

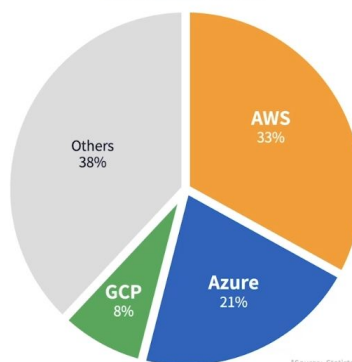
2. AWS Foundation

Introduction to AWS
 AWS Compute Services
 AWS Global Infrastructure
 AWS Region & AZ
 AWS Storage Services
 AWS Database Services

Why Amazon Web Services?

Amazon Web Services(AWS)는 세계적인 클라우드 서비스 사업자로 200개가 넘는 서비스를 제공하고 있으며 전 세계적으로 정부기관, 대기업을 비롯 스타트업과 중소기업 등 다양한 고객을 보유 중이며 국내에서도 압도적인 시장 점유율을 보이고 있음

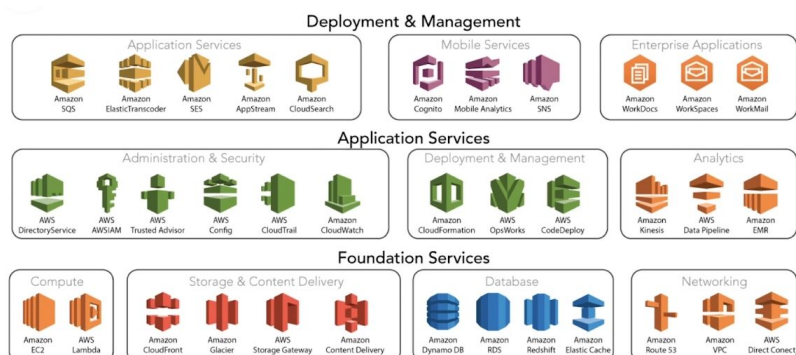
2022 Q1 Market share of Cloud services vendor



*Source: Statista 2022

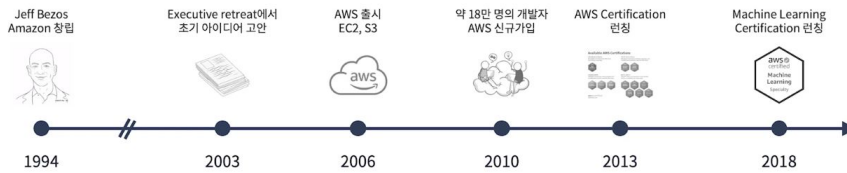
Amazon Web Services

AWS는 인프라와 기초 서비스 뿐만 아니라 사용자 니즈에 맞는 다양한 애플리케이션 서비스를 제공



History of Amazon Web Services

AWS는 Amazon을 운영하면서 쌓은 노하우를 바탕으로 기획, 개발된 서비스로 2006년 처음 출시



AWS Benefit: Worldwide Deployment

AWS 글로벌 인프라를 통해 짧은 시간 안에 전 세계에 애플리케이션을 배포할 수 있음

AWS 외에 글로벌 CSP인 GCP, Azure도 해외 리전으로 빠르게 서비스를 배포할 수 있어

글로벌 기업들은 글로벌 클라우드 서비스를 사용해 전 세계에 빠르고 안정적인 서비스 제공 가능합니다.



강의에서는 Foundation service에 해당하는 내용을 다룰 것이다.

Amazon Web Services

AWS는 인프라와 기초 서비스 뿐만 아니라 사용자 니즈에 맞는 다양한 애플리케이션 서비스를 제공

Copied



프로비저닝: 하드웨어를 구매해서 필요한 기본적인 os 소프트웨어들을 설치해서 deploy하는 일련의 과정을 말한다.

AWS Lambda는 GCP에서는 function에 해당한다.

자기에 필요에 맞는 컴퓨팅 환경을 고려해서 인스턴트 선택가능

General-purpose instance는 여러가능 기능을 균형적으로 제공하는데에 초점이 맞춰진 인스턴스

Compute-optimized: 컴퓨팅에 최적화된 인스턴스, c인스턴스들은 인텔칩종에서도 성능이 좋은 인스턴스

Memory-optimized: 메모리에 최적화된 인스턴스로 메모리 사용량이 많은 데이터베이스, 처리작업을 수행할때 그런 용도로 사용된다.

