AWS_210727

AWS Storage Service
블록 스토리지 , 데이터, 서버 인스턴스에 디스크 볼륨 형태로 제공되는 데이터를 의미.
EC2 인스턴스에 포함된 볼륨에 고속으로 접근가능. 블록 저장의 표적 서비스인 EBS, 즉 Elastic Block Store는 EC2 인스턴스를 위한 부트 볼륨 및 데이터 베이스로 널리 사용됨.

파일 스토리지 : 파일 스토리지에서 데이터란, 서버 인스턴스에 파일 시스템 인터페이스 또는 파일 시스템 시멘틱스 방식으로 제공되는 데이터를 의미하며, 서버 인스턴스에 파일 스토리지를 가하면 로컬 파일 시스템 처럼 작동한다. EFS, 즉 Elastic File System은 고속으로 다수의 EC2 인스턴스를 통해 데이터에 접근할 수 있도록 한 다.

아마존 심플 스토리지 서비스(S3)
전 세계 어디에서나 대규모 데이터를 저장하고 인출할 수 있는 인터페이스를 제공. 저장용량은 무제한, 99.99999999999%에 이르는 고신뢰성을 제공. S3는 두 개의 저장 시설에서 동시다발적인 손실 상황을 견디도록 설계됨.

아마존 심플 스토리지 서비스(S3) : S3는 다른 저장소와 근본적으로 다른 방식의 설계로 파일 시스템이 존재하지 않으며, 모든 객체는 S3 버킷(bucket)에 단순한 네임스페이스만으로 저장. 또 S3는 지역별 서비스로서, 지역별 재난 상황에 처할 수 있도록 자동 반복 저장됨.

아마존 S3의 주요 장점

간편성 : S3는 사용하기 편리하며, 직관적인 웹 기반 콘솔을 이용해 데이터 업로드, 다운로드, 관리 등의 업무 를 수행할 수 있으며, S3를 관리하기 위한 다수의 모바일 앱도 제공함. REST API와 SDK도 제공.

확장성 : S3는 무제한의 확장성이 있음. 별도의 준비 작업 없이도 무제한의 데이터를 저장가능. S3에 페타바이트 급 데이터를 저장하는 작업도 간단하게 가능하며, 비즈니스 니즈에 맞춰 언제든 저장 용량을 확장 및 축소기능 있음.

신뢰성 : S3는 99.999999999%의 신뢰성으로 객체를 저장할 수 있는 유일한 서비스로, 이를 위한 아마존만의 인프라 설계 및 운영 능력이 발휘되는중. 데이터는 반복 방식으로 다수의 데이터 센터와 다수의 스토리지 기기에 저장되며, 두 개 시설에서 동시에 재난 상황이 발생해도 데이터가 안전하게 보관됨.

보안성 : S3는 암호화 기능을 제공하며 모든 데이터는 업로드될 때 자동으로 암호화됩니다. • 또한 SSL을 통한 데이터 전송을 지원하고 아마존의 신분확인 및 접근관리 서비스인 IAM을 통해 S3 버킷에 대한 매우 세심한 접근권한 기능을 제공합니다.

고성능 : S3는 네트워크 성능을 최대한 활용할 수 있는 멀티파트 업로드 기능을 지원하고, 지역별 네트워크 지연을 최소화할 수 있도록 최종 사용자의 인접 리전을 고객이 선택가능. 또한 S3와 CloudFront를 연계해 고정 비용 없이 최종 사용자 에게 고속으로 콘텐츠를 전송할 수 있음.

가용성 : S3는 객체 처장과 관련해 연간 99 99%의 가용성을 지니도록 설계되어 SLA 상의 99.99% 활용성 및 가용성은 아래와 같은 비활용성 및 비가용성 수준을 의미. – 일간: 8.6초 – 주간: 1분 0.5초 – 월간: 4분 23.0초 – 연간: 52분 35.7초

저비용 : S3는 높은 비용대비효율성을 제공하며 대량의 데이터를 저비용으로 저장가능. 최소 약정 비용 및 고정 비용은 없으며 사용량에 대한 비용만 부담하면 됨.

관리용이성 : S3의 스토리지 관리 기능을 이용하면 스토리지 최적화를 이루고 보안 및 관리의 효율성을 증대시킬 수 있음. 결과적으로 데이터에 대한 인사이트를 높이고 더 개인화된 메타데이터를 활용가능.

연계성 : S3는 다른 서드 파티 도구와 연계가능. S3 기반의 다양한 애플리케이션을 신속하게 개발하고 배포가능. 또한 S3는 다른 AWS 서비스와의 연계성도 높으므로 다양한 AWS 제품군과 결합해 우수한 서비스 를 구현할 수 있습니다.

