

AWS

클라우드 컴퓨팅

- 클라우드 컴퓨팅은 컴퓨터의 각종 연산 및 처리를 개인 PC가 아닌 인터넷 공간, 즉 클라우드에서 처리하는 방식을 의미
- MIT 학생이었던 드루 하우스턴은 중요한 파일을 담은 USB 메모리를 책상에 두고 오는 불편함을 해결할 아이디어를 가지고 드롭박스 창업. 바로 자신의 파일이나 사진 등을 클라우드 서버에 저장해서 언제 어디서나 다양한 기기를 통해 사용
- 클라우드 컴퓨팅 정의
 - 인터넷 기반 컴퓨팅의 일종으로 정보를 자신의 컴퓨터가 아닌 인터넷에 연결된 다른 컴퓨터로 처리하는 기술을 의미
 - 공유 컴퓨터 처리 자원과 데이터를 컴퓨터와 다른 장치들에 요청 시 제공해 줌
 - 구성 가능한 컴퓨팅 자원에 대해 어디서나 접근이 가능한, 주문형 접근을 가능케하는 모델이며 최소한의 관리 노력으로 빠르게 예비 및 릴리스를 가능케 함
- 클라우드 컴퓨팅의 장점
 - 유연한 IT 인프라 관리
 - 신속한 인프라 도입
 - 간편한 글로벌 서비스
 - 예상치 못한 트래픽 폭주 대응

용어 정리

- 온프레미스(On-premise)

- 기업 시스템에서 지금까지 채택해온 전통적인 IT 서비스 방식
- 기업이 자체적으로 데이터 센터를 보유하고 시스템 구축부터 운영까지, 모두 수행하는 형태를 뜻 함
- 초기 시스템 투자에 드는 비용 부담이 크고, 시스템 가동 후 운영에 드는 비용도 시스템 사용량과 상관없이 일정 금액을 부담해야 함

- IaaS

- Infrastructure as a Service의 약자로 인프라는 클라우드 IT의 기본 요소들을 포함하고 일반적으로 네트워킹 기능, 컴퓨터 및 데이터 스토리지 공간을 제공
- 넷플릭스는 자체 데이터 센터를 구축하여 서비스를 운영하는 대신 AWS의 IaaS 서비스를 이용하는 방식을 채택
- 전 세계에 보다 빠른 서비스를 제공하기 위해 AWS에서 필요한 컴퓨팅 자원을 빌려서 서비스를 운영하며, 때에 따라 필요한 컴퓨팅 인프라를 몇 분 또는 몇 시간 안에 IaaS로 꾸려 운영할 수 있음

