



모듈 5 - 스토리지와 데이터베이스

Amazon Elastic Block Store(EBS)

EC2를 이용해서 비즈니스 애플리케이션을 수행하게 된다면 cpu, 메모리, 네트워크 그리고 스토리지에 액세스 할 수 있다

여기서 스토리지 액세스에 대해 설명을 해보겠습니다.

대부분의 경우에는 블록 수준 스토리지에 액세스를 해야합니다.



스토리지 : 데이터를 보관하는 장소 (예:usb, 외장하드)

블록 스토리지

파일을 저장하는 곳입니다. 이 파일은 바이트로 구성되며 디스크의 블록에 저장됩니다. 이때 파일을 수정되면 모든 블록에 덮어쓰기가 실행되지는 않습니다. 변경된 그 부분만 업데이트를 합니다.

EC2 인스턴스를 시작할 때에는 유형에 따라 인스턴스 스토어 볼륨이라는 로컬 저장소를 제공할 수 있습니다.

이 볼륨이 기본적으로 호스트에 물리적으로 직접 연결이 되어 있습니다. 따라서 EC2를 중지하거나 종료하면 인스턴스 스토어 볼륨에 작성한 모든 데이터가 삭제되어 임시 파일, 쉽게 재생성할 수 있는 데이터에 적합합니다.

휘발되어서는 안되는 데이터는 ec2와 함께 제공되는 인스턴스 스토어 볼륨에서 작성해서는 안 됩니다.

그럼 중요한 데이터는 어디에 두어야 할까요 ?

이런 중요한 데이터는 **Amazon Elastic Block Store(EBS)**라고 하는 서비스를 사용하면 됩니다. 여기에 저장하면 EC2 인스턴스의 수명 주기가 끝나도 계속적으로 지속될 수 있습니다.

EBS를 사용하면 EBS 볼륨이라고 하는 가상 하드 드라이브를 만들고 그 드라이브를 EC2에 연결할 수 있습니다.

다양한 크기와 유형으로 제공되는데 사용자가 원하는 대로 볼륨을 프로비저닝하고 ec2와 연결합니다. 그리고 나면 볼륨에 데이터를 작성하도록 구성할 수 있습니다.

- EBS 특징

EBS를 사용하면 **스냅샷**을 만들 수 있습니다.



스냅샷: 스냅샷은 말 그대로 특정 시점에 포인트를 찍어서 그 시점으로 되돌아갈 수 있는 지점을 만드는 기능입니다.

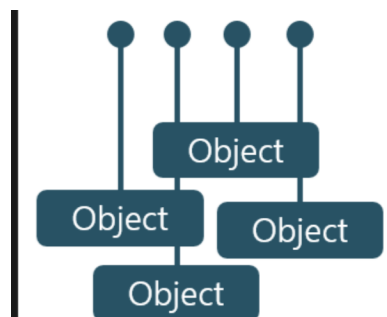
증분식 백업 방식을 이용하여 마지막 스냅샷 이후 변경되는 데이터 블록만 기록하고 복제하므로 저장 비용과 시간도 효과적으로 절감할 수 있습니다.

Amazon Simple Storage Service(Amazon S3)

AWS 서브수 중에서 EC2 서비스와 더불어 가장 오래되고, 기본이 되는 객체 스토리지 서비스입니다.



객체 스토리지: 각 데이터 조각을 가져와서 객체로 지정하고, 개별 단위로 저장합니다.



S3에 저장되는 데이터를 객체라고 하며, 이 객체 저장소를 버킷이라고 합니다.