데이터 백업 : S3는 기업용 데이터 백업 파일 저장 방식으로 활용. 데이터 손실 기능성이 매우 희박하며, 리전 내 여러 개의 AZ로 분산돼 세 개의 복제물로 관리되므로 특정 AZ가 자연재해 동으로 파괴돼도 데이터는 안전하게 보관가능. 버젼별 관리기능 또한 제공.

테이프 저장 장치 대체 : S3 의 주요 용도 중 하나로 전통적인 테이프 저장 장치의 대체가능. 다수의 기업이 기존의 테이프 드라이브 또는 테이프 인프라를 S3로 대체하고 있음.

정적 웹사이트 호스팅 : S3를 이용하면 웹 서버 설정이나 스토리지에 대해 고민할 필요 없이 정적 웹사이트 호스팅 가능. 또한 S3는 필요에 따라 언제든 확장 가능하므로 트래픽 제한이나 용량 제한에 대해 고민할 필요가 없음.

애플리케이션 호스팅 : S3는 고가용성 스토리지로서 다수의 고객이 모바일 앱 또는 인터넷 기반 앱을 호스팅하는 데 이용가능. 또 전 세계 어디에서나 접근 가능하므로 애플리케이션을 세계 어디에서나 접근 및 배포할 수 있음.

재난 복구 : S3는 재난 복구 전략으로도 활용되며, 크로스 리전 복제 전략을 통해 각각의 S3 객체를 서로 다른
리전, 서로 다른 버킷에 자동으로 복제하여 대비가능.

콘텐츠 배포 : S3는 인터넷을 통한 콘텐츠 배포 방법으로도 활용. 기업의 모든 스토리지 인프라를 클라우드로 이전해 S3의 고확장성, 사용량에 따른 비용부담의 혜택을 누릴 수 있음.콘텐츠는 어떤 유형의 파일이라도 가능.

데이터 레이크 : 기업에서 처리, 분석, 소비되는 막대한 양의 데이터를 보관하는 중앙 저장소역할. 로 데이터(raw data), 반 처리 데이터, 처리 데이터 등 다양한 단계의 기업 데이터를 저장가능. 빅데이터 저장소 활용. AWS의 다양한 서비스 포트폴리오를 통해 이와 같은 빅데이터 관리 비용 을 절감하고, 수요에 따라 확장가능 혁신의 속도를 높일 수 있음.

프라이빗 저장소 : S3를 이용해서 Git. Yum 또는 Maven과 같은 프라이빗 저장소 구현가능.

실생활에서 물을 양동이에 담듯이, 클라우드에서는 객체 데이터를 버킷에 담는다. S3에서 버킷은 객체를 저장하는 컨테이너 역할을 하며, 컴퓨터로 비유하자면 파일을 담는 폴더 역할을 수행. 폴더처럼 버킷에도 여러 개의 폴더를 생성가능.

버킷은 어떤 리전에서도 생성 할 수 있으며, 명시적으로 복제작업을 수행하거나 크로스-리전 복제를 수행하지 않는 한, 다른 리전에 특정 버킷의 데이터가 복제 되지 않음. 또한 S3 버킷은 버전 부여 기능을 제공하므로 객체가 버킷에 추가될 때마다 해당 객체에 유일한 ID가 할당됨. S3에서 객체는 가장 기본적인 요소로 버킷에 저장한 모든 것은 객체라 부르고, 각 객체는 데이터와 메타데이터를 지님. 이때 데이터는 S3에 저장되는 것이고, 메타데이터는 해당 객체를 설명하는 네임밸류 쌍으로 표시.

S3는 API를 통해 접근할 수 있으며, 개발자는 S3 기반의 애플리케이션을 개발가능. S3의 기본 인터페이스는 REST API이며, HTTPS 모드에서 SOAP API를 지원. S3 에서는 REST 방식을 사용권장.

REST API는 스테이트리스(stateless), 클라이언트-서버 기반, 캐시 커뮤니케이션 프로토콜, HTTP 프로토콜 기반의 API. 이를 통해 S3 버킷에서 파일 생성, 읽기, 갱신, 삭제, 목록 조회 등 모든 작업을 수행가능하며, 표준 HTTP/HTTPS 요청과 관련된 모든 작업도 수행할 수 있음. HTTP보다 HTTPS가 안전한 방법이므로, S3 에서 API를 요청을 할 때는 가급적 HTTPS를 사용하는 것이 권장됨. HTTP에서 가장 우선적으로, 많이 사용되는 명령은 POST, GET, PUT, PATCH, DELETE 이며, 이는 각각 생성, 읽기, 갱신, 삭제를 뜻하는 CRUD 작업에 대응.