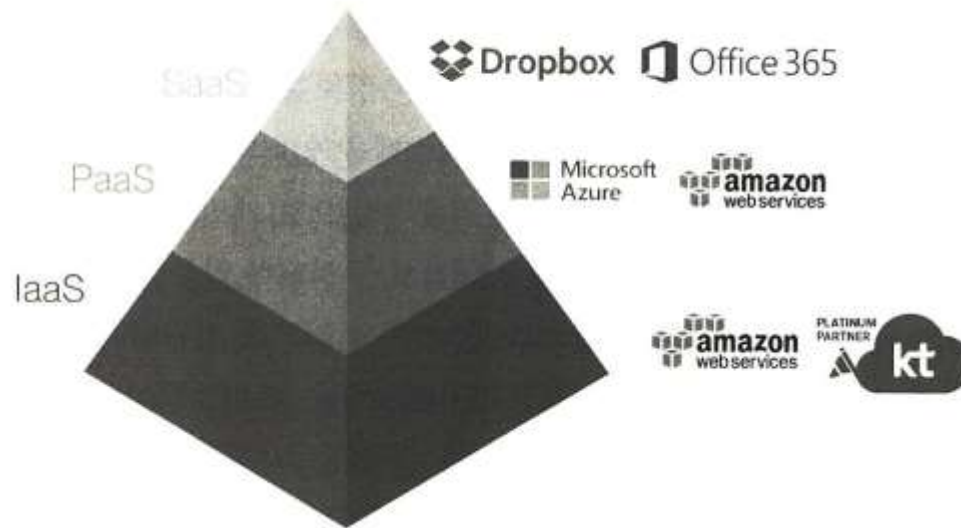
- PaaS

- Platform as a Service의 약자로 플랫폼은 소프트웨어 서비스를 개발할 때 필요한 플랫폼을 제공하는 서비스
- 사용자는 PaaS에서 필요한 서비스를 선택해 애플리케이션을 개발하고, PaaS 운영 업체는 개발자가 소프트웨어를 개발할 때 필요한 API를 제공해 개발자가 좀 더 편하게 앱을 개발할 수 있게 지원

용어 정리

- SaaS

- Software as a Service의 약자로 소프트웨어가 클라우드 환경에서 운영되는 애플리케이션 서비스를 의미
- 모든 서비스가 클라우드에서 이루어지는데, 소프트웨어를 구입해서 PC에 설치하지 않아도 웹에서 소프트웨어를 빌려 쓸 수 있는 환경
- SaaS를 사용하면 서비스가 어떻게 유지 관리되는지 또는 기본 인프라가 어떻게 관리되는지 고민할 필요가 없고, 해당 소프트웨어를 어떻게 사용할지에 대해서만 생각하면 됨



용어정리

• 라우터

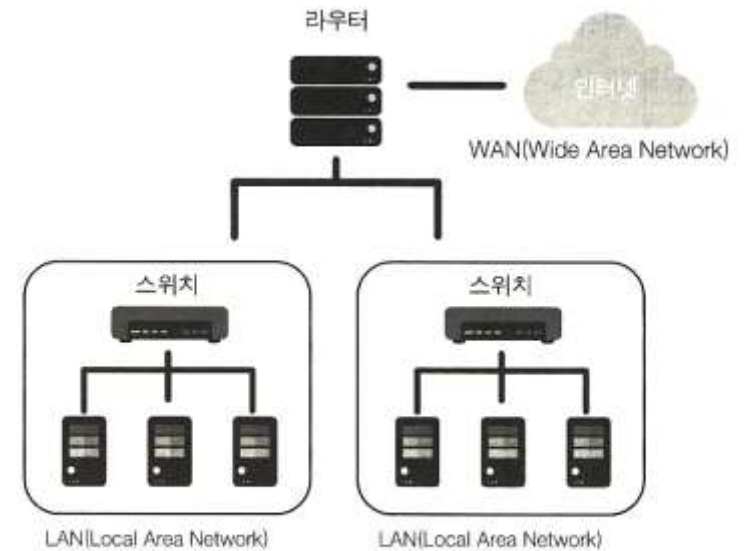
- 라우터는 서로 다른 지역의 컴퓨터를 연결하는 기기
- 2개 이상의 네트워크와 네트워크간 데이터 전송을 위한 최적 경로를 설정해주며, 데이터를 해당 경로를 따라 한 통신망에서 다른 통신망으로 통신 가능하게 도와주는 인터넷 접속 장비
- 네트워크를 통해 정보를 주고받을 때 데이터에 담긴 수신처의 주소를 읽고 가장 적절한 통신 경로를 이용해 다른 통신망으로 전송하는 장치로, 전화국의 교환기와 비슷한 개념

• 스위치

- 여러 컴퓨터를 연결하는 네트워크 분배기 역할을 함
- 네트워크 장비 중 허브와 스위치가 하나의 네트워크 라인에 여러 대의 컴퓨터에 랜 케이블을 꽂을 수 있도록 분배하는 장비

• 방화벽

- 네트워크의 보안 사고나 문제가 확대되는 것을 막고 격리하려는 것으로 이해면 됨
- 방화벽은 하나의 컴퓨터 시스템과 전체 인터넷을 구분시켜 주는 프로그램 또는 장비
- 시스템 사용자의 외부 접속을 제한하거나 보안 상의 문제로 인한 외부인의 사용을 제한



