

AWS S3

S3란?

- S3는 Simple Storage Service의 약자로 파일 서버의 역할을 하는 서비스다.
- S3는 AWS에서 제공하는 인터넷용 스토리지 서비스(=저장소)로, 원하는 양의 데이터를 저장, 검색, 삭제할 수 있다. 클라우드 서비스 관점에서 구글 드라이브와 비슷하게 생각하면 된다.
- 일반적인 파일서버는 트래픽이 증가함에 따라서 장비를 증설하는 작업을 해야 하는데 S3는 이와 같은 것을 대신해준다. 왜냐하면 S3는 AWS라는 클라우드 서비스로 제공되는 파일서버이기 때문에 관리자가 직접 장비를 증설할 필요성이 사라지기 때문이다.
- 또한, S3는 파일에 대한 접근 권한을 지정할 수 있다.

S3의 주요특징

- 사용자가 증가하여도 관리자는 장비 증설과 같은 시스템 적인 작업을 수행하지 않아도 된다.
- 저장할 수 있는 파일 수의 제한이 없다.
- 최소 1바이트에서 최대 5TB의 데이터를 저장하고 서비스할 수 있다.
- 파일에 인증을 붙여서 무단으로 접근하지 못하도록 할 수 있다.
- HTTP와 BitTorrent 프로토콜을 지원한다.
- REST, SOAP 인터페이스를 제공한다.
- 데이터의 손실이 발생할 경우 자동으로 복원한다.
- 버전관리 기능을 통해서 사용자에게 의한 실수도 복원이 가능하다.
- 정보의 중요도에 따라서 보호 수준을 차등할 수 있고, 이에 따라서 비용을 절감할 수 있다. (RRS)

S3의 주요용어

- 객체(Object)
: AWS는 S3에 저장된 데이터 하나 하나를 객체(=파일)라고 명명한다.
- 버킷(Bucket)
: 객체(=파일)들을 그룹핑한 최상위 디렉토리라고 할 수 있다. 즉, S3에 저장된 객체에 대한 컨테이너이다.
: 버킷 단위로 지역(region)을 지정할 수 있다.
: 버킷에 포함된 모든 객체에 대해서 일괄적으로 인증과 접속 제한을 걸 수 있다.
- 키(Key)
: 버킷 내 객체의 고유한 식별자로 버킷 내 모든 객체는 정확히 하나의 키를 갖는다.
: 예를 들어, data버킷에 키가 photos/puppy.jpg인 객체를 저장할 수 있다.

- 버전관리
: S3에 저장된 객체들의 변화를 저장한다.
: 예를 들어, A라는 객체를 사용자가 삭제하거나 변경해도 각각의 변화를 모두 기록하기 때문에 실수를 만회할 수 있다.
- BitTorrent
: 분산된 파일 배포 시스템이라고 정의할 수 있다.
: BitTorrent의 분산이란 하나의 서버에서 파일을 배포하는 것이 아니라, 파일을 가지고 있는 컴퓨터들에게 조금씩 파일을 다운받은 후에 이것을 붙여서 완전한 파일을 만드는 방식이다.
- RSS
: S3의 Storage Class중 하나이다.
: Reduced Redundancy Storage의 약자로 일반 S3 객체에 비해서 데이터 손실될 확률이 높은 형태의 저장 방식이다.
: 그럼에도 불구하고 물리적인 하드 디스크 대비 400배가량 안전하다는 것이 아마존의 주장이다.
- Glacier
: S3의 Storage Class중 하나이다.
: 영어로는 빙하라는 뜻으로 매우 저렴한 가격으로 데이터를 저장할 수 있는 아마존의 스토리지 서비스이다.

S3의 접근권한

- S3는 두 가지 범주의 작업(Operation)이 있다.
 - Bucket Operation : 버킷의 생성, 삭제, 속성변경 등 버킷 단위의 작업
 - Object Operation : 객체의 생성, 삭제, 다운로드 등 객체 단위의 작업
- 위 두 가지 범주의 작업을 제어하는 권한계층에는 다음 세 가지가 있다.

항목	Bucket ACL	Object ACL	Bucket Policy
Bucket 작업 제어	O	X	O
Object 작업 제어	X	O	O
접근정책 형태	허용정책만 가능	허용정책만 가능	허용/거부 모두 가능
권한정의 형식	XML	XML	JSON
정의하는 권한범위	일부	일부	전체

- Bucket ACL : 버킷 단위 작업 권한 제어, XML 형식으로 접근권한 정의 (예전방식)
- Object ACL : 객체 단위 작업 권한 제어, XML 형식으로 접근권한 정의 (예전방식)
- Bucket Policy : 버킷 및 객체 단위 작업 권한 제어, JSON 형식으로 접근권한 정의 (요즘방식)

S3의 Storage Class

- Storage Class는 버킷 단위로 고정된 것이 아니며, 동일한 버킷 내에서도 각 객체별로 다른 Storage Class를 선택하여 저장할 수 있다.
- S3에서 제공되는 Storage Class는 총 네가지가 있으며, 각 Class별로 보장하는 내구성과 가용성, SLA(서비스 수준 계약) 기준이 다르다.
- Storage Class

Storage Class	Standard	Standard-IA	RRS	Glacier
내구성	매우 높음	매우 높음	높음	매우 높음
가용성	매우 높음	높음	매우 높음	매우 높음
단가				
용도	범용	낮은 접근빈도	저중복	아카이빙

참고문헌

- 생활코딩 (2012). <https://opentutorials.org/course/608/3006>.
- seoyeon hwang (2020). <https://medium.com/@seoyeonhwng/aws-s3%EB%9E%80-%EB%AC%B4%EC%97%87%EC%9D%B8%EA%B0%80-b0da502b0504>.
- 악스토리 (2019). <https://acstory.tistory.com/33>.