Accelerated-computing: 딥러닝쪽에서 사용하는 인스턴스로 가속기를 사용한다. 어떠한 목적으로 얼마나 사용할 건지 정확히 알 고있어야 한다. gpu를 사용하므로 비용이 높다. 따라서 필요한 상황에만 사용한다면(인스턴스를 중지하고 다른 인스턴스 타입으로 변경) 효율적으로 사용가능

Storage-optimized: 데이터 웨어하우징 어플리케이션에 적합한 인스턴스

2. AWS Foundation

30

AWS Compute Services

AWS에서는 Amazon Elastic Compute Cloud (일명 EC2) 등의 Compute 서비스들을 통해 애플리케이션을 호스팅하거나 고객이 필요로 하는 컴퓨팅 파워를 제공

EC2와 같은 가상 서버를 통해 AWS 컴퓨팅 자원에 접근할 수 있고
EC2를 사용하게 되면 온프레미스로 자체 서버를 마련해 운영하는 것보다 프로비저닝에 들어가는 비용과 시간이 크게 절약됩니다.
AWS Lambda와 같이 서버를 프로비저닝 하거나 관리하지 않고 코드를 실행할 수 있는 서비스도 있습니다.

Amazon EC2

- 몇 분의 시간만으로도 EC2 인스턴스 프로비저닝 가능
- 작업을 완료했거나 사용하지 않는 인스턴스는 중지 가능
- 실행한 만큼만 비용을 지불하여 예산 절약 가능

General-purpose instances

Instance

Instances

Spot Instance

T3 instance

M4 instance

M1 Mac Instance

Compute-optimized

C4 instance

C5 instance

C5d instance

Memory-optimized

R5 instance

R4 instance

X1 instance

Accelerated-computing

P3 instance

G4dn instance

Inf1 instance

Storage-optimized

I3 instance

D3 instance

DB instance

리전을 많이 가지고 있어 예를들어 서울에서 문제가 발생한다 하더라도 근처에 있는 도쿄 리전을 통해서 migration을 통해서 운영을 바로 대응을 할 수있다. 현재 서울에도 리전이 있다.

한국에서 금융권이나 의료분야에서는 데이터 센터 한국이 물리적으로 위치해야한다! 그러다 보니 이러한 compliance를 그동안 만족하지 못해서 클라우드 도입이 쉽지않았는데 요즘 가능해 지고 있다.

특정 리전마다 제공하는 tpu가 다른경우가 있음(구글 gcp 의 경우) 어떤 리전만 좋은 tpu를 제공하는 경우가 존재한다. 이러한 경우를 고려하여 리전을 선택해야 한다

또한 같은 인스턴스라고 하더라도 나라마다 환경이 달라 전기료(+세금)가 다르게 책정된다. 따라서 리전에 따라 비용이 다르게 설정되므로 이것도 고려해야한다

AWS Global Infrastructure

AWS는 Region(리전)이라는 그룹으로 데이터 센터를 구축하여 하나의 건물에서 화재나 재난이 발생해도 안정적인 운영이 가능
오하이오, 더블린, 도쿄, 파리 그리고 더블린과 같은 전 세계 곳곳에 트래픽이 많이 발생하는 지역에 데이터 센터를 구축

서울 리전은 미국 동부, 서부, 도쿄에 이어 네번째로 가용영역을 4개 이상 가진 리전입니다.
각 리전은 격리되어 있지만 고속 광섬유 네트워크를 통해 다른 리전과 연결은 가능합니다.
비즈니스 요구사항 이외에도 리전을 선택할 때 여러가지 요인을 고려해야 합니다.

리전 선택 시 고려해야 할 요인

Compliance	데이터가 특정 국가/지역을 벗어나면 안되는 요구사항 있을 때 또는 리전을 지정해야 하는 Compliance 요구사항이 있는지
Proximity	서비스 제공 및 정보 전송에 지연이 발생하면 안되는지
Availability	가장 가까운 리전에서 원하는 기능을 제공하고 있는지
Price	동일한 하드웨어라고 하더라도 리전에 따라 비용이 어떻게 달라지는지

