- a. Region: 지리적으로 떨어진 지역들
- b. "Availability Zone: 각 지역 내에서 빠른 회선으로 연결되어 있으나 서로 최대한 격리되어 있는 소그룹.



- c. 고가용(Highly Available) 서비스를 제공하려면 필수. 하지만 우리는 “서비스”를 제공하는 것이 아니고 데이터 분석 용도로만 쓰려는 것이니 크게 중요하지 않음
 - d. 다만 Region이 가까우면 지연(latency)이 적을테니 상대적으로 쾌적할 수 있음
3. “Launch Instance” 클릭
 4. AMI 선택
 - a. “Ubuntu Server” 선택

- b. Ubuntu Server란? 윈도우, macOS 등과 같은 운영체제(OS; Operating System)의 일종. Linux 운영체제의 한 종류이며, Linux 중 (그나마) 사용하기가 편리하고 널리 쓰임
- 5. 인스턴스 선택
 - a. “t2.micro” 선택 후 “Next: Configure Instance Details” 클릭. 파란 버튼 아니고 **회색 버튼**

Review and Launch

Next: Configure Instance Details

(그냥 지나가도 되지만 기왕이면 각 단계별로 이해하고 넘어가기 위함)

- b. “t2.micro”란 EC2 인스턴스의 한 종류. 인스턴스에 따라 CPU, 메모리, 네트워크 대역폭 등이 결정됨
- c. t2.micro는 가장 성능이 낮은 인스턴스. 일정 기간 무료로 사용 가능
- 6. 인스턴스 설정
 - a. 기본 설정 유지하고 “Next: Add Storage” 클릭
- 7. 하드디스크(정확히는 SSD) 추가하기
 - a. 기본 설정 유지하고 “Next: Tag Instance” 클릭
- 8. 태깅하기
 - a. 인스턴스가 많아지면 관리가 어려워짐
 - b. “Name”에 “fastcamp” 입력
 - c. “Next: Configure Security Group” 클릭
- 9. 보안 설정
 - a. “Create new security group” 선택
 - b. “Security group name”에는 “SSH only” 입력
 - c. “Description”에는 “TCP 22 from anywhere”라고 입력
 - d. Type: SSH, Protocol: TCP, Port Range: 22, Source: Anywhere 0.0.0.0/0 입력
 - e. 포트? 외부의 컴퓨터와 연결할 때 쓰이는 플러그 같은 것. 불필요하게 포트를 열어놓지 않아야 함
 - f. 보안의 기본: “일단 열어놓고 위험하면 닫는다”가 아니라 “일단 다 막아놓고 필요하면 열어준다”. **최소 권한의 원칙(principle of least privilege)**이라고 부름
 - g. 데이터 분석가의 가장 중요한 직업 윤리 중 하나. 데이터가 외부로 유출되지 않도록 노력하기
 - h. SSH? Secure Shell의 약자. 안전한 원격 접속 프로토콜. 22번 포트를 사용함
 - i. “Next: Review and Launch” 클릭
- 10. 컴퓨터 켜기
 - a. 보안 경고:



Improve your instances' security. Your security group, SSH Only, is open to the world.

Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only.

You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)

세상 어디서든 들어올 수 있는 포트가 있어서 위험하다는 경고. 좀 더 안전하게

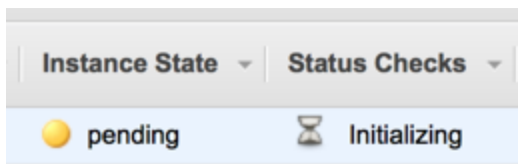
하려면? 특정 IP 주소에서만 접속할 수 있게 제약. 우리 상황에서는 지나친 우려.
보안성과 사용성 사이의 균형이 중요.

b. "Launch"