[그림 1-6-1] 일반적인 네트워크 구성

용어정리

• IP주소

- 인터넷을 사용하려면 NIC에 할당하는 IP 주소와 인터넷 연결 서비스를 관리하는 회사에 가입해야 함
- 현재 IP 주소는 IPv4를 사용하고 있으며, 세계적으로 IP는 중복되서는 않아야 하며, 이를 공인 IP라고 함
- IPv4는 32비트 형식으로 xxx.xxx.xxx.xxx 형태로 사용

• 도메인

- 도메인 네임이란 사용자의 입장에서 쉽게 이용할 수 있도록 숫자로 표현된 주소 대신 문자열로 나열된 주소
- 숫자형태의 IP를 기억하기 힘들기 때문에 도메인 네임을 지정하여 사용

• IPv6

- 128비트 체계의 IP 주소
- IPv4가 약 42억개의 IP 주소를 사용할 수 있기 때문에 부족현상을 해결하기 위한 대안

• 포트

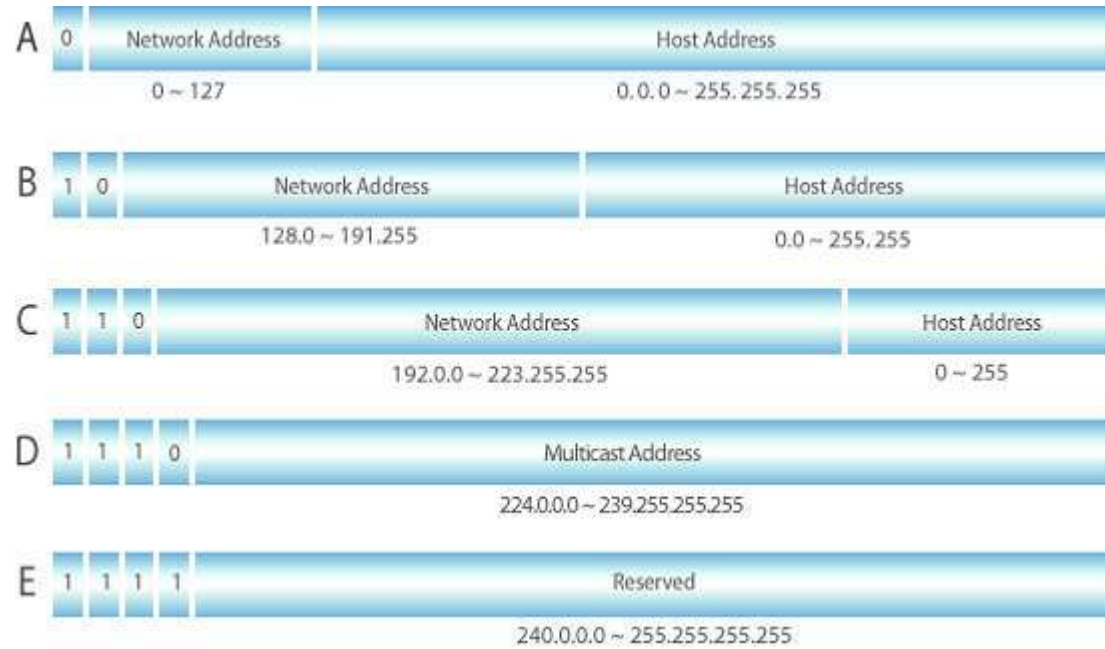
- 컴퓨터에서 통신에 사용하는 프로그램을 식별하는 번호
- well-known port : 0 ~ 1023번
- registered port : 1024 ~ 49151번