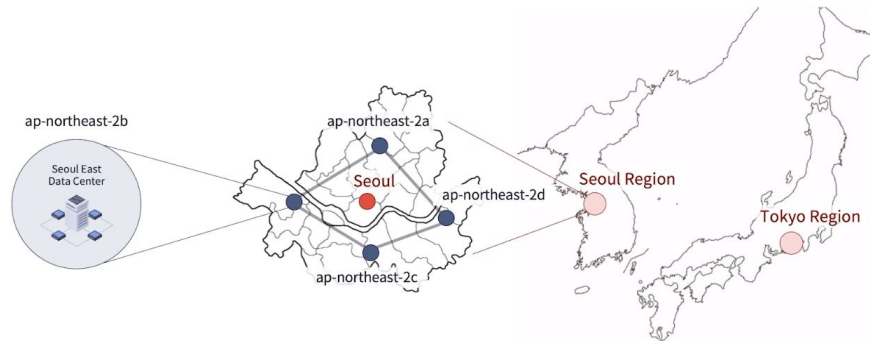
리전안에 가용영역이 2개이상 구성되어있을때 리전이라고 말한다, 서울에 4개의 가용영역이 있다고 가정 하나의 데이터 센터가 고장나도 나머지 3개의 데이터 센터가 빠르게 대응할 수있도록 구성되어있음 따라서 장애가 생겨도 빠르게 대응할 수있도록 인프라가 구축되어있다.

Regions & Availability Zones

Region(리전)은 지리적 위치를 의미하며 최소한 2개 이상의 가용 영역으로 구성됨

Availability Zone(가용 영역)은 리전 내 단일 또는 그룹으로 구성된 데이터 센터를 의미

각 가용 영역은 다른 가용 영역의 장애로부터 영향을 받지 않을 수 있도록 분리되어 있으나 지연 시간은 짧은 링크로 연결되어 있음



S3의 경우 최소 3개의 가용영역에 데이터를 저장하기 때문에 안정성이 높다

AWS Storage Services

AWS에는 크게 EBS와 S3 두 가지 스토리지를 많이 사용

Elastic Block Store(EBS)는 Amazon EC2 인스턴스에서 사용할 수 있는 volume을 제공하는 스토리지 서비스

Simple Storage Service(S3)는 데이터, 메타데이터, 키 세 개로 구성된 객체를 bucket에 저장하는 스토리지 서비스

복잡한 읽기, 쓰기, 변경 기능을 수행하는 경우에는 Amazon EBS를 사용하는 것이 일반적이고,
완성된 객체를 사용하고 변경 횟수가 적은 경우에는 Amazon S3를 사용 하는 것이 일반적입니다.

Amazon EBS

- EBS volume은 데이터를 보존해야 하는 경우에 적합
- Snapshot을 생성하여 EBS volume을 백업
- incrementally 최근의 스냅샷 이후에 변경된 데이터 블록만 저장
- EC2 instance를 중지/종료하더라도 데이터는 보존됨

Snapshot

Volume

Multiple volumes

Amazon S3

- 데이터에 자주 접근하지만 변경은 하지 않는 경우에 적합
- Object는 데이터, Bucket은 파일 디렉토리라고 볼 수 있음
- 이미지, 동영상, 텍스트 파일 등 모든 유형의 파일 저장 가능
- 최소 3개의 가용 영역에 데이터를 저장

Bucket

Object

Bucket with objects

Access points

AWS Database Services

AWS 클라우드에서 관계형 데이터베이스를 실행할 수 있는 RDS와 비관계형 데이터를 위한 NoSQL DB, Document DB 등 다양한 서비스를 제공
Amazon Relational Database Service는 하드웨어 프로비저닝, DB 설정, 패치 업데이트 적용, 백업 등의 작업을 자동화하여 관리해주는 서비스
Amazon DynamoDB는 key-value의 구조로 된 데이터를 저장하고 쿼리할 수 있는 데이터베이스 서비스

Amazon RDS는 6개의 데이터베이스 엔진을 사용할 수 있습니다.

- Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database, Microsoft SQL Server)



Amazon Aurora

- MySQL, PostgreSQL과 호환되는 엔터프라이즈 레벨의 RDB
- 표준 관계형 데이터베이스보다 최대 3배 ~ 5배 빠름
- Amazon Aurora는 6개의 데이터 복사본을 3개의 가용 영역에 복제하고 지속적으로 Amazon S3에 데이터를 백업



Amazon DynamoDB

- 테이블을 생성해 키(attribute), 값(item)의 pair로 데이터 저장
- 서버를 사용하지 않아 프로비저닝, 패치 적용 등의 관리 불필요
- 소프트웨어 설치, 유지보수, 운영 불필요
- 데이터베이스 용량 변화에 맞게 크기를 조정하면서 성능 유지



Table



Attributes



Items