

분산 스토리지에 대한 완벽 가이드



기존 스토리지의 비용 효율적인 대안

목차:

- 1 분산 스토리지에 대한 완벽 가이드를
읽어야 하는 이유
 - 3 1부: 분산 스토리지의 정의
 - 4 2부: 분산 스토리지의 원리
 - 15 3부: 분산 스토리지의 이점
 - 16 4부: 분산 스토리지의 이용 사례
 - 19 5부: 결론
-

분산 스토리지에 대한 완벽 가이드를 읽어야 하는 이유

기존 스토리지 시스템이 페타바이트 규모를 처리할 수 있습니까?

*분산 스토리지에 대한 완벽 가이드*는 대량의 비정형 데이터와 관련된 문제 즉, 기존의 스토리지 기술로 처리할 수 없는 문제를 해결하기 위한 새로운 방식에 대해 설명합니다.

비정형 데이터의 급증

IDC에 따르면 2013년에 생성되고 복제된 디지털 정보의 총량이 4.4제타바이트(제타바이트 = 1,000엑사바이트)를 초과했습니다. 디지털 유니버스의 크기는 2년마다 2배 이상 증가하고 있으며, 2020년이면 44제타바이트에 가깝게 증가할 것으로 예상됩니다.¹

이러한 데이터는 대부분 개인에 의해 생성되지만, IDC의 추정에 따르면 디지털 유니버스의 정보 가운데 85%는 수명 주기의 어느 시점에는 기업이 관할하게 됩니다.² 이는 수요를 충족하기 위한 정보 기술 시스템과 데이터 스토리지 시스템을 구성, 제공, 유지 관리할 책임이 기업에 있음을 의미합니다.



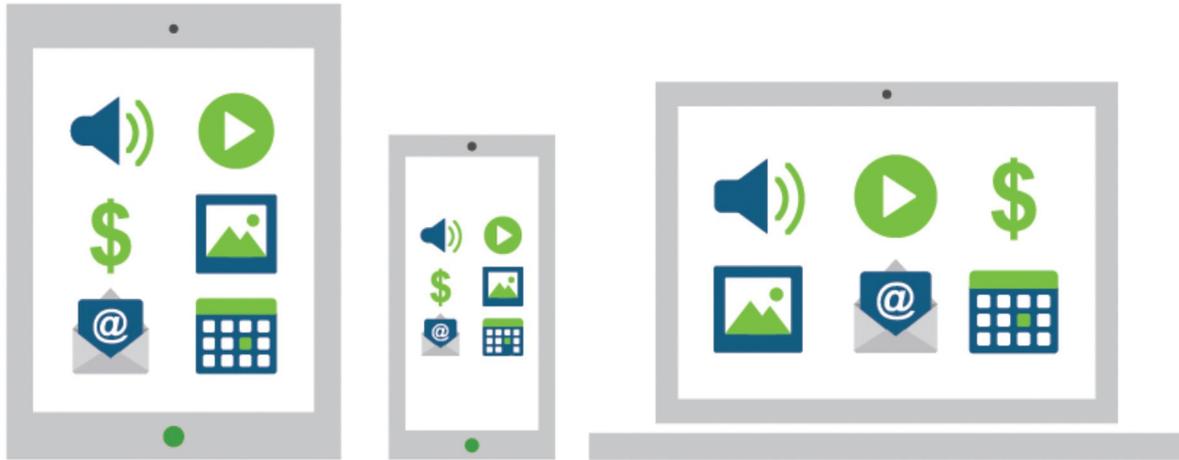


그림 1: 데이터 증가를 촉진하는 비정형 디지털 정보의 유형 및 소스

기존 스토리지 방식은 부적합

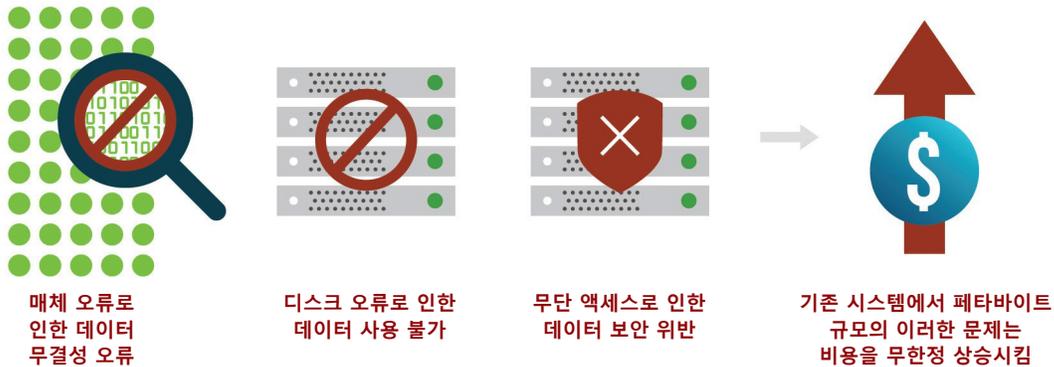


그림 2: 하드웨어 오류 및 무단 액세스로부터 데이터를 보호하고 유지하기 위한 패리티, RAID 및 복제와 같은 기존 스토리지 시스템 방식은 페타바이트 규모에 부적합