서비스명	서비스 내용	포트 번호	전송 계층
TCPMUX	TCP port service multiplexr	1	UDP/TCP
ECHO	echo	7	UDP/TCP
DAYTIME	daytime	13	UDP/TCP
FTP - DATA	FTP 데이터 전송	20	UDP/TCP/SCTP
FTP - CONTROL	FTP 데이터 전송 제어	21	UDP/TCP
TELNET	telnet 터미널 에뮬레이션	23	TCP
SMTP	메일 메시지 전송 프로토콜	25	UDP/TCP
DNS	DNS 질의 응답	53	UDP/TCP
HTTP	웹 페이지	80	UDP/TCP
NTP	Network Time Protocol	123	UDP/TCP
BGP	BGP 라우팅 프로토콜	179	TCP

[표 1-6-1] 주요 포트

용어정리

- IPv4의 주소 클래스



용어정리

- 프로토콜

- 네트워크 사용에도 규약이나 규칙이 필요
- 네트워크 통신을 할 때도 두 기기가 서로 통신하기 위한 사전 약속이 필요한데 이를 프로토콜이라고 함

- TCP/IP(Transmission Control Protocol/Internet Protocol)

- 컴퓨터 통신에 이용되는 프로토콜 중 가장 중요한 프로토콜
- 다른 프로토콜보다 신뢰성이 높은 방식으로, 정보를 전달할 때 중간 유실이나 흐름 장애 시 재전송을 하는 방식
- 이는 한 정보를 여러 패킷으로 나누어 보낼 때 중간에 유실되었을 때 수신자는 불완전한 정보를 받게 되는데 수신자는 다시 발신자에게 중간에 정보가 유실되었음을 알려주고, 다시 정보를 받을 수 있기 때문에 신뢰성이 높은 방식

- 가상 서버

- 클라우드 서비스에서 제공하는 가장 기본적인 기능이 가상 서버임
- 가상 서버는 하나의 물리적인 서버와 마찬가지로 개별 운영체제나 애플리케이션을 실행시킬 수 있음
- 가상 서버는 하나의 물리적 서버를 논리적으로 나누어 CPU, 메모리 등의 자원을 할당



전통적인 서버 구조

가상화된 서버 구조

[그림 1-7-1] 일반 서버와 가상 서버

용어정리

가상화 기술

- 가상화는 물리적인 하드웨어 장치를 논리적인 객체로 추상화하는 것을 의미
- 마치 하나의 장치를 여러 개로 나눠 동작시키거나, 반대로 여러 개의 장치를 묶어 마치 하나의 장치인 것처럼 사용자에게 공유 자원으로 제공할 수 있기 때문에 클라우드 컴퓨팅 구현을 위한 핵심 기술
- 1960년대부터 시작했지만, 1996년 컴팩에서 처음 사용하며 등장한 이래로, 이후 2006년 구글 내부에서 유틸 서버를 활용하는 방식을 사용하면서 발전
- 같은 해 8월 아마존이 EC2라고 불리는 서비스를 개시하면서, 대중에게 널리 알려지기 시작
- 방식에 따라 하이퍼바이저형과 호스트형으로 나눌 수 있음



[그림 1-7-2] 호스트 · 하이퍼바이저 · 컨테이너 가상화

용어정리

• 호스트 가상화

- 하드웨어 위에 기본이 되는 호스트 운영체제를 설치하고, 호스트 운영체제에 가상화 SW를 설치
- 호스트 가상화는 가상화 SW 위에 게스트 운영체제를 작동시키는 가상화를 말함
- 구축하기 쉬운 장점이 있으나, 호스트 운영체제 위에서 실행하기 때문에 CPU나 메모리 사용이 증가하는 오버헤드가 발생
- VirtualBox, Vmware, Parallels

• 하이퍼바이저 가상화

- 하드웨어에 가상화를 전담하는 SW인 하이퍼바이저를 배치해 HW와 가상 환경을 제어
- 호스트가 없는 HW를 직접 제어하기 때문에 컴퓨터 자원을 효율적으로 사용할 수 있음
- Hyper-V, XenServer