HTTP Verb	CRUD Operations in Amazon S3
GET	Read
PUT	Create
DELETE	Delete
POST	Create

API 외에도 브라우저와 Android와 iOS 등 모바일 기기, Java, .NET, Node.js, PHP, Python, Ruby, Go, C++ 등 다양한 언어를 지원하는 SDK를 제공. 모바일 SDK와 IoT SDK도 제공. SDK와 API를 함께 적용도 가능. AWS CLI, 즉 커맨드 라인 인터페이스는 전통적인 텍스트 기반 명령어 인터페이스이며, 모든 AWS 서비스를 통합 관리할 수 있음.

키 이름에 Hex Hash 프리픽스 추가하기
최적의 파티셔닝 전략이라고 하기는 어렵습니다. 키 이름에 16진수 계열의 Hex Hash를 픽스로 추가해 적절한 무작위성을 반영하는 것.

아마존 S3의 암호화
아마존 S3에서 데이터를 암호화하는 방법 2가지
1.전송 중인 데이터를 암호화
2.저장된 데이터를 암호화

저장된 데이터의 암호화는 SSE, 즉 Server Side Encryption 을 이용하며, 데이터 를 작성할 때 자동으로 암호화되고 데이터를 출할 때 자동으로 복호화. 이 때 AES 256-비트 대칭키사용

이 때는 AES 256-비트 대칭키가 사용되며 아래와 같은 세 가지 방식으로 키를 관리

AWS KMS 기반의 SSE(SSE-KMS}
AWS의 키 매니지먼트 전문 서비스인 KMS 이용.
마스터 키 사용 맥락에 따라 다양한 퍼미션 설정이 가능하고, S3에 저장된 객체에 대한 승인 접근을 효과적으로 통제할 수 있는 추가 보안 레이어를 제공.
KMS의 감사 기능을 이용해 누가, 언제, 어떤 데이터에 접근했는지 확인하거나 암호화된 데이터를 승인 없이 접근하려는 시도가 있었는지 여부 또한 상세히 파악할수있음. 또한 산업 보안 표준인 PCI-DSS, HIPAA/I-TECH, FedRAMP 등 다양한 요건을 충족하기 위한 추가적 인 보안 통제 기능을 제공.

아마존 S3의 접근성 통제
액세스 컨트롤, 즉 접근성 통제란 S3 버킷에 누가, 어떻게 접근하도록 할 것인지 정의하는 것.
세분화된 통제가능
1접근 정책 : IAM, 즉 신분 및 접근 관리 정책을 통해 S3 의 객체에 대한 매우 세분화된 통제가 가능
2버킷 정책 : 이 외에도 조건을 가해 여러 개의 계정에 한꺼번에 접근을 허용하거나, MFA 인증을 반드시 거치게 하거나, 아마존 클라우드프론트 서비스를 통한 접근만 허용하거나, 특정 HTTP 참조기관의 트래픽을 거 부하는 것도 가능.
3접근 제어 목록 : 세 번째 접근 제어 방식은 ACL, 즉 접근 제어 목록을 이용하는 것.

아마존 S3 스토리지 클래스
아마존 S3는 다양한 상황에 대응할 수 있도록 다양한 스토리지 클래스 제공.

아마존 S3 스탠다드 IA : IA는 Infrequent Access 의 약자로 다른 클래스에 비해 상대적으로 접근 빈도가 낮은 데이터를 위한 스토리지 클래스.

아마존 S3 One Zone-IA : 최근 추가된 One Zone-IA는 평소에는 낮은 속도로 접근하지만, 때에 따라서는 매우 신속하게 접근 할 수 있는 스토리지 클래스.

아마존 글레이셔 : 데이터 아카이브, 즉 데이터 장기보관을 목적으로 활용.

아마존 S3에서의 객체 버전 관리 : 버저닝(Versioning)은 동일한 파일의 다양한 업데이트 상태를 관리하는 방법으로, 적절한 버저닝 기법을 통해 동일 파일의 서로 다른 10가지 버전을 업로드할 수 있음. 아마존 S3에서의 객체 버전 관리 버저닝과 라이프사이클 규칙을 함께 사용해 전 버전의 파일을 좀 더 저렴하거나 중요도가 낮은 스토리지 클래스로 시킬 수 있음. 단, 일단 버저닝을 활성화하면 다시비활성화 할 수 없음.

아마존 S3는 특정 OS 기반의 파일 시스템이 아닌, 웹 기반의 데이터 저장소
S3 서비스는 기본적으로 "한 번 기록하고, 여러 번 읽는" 서비스이므로, S3 의 아 키텍처는 전통적인 파일 시스템 또는 SAN 아키텍처와는 차이가 있음.

아마존 S3에서 정적 웹사이트 호스팅하기