S3 특징

- 99.999999999% 내구성을 가지고 있어 데이터 손실을 최소화
- 최대 5TB 크기의 객체를 업로드 가능
- 객체 버전 관리
- 여러 버킷을 생성
- 액세스 접근 권한을 설정 가능
- 웹 접속이 가능하기 때문에 간단한 정적 웹 콘텐츠를 S3에 올려 웹 서버의 도움없이 바로 웹 서비스가 가능

S3 스토리지 클래스 유형

- Standard: 자주 액세스 하는 데이터 (빅데이터 분석용 데이터, 작업용 임시 백업 데이터)
- Standard-Infrequent Access: 액세스 빈도가 낮지만 빠른 접근이 필요한 데이터 (파일 동기화, 백업, 재해 복구용 데이터)
 - standard-infrequent access: 최소 세개 이상의 가용 영역에 데이터가 저장되기 때문에 내구성이 높은 클래스
 - one zone-infrequent access: 하나의 가용 영역에만 데이터가 저장되어 상대적으로 내구성이 낮고, 삭제되더라도 다시 생산할 수 있는 데이터를 위한 클래스
- glacier : 거의 액세스하지 않는 데이터 (감사 데이터, 장기 보존용 데이터)

S3 수명 주기 정책을 사용할 수 있습니다.



S3 수명 주기 관리: 계층 사이에서 데이터를 자동으로 이동

수명 주기를 사용하면 데이터는 여러 계층을 자동이동하게 됩니다. standard 90일 유지 > IA 로 30일 동안 이동 > 120일 지나면 Glacire 에 옮길 도록 지정할 수 있다.

Amazon Elastic File System(Amazon EFS)

관리형 파일 시스템입니다. EFS를 사용하면 기존 파일 시스템을 그대로 유지하면서도 규모를 조정하거나, 복제하는 등의 작업을 AWS에게 맡길 수 있습니다.

특징

- 여러 인스턴스가 EFS의 데이터에 동시에 액세스 가능
- 사용자가 작업을 수행하지 않아도 필요에 따라 확장되거나 축소됨

EBS와 EFS의 차이점

아마존 EBS 은 EC2 인스턴스에 연결이 되고 가용 영역 수준의 리소스입니다. EC2 인스턴스를 연결하려면 동일한 가용 영역에 있어야 하며 볼륨이 자동으로 확장되지 않습니다.

EFS는 여러 인스턴스를 동시에 읽기와 쓰기가 가능합니다. 리눅스를 위한 파일 시스템이며 리전 기반의 리소스이기도 합니다. 리전에 있는 모든 EC2 인스턴스가 EFS 파일 시스템의 데이터를 사용할 수 있다는 말입니다. 필요에 따라 자동으로 확장됩니다.

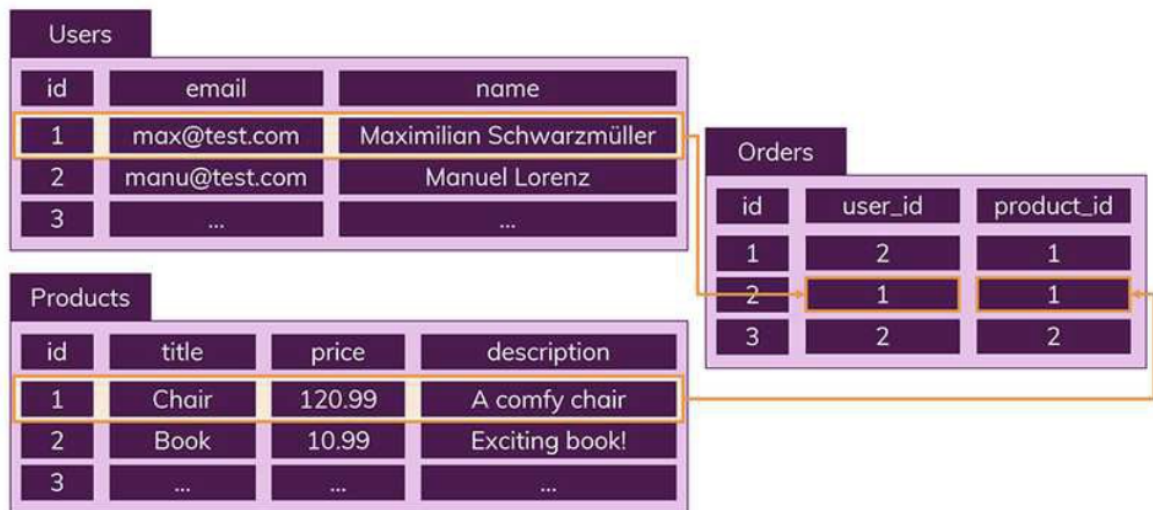
Amazon Relational Database Service(Amazon RDS)