11. 키 생성하기

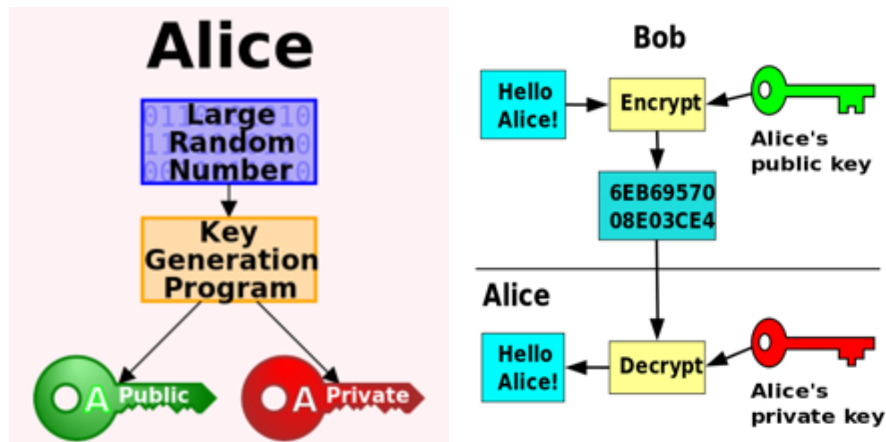
- "Create a new key pair" 선택
- "Key pair name"에 "fastcamp" 입력
- "Download Key Pair" 클릭. 맥 사용자는 반드시 "Downloads" 폴더에 저장할 것. 파일을 분실하면 컴퓨터에 접속할 수 없음. 반대로, 이 파일을 가진 사람은 누구나 컴퓨터에 접속할 수 있음
- "Launch Instance" 클릭
- "View Instance" 클릭

12. 컴퓨터가 켜질동안 기다리기

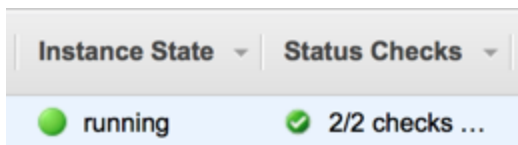


13. (기다리는 동안...) Key Pair란?

- 내 비밀키로 암호화를 하면 내 공개키로만 해독이 가능
- 내 공개키로 암호화를 하면 내 비밀키로만 해독이 가능
- 비밀키는 나만 보관하고 공개키는 나와 이야기하길 원하는 모든 사람 및 컴퓨터에게 배포
- 뭐가 좋은가? 도청 방지, 본인 인증, 위조 방지



14. 축하합니다 :-)



원격 컴퓨터의 Public IP 주소를 적어두세요.



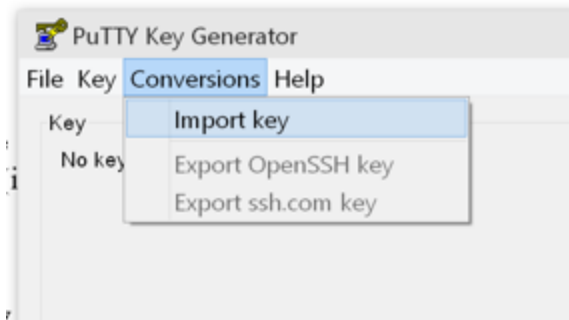
15. Stop vs. Terminate

- Stop은 전원 끄기. 컴퓨터를 켜다가 켜면 IP 주소가 바뀝니다. 두달 동안 끄지 맙시다 :-)
- Terminate는 컴퓨터 부숴버리기

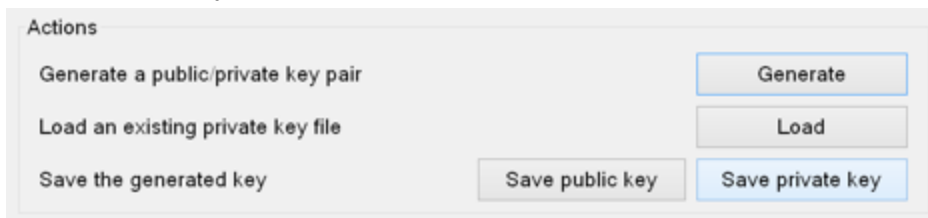
원격 컴퓨터에 접속하기 (20m)

키파일 준비하기 (윈도)

- <http://www.putty.org/> 접속 후 Download 링크 클릭
- "A Windows installer for everything except PuTTYtel" 섹션에서 [putty-0.64-installer.exe](#) 다운로드
- 설치
- 시작메뉴에서 PuttyGen 실행
- "Import key" 메뉴를 실행하여 fastcamp.pem 파일을 읽어오기



- "Save private key" 버튼을 클릭하여 fastcamp.ppk 파일로 저장하기 (경고는 무시)



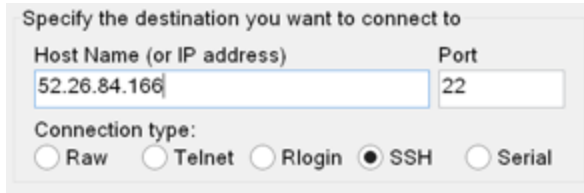
키파일 준비하기 (맥)