• 컨테이너 가상화

- 호스트 가상화, 하이퍼바이저 가상화처럼 운영체제를 여러 개 가동시키면 이것만으로도 많은 자원이 필요
- 따라서 운영체제에 논리적인 영역(컨테이너)을 만들고, 애플리케이션을 작동하는 데 필요한 라이브러리와 애플리케이션 등을 컨테이너 안에 넣어, 마치 개별 서버처럼 사용할 수 있게 한 것이 이 컨테이너 가상화
- OpenVZ, LXC, FreeBSD Jail, Docker등이 있음

용어정리

- 분산 처리 기술

- 분산 처리 기술과 클라우드 서비스를 이용해서 데이터를 여러 개 서버에 나누어 병렬 처리 할 수 있게 함
- 대량의 데이터를 분산 처리하는 방법으로 여러 개의 서버를 결합하여 하나의 컴퓨터처럼 보이게 만드는 기술을 클러스터링이라고 함
- 대표적인 기술로 Hadoop과 Spark가 있음

- 서버리스

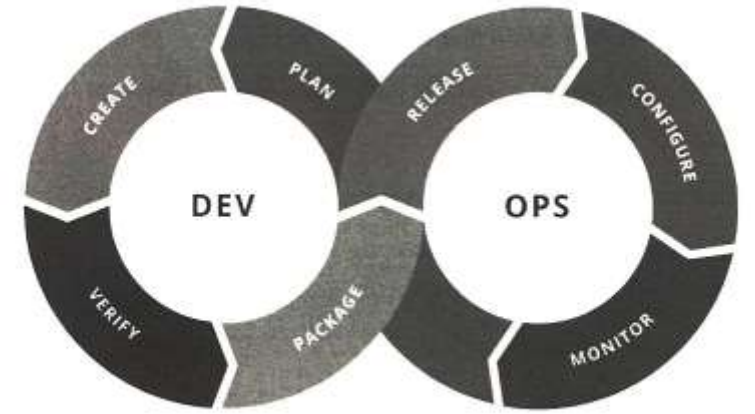
- Serverless 컴퓨팅은 클라우드 컴퓨팅 실행 모델의 하나로, AWS와 같은 클라우드 제공자가 동적으로 가상 머신 자원의 할당을 관리해 미리 설정된 서버리스 코드를 실행시키는 기술을 의미
- 2014년 AWS의 람다 서비스가 시초

- 로드 밸런싱

- Load Balancing은 컴퓨터 자원 분산 기술의 일종으로 둘 혹은 셋 이상의 중앙처리장치 혹은 저장 장치와 같은 컴퓨터 자원들의 작업을 나누는 것을 의미
- 웹 서비스를 운영할 때 로드 밸런싱을 고민해야 한 서버에 트래픽이 과도하게 집중되는 것을 방지할 수 있음

용어정리

- 스케일 업(Scale-Up)
 - 서버 자체의 성능을 업그레이드하여 처리 능력을 향상시키는 방법
- 스케일 아웃(Scale-Out)
 - 물리적으로 서버를 늘려서 시스템을 확장하는 방법
 - 서버가 늘어나기 때문에 로드 밸런싱 기술이 필수
- 오토 스케일
 - 서비스가 집중되었을 때 서버 CPU 및 메모리 사용량이 일정 사용량을 초과하면 자동으로 가상 서버 대수를 증감시키는 기능
 - 접속량이 돌발적으로 혹은 대량으로 증가하면 서버의 대수를 늘려서 대응하고, 접속량이 줄어들면 서버의 대수를 줄여서 불필요한 비용을 절약할 수 있게 함
- 데브옵스
 - 개발팀은 보통 웹 사이트 개발, 앱 개발 그리고 알고리즘 개발 하는 부서
 - 운영팀은 개발된 아이템을 고객들에게 전달해주는 부서
 - 데브옵스는 개발팀과 운영팀을 의미. 두 팀 간의 소통이 원활히 이루어질 수 있도록 하는것을 의미