: 클라우드 환경에서 관계형 데이터베이스를 간편하게 설정하고 운영할 수 있는 서비스

- 데이터: 문자와 숫자, 그림, 영상 등의 형태로된 단위
- 데이터베이스: 데이터를 하나로 모아두는 것
- DBMS: 데이터베이스를 관리하는 시스템



관계형 데이터베이스 모델: 데이터의 논리적 관계를 초점으로 특정 개체 정보를 이용하여 열과 행으로 된 테이블 이라는 최소 단위로 구성



RDS 데이터베이스 엔진

- Aurora
- postgresql
- mysql
- mariadb
- oracle
- microsoft SQL server

리프트 앤 시프트

데이터센터에서 관리하는 데이터를 클라우드에 쉽게 옮기는 방법입니다.

이 작업을 수행하여 데이터베이스를 마이그레이션해서 EC2에서 실행하게 됩니다. 메모리, cpu, 스토리지 용량 같은 변수를 제어할 수 있습니다.

RDS 특징

- 자동 패치
- 백업
- 이동화
- 장애 조치
- 재해 복구

Amazon Aurora

aws 자체의 클라우드 데이터베이스 엔진입니다.

- mysql, postgresql 호환
- 상용 db 비용의 1/10
- 최소 3개의 가용 영역에서 두 개씩 총 여섯 개의 복제 데이터를 가짐
- 읽기 전용 복제본을 15개 제공
- S3 지속적인 백업 가능
- 특정 시점으로 복구 가능

Amazon DynamoDB

비관계형 데이터베이스의 유형으로, 키-값 메소드를 사용하는 key-value 데이터베이스입니다.

키	값
1	<p>이름: John Doe</p> <p>주소: 123 Any Street</p> <p>좋아하는 음료: 미디엄 라떼</p>
2	<p>이름: Mary Major</p> <p>주소: 100 Main Street</p> <p>생일: 1994년 7월 5일</p>

여기에서 비관계형 데이터베이스는 관계형 데이터베이스와 다르게 데이터가 서로 연결되지 않는 개별 형태로 저장하여 복잡하고 구조화되지 않은 데이터 유형에 적합한 데이터 베이스

입니다.

sql문을 사용하지 않기 때문에 **NoSQL 데이터베이스**라고도 합니다.

특징

- 구조가 단순하여 빠른 처리 가능 **10밀리초 미만의** 처리 성능 제공
- **서버리스**
- 용량에 따라 테이블을 자동으로 확장, 축소하여 관리 편의성이 높음

SQL을 사용하는 경우, NoSQL을 사용하는 경우

6. 알겠습니다. 그래서 뭘 언제 사용해야 하나요?

SQL를 사용해야 할 경우:

- 복잡한 쿼리나 리포트를 만들어내야 하는 경우
- 높은 부하가 예상될 경우(탄탄함!)
- 데이터가 크게 쌓이게 될 여지가 별로 없는 경우(정말 데이터가 많으면 압도적으로 NoSQL이 빠름)