Terminal 앱 실행 후 다음 명령을 입력

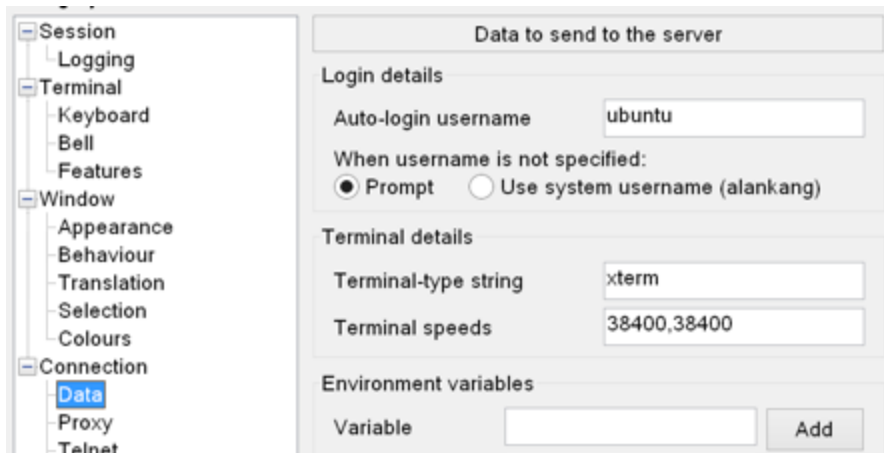
```
mkdir ~/.ssh
mv ~/Downloads/fastcamp.pem ~/.ssh
chmod 600 ~/.ssh/fastcamp.pem
```

접속 준비하기 (윈도)

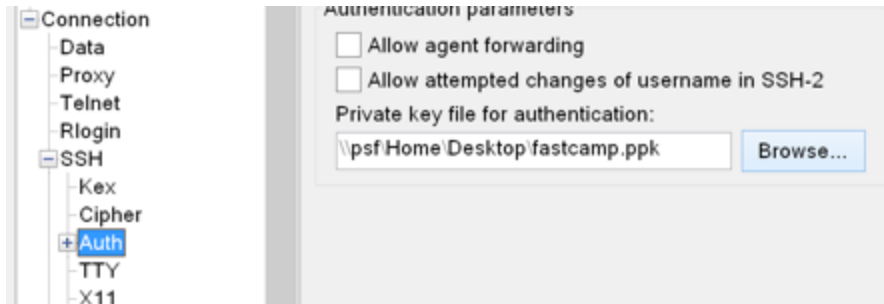
1. 시작메뉴에서 “PuTTY” 실행
2. “Host Name” 부분에 원격 컴퓨터의 IP 주소를 입력



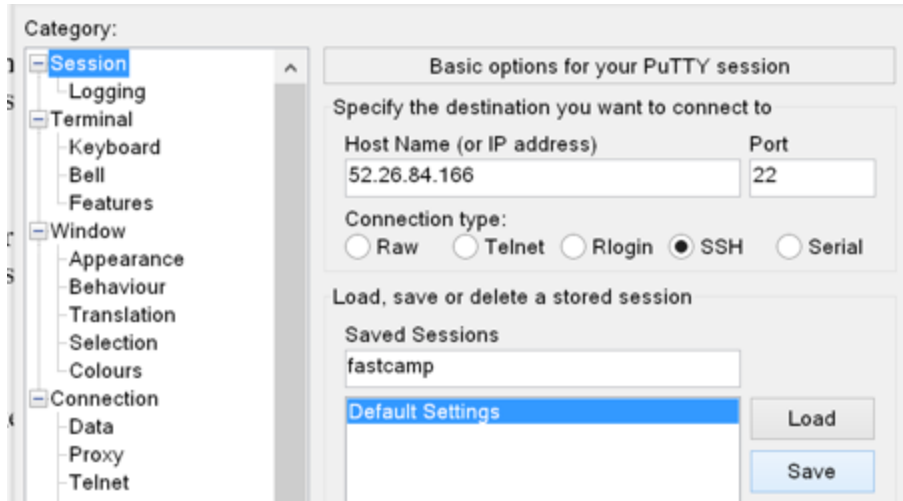
3. “Connection -> Data” 메뉴로 이동하여 “User name”에 “ubuntu” 입력



4. “Connection -> SSH -> Auth” 메뉴로 이동하여 “Private key file for authentication”에 fastcamp.ppk 파일을 지정