[그림 1-7-7] 데브옵스

AWS

- 리전

- AWS는 물리적으로 떨어진 지역에 여러 개의 클라우드 인프라를 운영하는데, 이 지역을 리전(region)이라고 부름
- 리전이 이렇게 여러 곳에 있는 이유는 네트워크 속도 때문임
- AWS는 전 세계 20여 곳의 리전을 제공하고 있고, 각 리전은 물리적으로 완전히 분리되어 있으며, AWS 콘솔 상에서도 완전히 다른 리소스로 구분
- 내가 만약 서울에서 만든 EC2나 서비스등은 도쿄나 싱가포르 리전에서 만든 EC2와는 완전히 별개의 리소스

- EC2

- Elastic Compute Cloud의 약자로, 종량제 형태로 과금이 되는 가상 서버
- 인스턴스는 AWS에 구축된 1대의 가상 서버를 의미

- Lambda

- AWS 람다는 특정한 이벤트를 통해 입력 인자가 발생하면 연산 과정으로 출력 인자를 바꾸는 이벤트 기반의 컴퓨팅 서비스
- EC2와는 다르게 오직 소스코드만 사용하여 인프라에 대한 관리가 필요하지 않은 완전 관리형 컴퓨팅 서비스

- S3

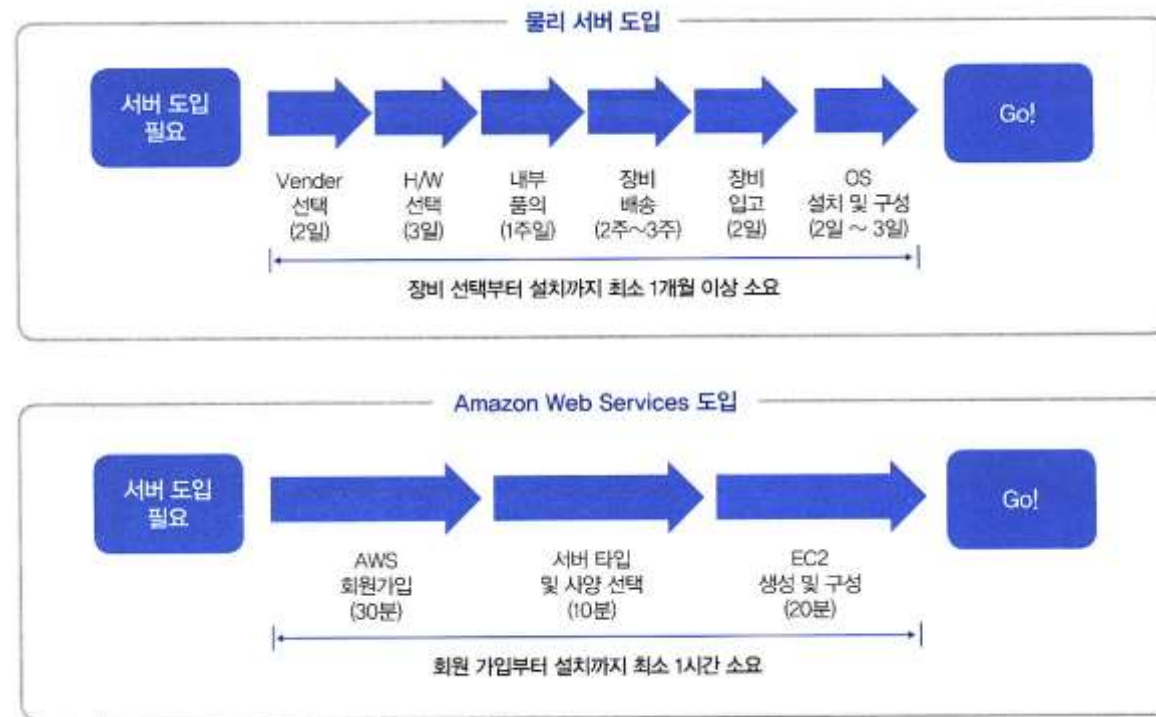
- AWS의 저장소
- 뛰어난 안정성과 보안이 완벽한 스토리지
- 서비스 운영시 생성되는 이미지, 동영상, 오디오, 머신러닝 모델 산출물등 모든 것을 저장할 수 있음