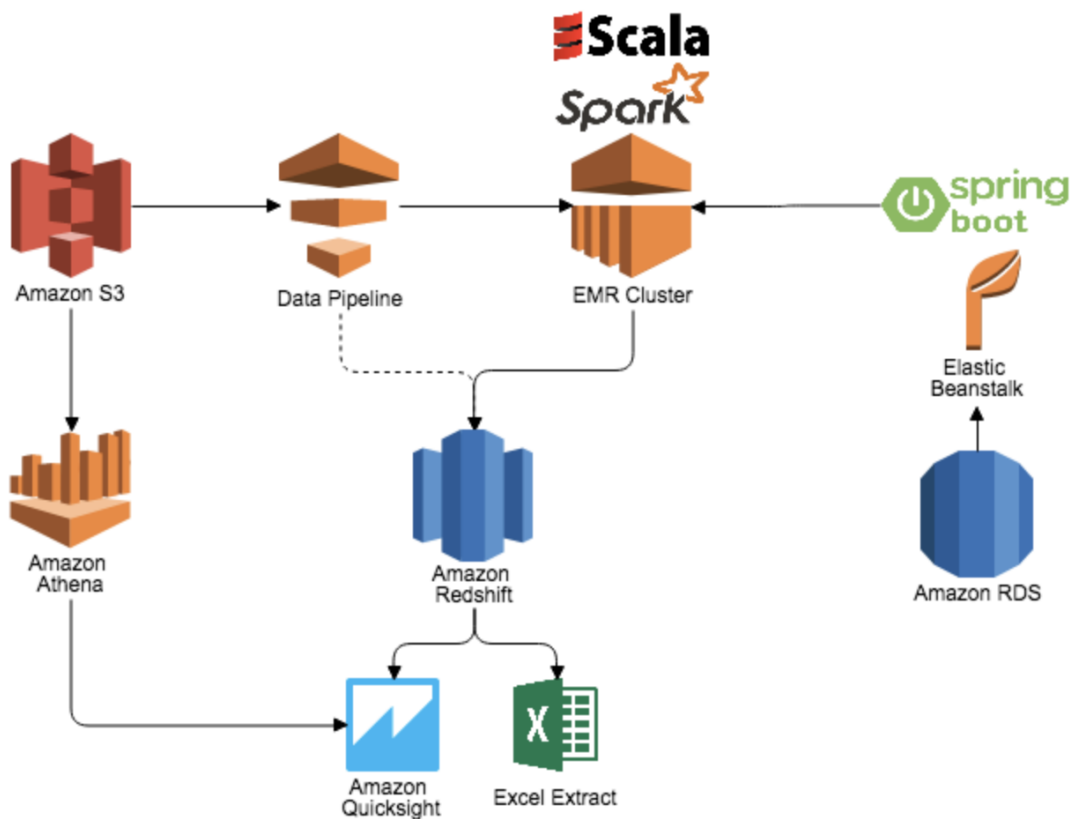
크게 = 수억개~ (사실 10억도 괜찮긴 한데..)

NoSQL을 사용해야 할 경우

- 지속적으로 새로운 기능들이 추가되고 변경되어 데이터셋을 사전정의하기 힘든 경우
- 데이터 정합성이나 일관성이 최고 우선순위가 아닌 경우
- 데이터가 너무너무 많을 경우

Amazon Redshift

AWS에서 제공하는 **관계형 데이터베이스 서비스**로서, 대용량 데이터를 신속하게 분석하고 처리할 수 있는 **데이터 웨어하우스**입니다. 열 지향형 컬럼 스토리지를 사용하여 데이터 압축 및 처리 성능을 최적화하고, 병렬 처리를 통해 대용량 데이터셋을 효과적으로 처리할 수 있습니다.



대용량 데이터셋에서의 분석 및 BI 작업에 효과적입니다.

AWS Database Migration Service(AWS DMS)

관계형 데이터베이스, 비관계형 데이터베이스 및 기타 유형의 데이터 저장소를 마이그레이션할 수 있는 서비스입니다.



마이그레이션: 한 형식에서 다른 형식으로 이동하는 프로세스

DMS를 사용하면 원본 데이터베이스와 대상 데이터베이스 간에 데이터를 이동할 수 있습니다.

예를 들어 온프레미스에서 EC2 or RDS에 저장된 mysql이 있다고 가정해보면, mysql db를 DMS를 이용해서 aurora db와 같은 db로 데이터를 마이그레이션할 수 있습니다.

추가 데이터베이스 서비스

- Amazon DocumentDB
- Amazon Neptune
- Amazon Quantum Ledger Database(QLDB)
- Amazon Managed Blockchain
- Amazon ElasticCache
- Amazon DynamoDB Accelerator