여러 가지 기술 발전이 이러한 데이터 증가 문제에 어느 정도 도움을 주고 있습니다. 컴퓨팅이 점점 더 빨라지고 저렴해지고 있으며, 가상화가 효율성과 이용률을 향상시키고 있습니다. 스토리지 장치는 용량이 증가하는 동시에 가격이 낮아지고 있으며(더 낮은 가격으로 장치당 더 많은 비트 수를 이용 가능), 최근 솔리드 스테이트 기술이 등장하여 속도가 더 빨라졌습니다(다만 현재는 모든 워크로드에 적합한 가격점이 아님). 또한 클라우드 컴퓨팅과 같은 서비스 체계가 비용을 낮추고 효율성을 높이는 데 도움이 되고 있습니다. 하지만 경우에 따라 기술의 발전, 특히 스토리지 장치의 용량 확대가 디지털 정보를 보호하고 보존하는 기존 방법에 제약이 되기도 합니다.

RAID 같은 기존의 스토리지 보호 기술은 페타바이트 규모 이상의 디지털 정보를 데이터 손실로부터 보호하는 데 전혀 적합하지 않습니다. 기존 스토리지 아키텍처는 페타바이트 규모로 확장하도록 설계되지 않았습니. 이러한 아키텍처는 안전성과 신뢰성은 더 낮으면서 가격은 더 높습니다.

기존 스토리지 시스템이 페타바이트 규모에 도달할 경우에 다음과 같은 문제에 직면하게 됩니다.

- 시스템 크기가 하드 드라이브의 비트 오류율보다 100억 배 큰 경우 데이터 무결성 문제 발생
- 수백 개의 드라이브에서 매일 오류가 발생하고 재작성하는 데 1주일 이 필요한 경우 데이터 가용성 문제 발생
- 수백만 개의 장치 및 여러 개의 복사본이 여러 위치에 있는 경우 데이터 보안 문제 발생

많은 양의 비정형 데이터를 저장해야 하는 기업은 기존 스토리지 솔루션 이외의 옵션을 찾아보고 새로운 방식을 검토해야 합니다. 이 가이드를 통해 분산 스토리지의 원리와 고유한 이점이 어떻게 다른 기업들로 하여금 높은 수준의 확장성, 가용성 및 보안성을 확보하고 스토리지 비용을 관리하도록 지원했는지 이해할 수 있습니다.

1부 분산 스토리지의 정의

분산 스토리지의 정의

IBM® Cloud Object Storage(COS) System은 대량의 비정형 데이터를 저장하고 보안, 가용성 및 신뢰성을 유지하는 데 도움이 되는 혁신적인 방식을 사용합니다. IBM COS 스토리지 기술은 정보 분산 알고리즘(IDA, Information Dispersion Algorithm)을 사용하여 데이터를 인지 불가능한 "슬라이스(slices)"로 나누며, 이 슬라이스는 네트워크 연결을 통해 로컬 또는 전 세계의 스토리지 노드에 분산됩니다. 이렇게 분산된 스토리지 어플라이언스의 집합이 IBM COS System을 구성합니다. IBM COS 분산 스토리지 기술을 사용하면 데이터의 전송과 저장이 본질적으로 비공개로 안전하게 이루어집니다. 단일 스토리지 노드에 데이터의 완전한 복사본이 존재하지 않으며, 노드의 부분 집합만 사용할 수 있어도 네트워크의 데이터를 완벽하게 검색할 수 있습니다.

배경

분산 스토리지는 인터넷이 데이터 네트워킹에 사용하는 방법을 데이터 스토리지에 적용하여, 기업이 방대한 콘텐츠(비디오, 오디오, 사진, 텍스트)를 안전하고 확실하게 저장할 수 있게 지원하도록 설계되었습니다.

인터넷이 이전의 회선 교환 네트워크에 사용된 전화 통신 프로토콜과 비교하여 향상된 패킷 전환을 기반으로 하는 오픈 프로토콜(TCP/IP)을 사용한 것과 유사하게, 분산 스토리지는 정보 분산 알고리즘(IDA, Information Dispersion Algorithms)이라고 하는 데이터 스토리지를 위한 기술을 상업적 수준으로 구현한 것입니다.