5. “Session” 메뉴로 이동하여 “Saved sessions”에 “fastcamp” 입력 후 “Save” 버튼 클릭



6. “Saved sessions”에서 “fastcamp” 더블클릭

접속 준비하기 (맥)

Terminal 앱에서 다음 명령을 입력:

```
ssh -i ~/.ssh/fastcamp.pem ubuntu@원격컴퓨터주소
```

다음부터는... (윈도)

1. 시작메뉴에서 PuTTY 실행
2. “fastcamp” 더블클릭

다음부터는... (맥)

Terminal 앱에서 다음 명령을 입력:

```
ssh -i ~/.ssh/fastcamp.pem ubuntu@원격컴퓨터주소
```

휴식 (10m)

명령행 인터페이스 기초 (50m)

개요

명령행 인터페이스란:

- 명령행 인터페이스(CLI; Command-line Interface) 또는 명령행 사용자 인터페이스(CUI; Command-line User Interface)
- 그래픽 사용자 인터페이스(GUI; Graphic User Interface)와 대비되는 개념
- 키보드로 명령어를 입력 → 명령이 실행됨 → 결과가 화면에 출력됨

예시:

```
// 현재 작업디렉토리 출력하기 (print working directory)
pwd
```

```
// 새 디렉토리 만들기 (make directory)
mkdir fastcamp
```

```
// 해당 디렉토리로 이동하기 (change directory)
cd fastcamp
```

```
// 현재 작업디렉토리 출력하기 (print working directory)
pwd
```

```
// 인터넷에 있는 파일 다운로드
curl http://s.g15e.com/pride.txt > pride.txt
```

```
// 현재 디렉토리에 있는 파일 목록 보기
ls
```

```
// 파일 내용 출력하기
cat pride.txt
```

```
// 앞 부분만 보기
head pride.txt
```

```
// 뒤 부분만 보기
tail pride.txt
```

```
// 몇 줄인지 알아보기
wc -l pride.txt
```

```
// 남주 Darcy가 나오는 문장들만 보기
grep -P Darcy pride.txt
```

```
// 남주 Darcy가 몇 번 나오는지 알아보기
grep -P Darcy pride.txt | wc -l
```

CLI의 단점:

- 명령어를 외워서 입력하기 때문에 처음에 배우기 어렵다

CLI의 장점:

- 일단 익숙해지면 GUI에 비해 훨씬 쉽고 빠르다
- 명령들을 조합해서 원하는 결과를 만들어낼 수 있음. 명령 중 하나는 “RScript” (R과 조합해서 사용할 수 있음)
- 자동화 용이

두 종류의 쉬움:

- 배우기 쉬움(ease of learn) - 비전문가용, 아주 가끔만 쓰이는 시스템(현금인출기)에 적합
- 쓰기 쉬움(ease of use) - 전문가용, 빈번하게 쓰이는 시스템(은행원용 단말기)에 적합

파일, 표준입출력, 출력전환, 파이프

유닉스 계열 운영체제는 거의 모든 것을 파일로 취급:

- 키보드: 읽기 전용 파일
- 모니터: 쓰기 전용 파일
- 네트워크: 읽고 쓰기가 가능한 파일

표준입출력:

- 표준입력: Standard Input. 보통은 키보드
- 표준출력: Standard Output. 보통은 모니터
- cat 명령: 표준입력으로 들어온 내용을 그대로 표준출력으로 내보낸다.
- sort 명령: 표준입력으로 들어온 내용을 정렬하여 표준출력으로 내보낸다.

출력전환 - 표준출력 혹은 표준입력을 다른 파일로 대체

```
# 키보드로 입력받은 내용을 정렬하기
sort

# 키보드로 입력받은 내용을 정렬하여 파일로 쓰기
sort > test.txt
```

파이프 - 한 프로그램의 표준출력을 다른 프로그램의 표준입력으로 연결

```
# 파일 내용을 출력 → Darcy 가 있는 줄만 거르기 → 몇 줄인지 계산하기
cat pride.txt | grep Darcy | wc -l
```

Pipe and Filter architecture:

- 작은 부품들을 조합하여 원하는 작업을 수행
- 유닉스 계열 운영체제의 핵심 철학 중 하나

예시. 오만과 편견에서 가장 빈번하게 나타나는 단어 50개는?

```
cat pride.txt | tr '[:space:]' '\n' | sort | uniq -c | sort -rn | head -n 50
```

참고자료

- 리눅스 텍스트 처리 명령어
<https://wiki.kldp.org/HOWTO/html/Adv-Bash-Scr-HOWTO/textproc.html>