AWS

- EBS
 - Elastic Block Storage의 약자로 EC2와 연결할 수 있는 저장 장치 서비스
 - EC2의 하드디스크, SSD 스냅샷을 통해 언제든지 EC2를 백업할 수 있고 문제가 생겼을 때 복원할 수도 있음
- VPC
 - Virtual Private Cloud의 약자로 AWS의 VPN(Virtual Private Network) 구축 지원 서비스
 - 서비스 보안 수준을 결정하거나 EC2나 RDS같은 AWS 서비스들에 적합한 권한이 있는 사용자들만 접속할 수 있게 함
 - 가상 네트워크망이란 실제 물리적으로 다른 장비를 사용하거나 떨어져 있지는 않지만, 논리적으로 실제 다른 네트워크망에 있는 것처럼 분리해주기 때문에 '가상'이라는 개념을 사용한 네트워크 전용망
- EMR
 - 빅데이터 분석 플랫폼인 하둡 프레임워크를 AWS의 다른 서비스들과 연동하여 더욱 쉽고 빠르게 사용할 수 있는 서비스

EC2

- 크기를 조정 가능한 컴퓨터 파워를 제공하는 웹 서비스
- 가상화 서버를 인스턴스라고 부르며, 한개에서 수천 개의 인스턴스로 손쉽게 컴퓨터 파워를 확장할 수 있음



[그림 2-8] 물리 H/W 도입 vs Amazon Web Services 도입

EC2 - 인스턴스 유형

- 인스턴스 유형

- 범용(M), 컴퓨팅 최적화(C), 스토리지 최적화(I, D), GPU 최적화(G), 메모리 최적화(R)



[그림 2-9] EC2 인스턴스 유형



[그림 2-10] EC2 인스턴스 유형에 대한 설명

EC2 - EBS

- Amazon EBS 볼륨 유형
 - 5가지의 서비스로 분류

유형	범용 SSD	프로비저닝된 IOPS	처리량 최적화 HDD	콜드 HDD	마그네틱
설명	다양한 트랜잭션 워크로드 처리	지연 시간에 민감한 고성능 처리	자주 액세스 하며 처리량 집약적 HDD	액세스 빈도 낮은 저비용 HDD	빈도가 낮으며 성능 낮은 HDD
사례	부트 볼륨	I/O 집약적인 NoSQL, RDBMS	빅데이터, 로그 처리	일별 스캔 횟수 작업 데이터	드문 데이터 액세스
API 이름	gp2	io1	st1	sc1	standard
볼륨 크기	1GB~16TB	4GB~16TB	500GB~16TB	500GB~16TB	1GB~1TB
최대 IOPS	10,000	32,000	500	250	200
최대 처리량	160MB/S	500MB/S	500MB/S	250MB/S	4 ~ 90MB/S

- IOPS
 - HDD, SSD 같은 저장장치의 속도를 나타내는 데 사용되는 단위

CLI

- CLI
 - Command Line Interface(CLI)를 사용하여 aws의 기능을 명령어를 사용하여 시작, 나열 및 종료
- 설정
 - 보안 자격 증명 메뉴에 들어간다.
 - 새 액세스 키 만들기
 - 다운로드 받은 CSV 파일안에 KeyId와 Key 값이 들어 있다
 - command 창을 열어 아래 명령어를 실행한다.
 - aws configure
 - KeyId와 Key값을 csv에 있는 내용을 입력하고 리전 네임을 입력한다.
 - 참고로 서울의 리전 이름은 아래와 같다.
 - ap-northeast-2
 - output format은 아래와 같이 설정한다.
 - json

CLI

- ec2 instances list
 - `aws ec2 describe-instances --no-cli-pager`



Git server



Was



Web Server



개발용