IDA 기술은 데이터를 여러 개의 슬라이스로 변환하는데, 이때 슬라이스의 부분 집합만으로 원본 데이터를 재생성할 수 있는 방식이 사용됩니다. 슬라이스는 패킷과 유사하지만, 데이터

스토리지의 경우 여러 개의 스토리지 어플라이언스에 저장됩니다. 슬라이스는 삭제 코딩(erasure coding), 암호화 및 정교한 분산 알고리즘의 조합을 사용하여 생성됩니다.

분산 스토리지 시스템은 데스크톱 생산성 애플리케이션에서 생성되는 작은 크기의 문서와, 일반적으로 더 큰 크기의 서버 로그 파일을 비롯하여 다양한 유형 및 크기의 디지털 미디어와 같은 비정형 데이터를 저장하는 데 아주 적합합니다. 현재 사용가능한 업계 표준 하드웨어, 소프트웨어 및 네트워킹 기술은 슬라이싱 및 분산과 관련된 처리의 오버헤드 때문에 트랜잭션 기반 데이터베이스처럼 지연에 민감하고, 고도의 IOPS 워크로드에서 구조화된 데이터를 분산하는 데 비용 효율적이지 않습니다.

정보 분산이란?

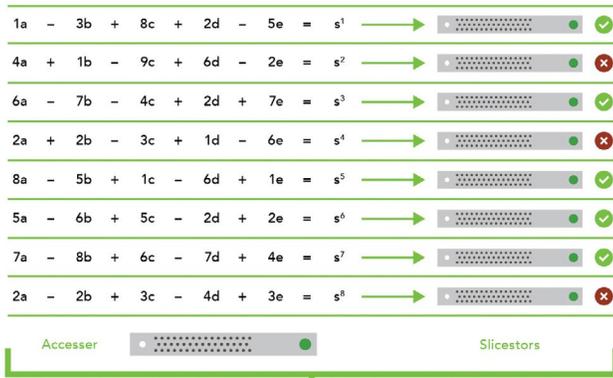
IBM COS System의 기반은 정보 분산(Information Dispersal)이라는 기술입니다. 정보 분산은 데이터의 전송 및 저장을 위해 중복을 생성하는 수단으로 삭제 코드를 사용하는 방식입니다.

삭제 코드는 k 기호가 포함된 메시지를 n 기호가 포함된 더 긴 메시지로 변환하여, n 기호의 부분 집합(k 기호)에서 원본 메시지를 복구할 수 있도록 하는 방식의 정방향 오류 수정(FEC, Forward Error Correction) 코드입니다.

간단히 설명하면, 삭제 코드는 사용자가 "코딩된 데이터"의 부분 집합만 있으면 "원본 데이터"를 재생성할 수 있는 "추가 데이터"를 "원본 데이터"에 삽입하는 고급 결정론 수학(Deterministic Math)을 사용합니다.

IDA는 모든 정방향 오류 수정(FEC, Forward Error Correction) 코드에서 생성될 수 있습니다. IDA의 추가 단계는 코딩된 데이터를 여러 세그먼트로 분할하는 것으로, 그런 다음 이를 여러 장치 또는 미디어에 저장하여 높은 수준의 오류 독립성을 달성할 수 있습니다. 예를 들어, 컴퓨터의 파일에만 FEC를 사용하면 하드 드라이브가 실패할 경우 도움이 될 가능성이 적지만, IDA를 사용하여 정보 슬라이스를 여러 시스템에 분산시키면 데이터를 재조합하는 능력을 손상시키지 않으면서 여러 개의 오류를 허용할 수 있습니다.

정보 분산의 기반이 되는 수학



K개 변수를 알아내려면
임의의 k개 결과를 알아야 하므로
(n-k)개의 결과를 놓쳐도 문제 없음
전송되는 디스크의 데이터 오버헤드는 (n/k)

그림 3. 정보 분산 및 검색 알고리즘의 기반이 되는 수학 방정식

이러한 IDA의 원리는 아주 간단합니다. 고등학교에서 대수학 시간에 배운 내용을 떠올려 보십시오. 5개의 변수를 사용하는 방정식의 집합이 있을 때, 이러한 변수를 사용하는 서로 다른 방정식이 5개 이상이면 변수의 해를 구할 수 있습니다.

위의 그림 3과 같이, 5개의 변수(a부터 e까지)와 이러한 변수를 사용하는 8개의 서로 다른 방정식이 존재하고 각 식이 서로 다른 결과를 산출합니다. 정보 분산의 원리를 이해하기 위해 5개의 변수가 바이트라고 생각할 수 있습니다. 8개의 방정식을 통해 8개의 결과를 계산할 수 있으며, 각 결과가 하나의 바이트에 해당합니다. 원래의 5바이트를 알아내기 위해 결과에 해당하는 8바이트 중에서 임의의 5개를 사용할 수 있습니다.

이것이 바로 정보 분산이 k 값을 지원하는 방식이고, 여기서 n-k는 변수의 개수, n은 방정식의 개수입니다.

2부

분산 스토리지의 원리

분산 스토리지의 원리: 단계별

기본적으로 IBM COS System은 데이터의 슬라이싱, 분산 및 검색을 위해 3단계를 사용합니다.

1단계

데이터가 IDA를 사용하여 가상화, 변환, 슬라이싱 및 분산됩니다. 그림 4의 예에서는 데이터가 12개의 슬라이스로 분할되었습니다. 따라서 시스템의 "너비"(n)는 12입니다.

2단계

슬라이스가 별도의 디스크, 스토리지 노드 그리고/또는 지리적 위치에 분산됩니다. 이 예에서는 슬라이스가 3개의 다른 사이트에 분산되었습니다.

3단계

슬라이스의 부분 집합에서 데이터가 검색됩니다. 이 예에서 데이터를 검색하기 위해 필요한 슬라이스의 개수는 7입니다. 따라서 시스템의 "임계값"(k)이 7입니다.

너비가 12이고 임계값이 7인 경우, 이러한 예를 가리켜 "12 중 7"(n 중 k) 구성이라고 합니다.

시스템의 구성은 필요한 신뢰성 수준에 따라 결정됩니다. "12 중 7" 구성에서는 5개의 슬라이스가 손실되거나 사용할 수 없더라도 임계값인 7개의 슬라이스가 충족되었으므로 여전히 데이터를 검색할 수 있습니다. "8 중 5" 구성에서는 3개의 슬라이스만 손실이 허용되므로 신뢰도가 더 낮습니다. 반면 "32 중 20" 구성에서는 12개의 슬라이스가 손실이 허용되므로 신뢰도가 더 높습니다.

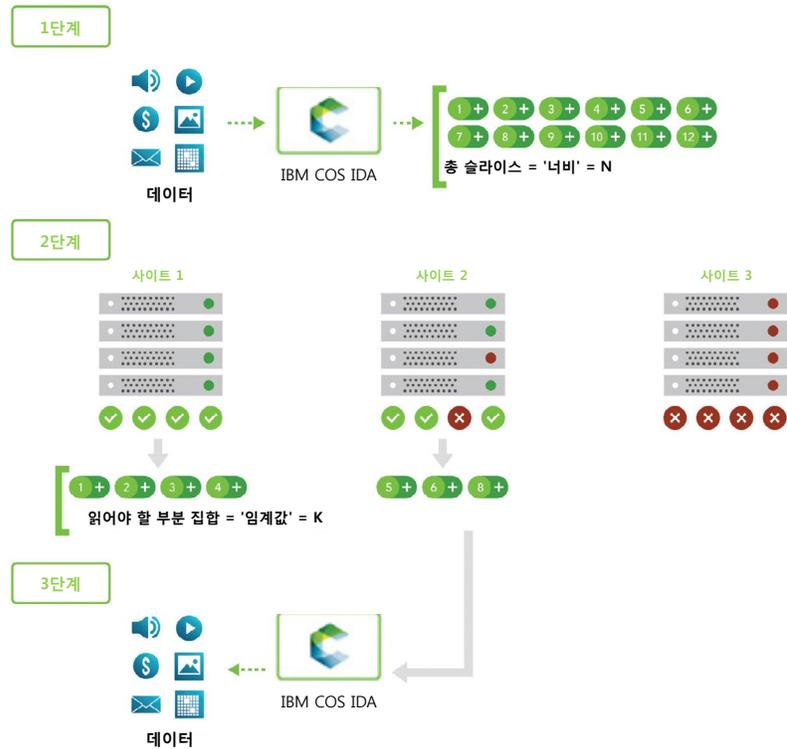


그림 4. IBM COS System에서의 단계별 데이터 슬라이싱, 분산 및 검색

다중 사이트 오류 예제

분산 스토리지를 사용할 경우, 데이터를 검색하려면 슬라이스의 부분 집합만 있어도 됩니다. 따라서 분산 스토리지 시스템은 단일 사이트 내에서 그리고 여러 사이트에 걸쳐서 어플라이언스 오류를 허용할 수 있습니다.

1단계

IDA를 사용하여 데이터가 가상화, 변환, 슬라이싱 및 분산됩니다. 이 예에서 시스템의 "너비"(n)는 12입니다.

2단계

슬라이스가 별도의 디스크, 스토리지 노드 및/또는 지리적 위치에 분산됩니다. 그림 5의 예에서는 슬라이스가 지리적으로 떨어진 4개의 사이트에 분산되어 있습니다.

3단계

데이터가 슬라이스의 부분 집합에서 검색됩니다. 이 예에서 데이터를 검색하기 위해 필요한 슬라이스의 개수는 7개입니다. 따라서 3개의 사이트에서 모두 오류가 발생하더라도, 사용 가능한 슬라이스가 7개라는 "임계값"에 도달했으므로 여전히 데이터를 검색할 수 있습니다.

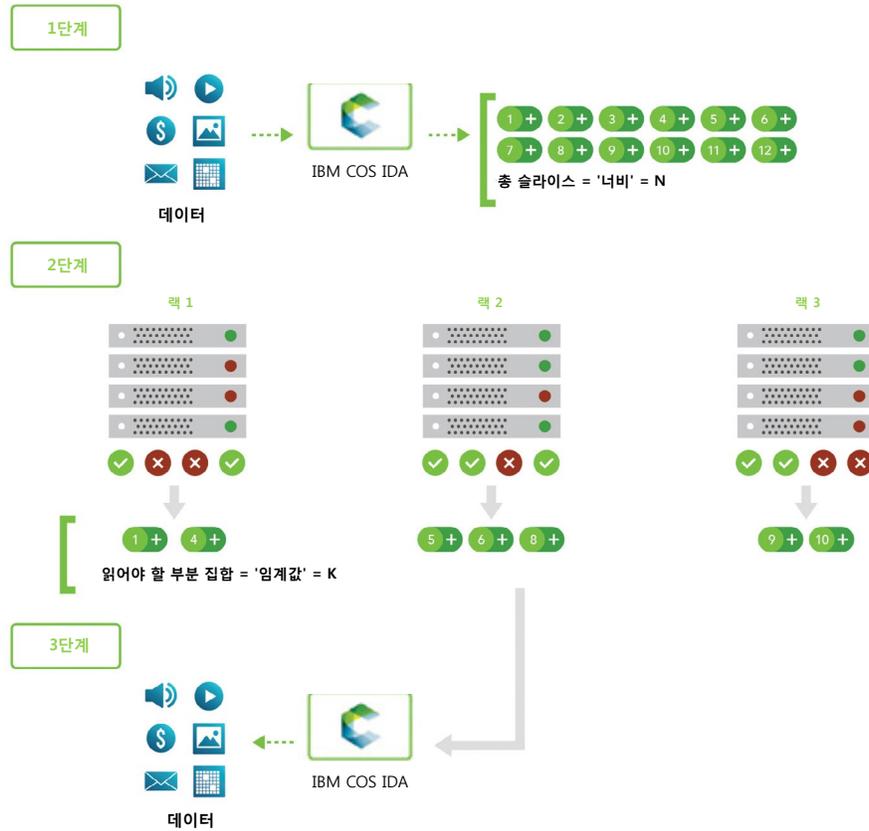


그림 5. 단일 사이트 내에서 또는 여러 사이트에서 오류를 허용하기 위한 IBM COS System의 데이터 슬라이싱, 분산 및 검색

단일 사이트/다중 장치 오류 예제

단일 사이트 내의 여러 장치에서 오류를 허용하는 기능을 포함하여 분산 스토리지 시스템을 단일 사이트에 배포할 수도 있습니다.

1단계

데이터가 IDA를 사용하여 가상화, 변환, 슬라이싱 및 분산됩니다. 그림 6의 예제에서는 시스템의 "너비"(n)가 12입니다.

2단계

슬라이스가 별도의 디스크, 스토리지 노드 및/또는 지리적 위치에 분산됩니다. 이 예제에서는 슬라이스가 단일 사이트 내에서 서로 다른 4개의 랙에 분산되었습니다.

3단계

데이터가 슬라이스의 부분 집합에서 검색됩니다. 이 예제에서 데이터를 검색하는 데 필요한 슬라이스의 개수는 7개입니다. 따라서 각 랙에서 하나 이상의 장치 오류가 발생했다 하더라도, 7개의 슬라이스라는 "임계값"이 충족되었으므로 데이터를 검색할 수 있습니다. 5개의 슬라이스를 사용할 수 없어도 데이터의 모든 비트를 완벽하게 재생성할 수 있습니다.

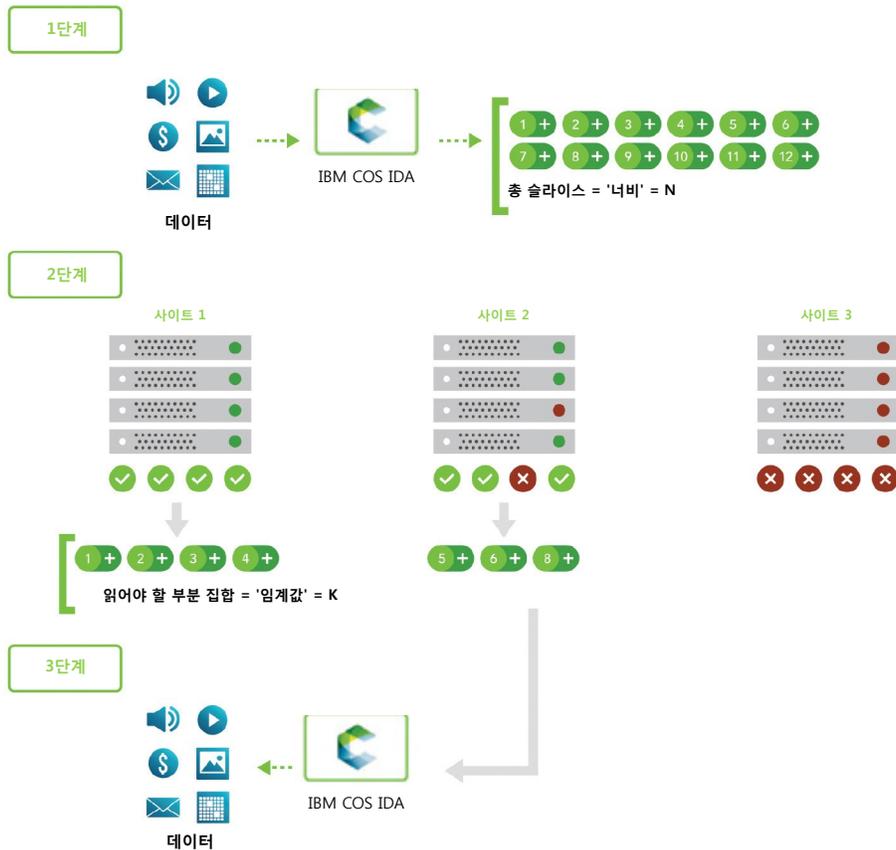


그림 6: 단일 사이트의 여러 오류를 허용하기 위한 IBM COS System의 데이터 슬라이싱, 분산 및 검색

IBM COS System의 구성요소

IBM COS System은 세 가지 소프트웨어 스토리지 시스템 - IBM® Cloud Object Storage Manager, IBM® Cloud Object Storage Accesser® 및 IBM® Cloud Object Storage Slicestor®의 생성을 지원합니다. 이 소프트웨어 구성요소들은 호환 가능한 광범위한 업계 표준 하드웨어 플랫폼에 가상 시스템으로 배포할 수 있고, IBM COS Accesser®의 경우는 Linux OS에서 실행되는 애플리케이션으로 배포할 수 있습니다. 또한, 물리적 배포와 가상 배포를 단일 시스템으로 결합할 수 있습니다(예: IBM COS Manager 및 IBM COS Accesser®의 가상 시스템과 IBM COS Slicestor®의 물리적 서버).

소프트웨어

세 가지 소프트웨어 구성요소는 각각 IBM COS의 일부로서 특정한 기능을 수행합니다.

- IBM COS Manager는 시스템의 상태 및 성능 모니터링, 시스템 구성 및 스토리지 프로비저닝, 오류 관리, 기타 관리 및 운영 기능을 담당합니다.
- IBM COS Accesser®는 데이터 수집 시 데이터의 암호화/인코딩과, IBM COS Slicestor® 노드의 집합에서 이러한 프로세스로 인해 발생한 데이터 슬라이스의 분산을 관리하거나 읽을 때 데이터의 디코딩/복호화를 담당합니다.
- IBM COS Slicestor®는 데이터 슬라이스의 저장을 담당합니다.

IBM COS Manager, IBM COS Accesser® 및 IBM Slicestor® 소프트웨어를 IBM이 인증한 업계 표준 하드웨어 플랫폼에 배포하면 다음과 같은 몇 가지 이점이 있습니다.

- 하드웨어 및 소프트웨어 호환성과 구성이 IBM COS에 의해 사전 정의되고 검증되기 때문에 최초 배포 시 프로덕션에 소요되는 시간이 크게 절감됩니다.
- 하드웨어 구성이 IBM COS의 효율을 극대화하도록 최적화됩니다.
- 하드웨어 상태를 더 낮은 구성요소 수준에서 모니터링하고 관리하므로 시스템 안정성이 향상됩니다.
- 시스템의 하드웨어 및 소프트웨어 구성요소에 모두 정통한 IBM COS 지원 담당자에게 문의할 수 있습니다.

오브젝트 스토리지 기반

IBM COS는 단순 오브젝트 스토리지 방식에 기반하여 수십억 개의 데이터 오브젝트를 단일 플랫폼 네임스페이스에 저장하고 HTTP 기반 프로토콜을 사용하는 REST 인터페이스를 통해 노출합니다.

기존 방식: 파일 기반

기존 스토리지 시스템은 데이터를 계층 파일 시스템으로 구성하고 NFS 및 SMB 같은 NAS 기반 프로토콜을 통해 데이터를 제공합니다. 이러한 파일 시스템 방식은 사용자(사람)가 적은 양의 데이터를 저장하는 데 적합합니다.

향상된 방식: 오브젝트 기반

확장성	페타바이트 이상의 확장성으로 비용, 용량, 액세스 가능성이 개선됩니다.
보안	미션 크리티컬 데이터를 제로 터치 암호화와 내장된 캐리어급 보안으로 보호합니다.
가용성	계획된 또는 계획되지 않은 다운타임 시간에 관계없이 데이터를 항상 사용할 수 있도록 합니다.
효율성	기존 스토리지 운영보다 약 15배 더 효율적인 직관적인 플랫폼으로 관리를 간소화합니다.
경제성	범용 하드웨어에서 실행되는 프리미엄 소프트웨어 기반 솔루션으로 장기간의 총소유비용을 크게 절감합니다.

오브젝트 기반 스토리지의 동적 데이터 처리는 기존 스토리지보다 많은 이점을 제공합니다. 이 중에는 우수한 확장성, 스토리지 효율성 향상, 데이터 마이그레이션 및 이동의 용이성이 있습니다. 또한 오브젝트 기반 스토리지는 기존 스토리지보다 더 많은 메타데이터를 지원하므로 데이터 계층, 보안 및 마이그레이션을 더 쉽게 관리할 수 있습니다.¹

그림 7. 기존 스토리지와 비교한 페타바이트 스토리지 시스템의 오브젝트 기반 스토리지의 이점

파일 시스템 스토리지에서 데이터는 위치와 밀접하게 연관됩니다. 오브젝트 기반 스토리지는 데이터를 스토리지 시스템의 물리적 위치와 디커플링하여 이러한 제한을 극복합니다. 애널리스트 기업인 Forrester Research는 오브젝트 기반 스토리지에 꼭 맞는 비유로 발렛 파킹을 들었습니다.⁴ 발렛 파킹 서비스에 차량을 맡기면 직원에게 필요할 때 차를 찾을 수 있는 주차 티켓을 받습니다. 직원은 차량을 보관하는 동안 주차장이나 차고의 공간을 최적화하기 위해 필요에 따라 차량을 이동할 수 있습니다. 주차 티켓은 특정한 주차 공간이 아니라 차량 자체를 식별합니다. 오브젝트 기반 스토리지에서도 오브젝트 ID는 스토리지 시스템의 특정한 위치가 아니라 특정한 데이터 조각을 식별합니다. 데이터를 필요에 따라 시스템에서 이동할 수 있으며, 오브젝트 ID는 데이터가 어디에 있는 데이터를 검색하기 위해 필요한 "주차 티켓"입니다.

액세스 방법

오브젝트 기반 액세스 방법

분산 스토리지 시스템의 기본 스토리지 풀은 공유할 수 있고 여러 액세스 프로토콜에서 공동으로 액세스할 수 있습니다. 단순 오브젝트 인터페이스는 HTTP/REST API를 사용하여 액세스됩니다. 단순 PUT, GET, DELETE, LIST 명령은 애플리케이션이 디지털 콘텐츠에 액세스하도록 허용하고, 그 결과로 발생한 오브젝트 ID는 바로 애플리케이션 내에 저장됩니다. IBM COS Accesser®는 애플리케이션이 오브젝트 ID를 사용하여 바로 IBM COS Slicestor®와 통신하므로 전용 어플라이언스가 필요하지 않습니다.

스토리지에 대한 REST API 액세스

REST는 월드 와이드 웹(WWW)과 같은 분산 하이퍼미디어 정보 검색 시스템을 위한 소프트웨어 아키텍처의 일종입니다. REST 형태의 아키텍처는 클라이언트와 서버로 구성됩니다. 클라이언트가 서버에 요청을 보내면, 서버는 요청을 처리하여 관련된 응답을 회신합니다. 요청과 응답은 리소스의 다양한 표시를 전송하는 것을 기반으로 이루어집니다.

REST API는 URL(Universal Resource Locator) 검색과 유사한 방식으로 작동합니다. 단, 웹 페이지를 요청하는 대신에 애플리케이션이 오브젝트를 참조합니다.

스토리지에 대한 REST API 액세스는 몇 가지 이점이 있습니다.

- 인터넷 지연을 허용
- "프로그래밍 가능한" 스토리지를 지원
- 대량의 데이터에 대하여 효율적인 글로벌 액세스 제공

파일 기반 액세스 방법

분산 스토리지는 써드파티 게이트웨이 어플라이언스와의 통합을 통해 기존의 NAS 프로토콜(SMB/CIFS 및 NFS)도 지원합니다. 사용자와 스토리지 관리자는 표준 파일 프로토콜을 통해 데이터 자산을 쉽게 전송, 액세스 및 보존할 수 있습니다.

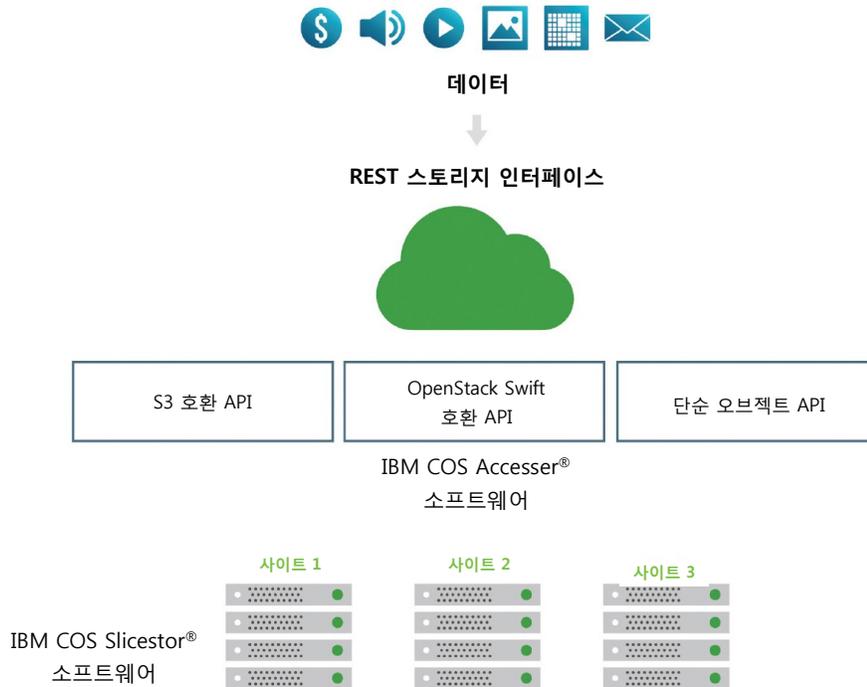


그림 8: 애플리케이션은 업계 표준 REST 스토리지 인터페이스를 사용하는 IBM COS Slicestor®에서 바로 오브젝트 기반 스토리지에 액세스할 수 있음

보안 기능

데이터 보안

IBM® COS SecureSlice™는 IBM COS System에 저장된 데이터의 기밀성, 무결성 및 가용성을 위해 사용되는 기술입니다. SecureSlice는 정보 분산 알고리즘과 AONT(All-or-Nothing Transform)라는 두 가지 알고리즘을 결합했습니다. AONT는 모든 정보를 알아야만 정보를 해독할 수 있는 암호화 모드입니다. 그림 9와 10의 다이어그램은 SecureSlice를 사용하는 기본적인 읽기 및 쓰기 작업을 보여줍니다.

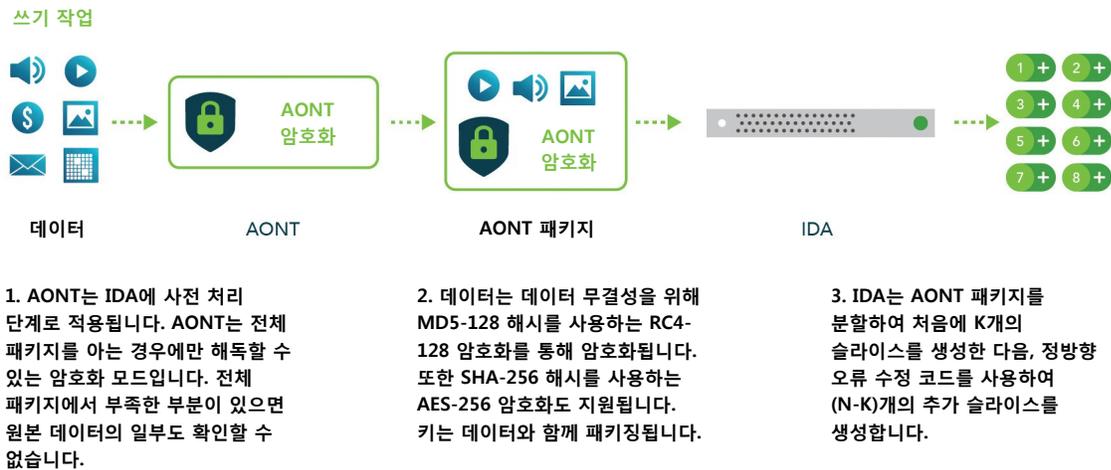


그림 9. IBM COS System의 기본적인 쓰기 작업



그림 10. IBM COS System의 기본적인 읽기 작업

AONT 인코딩의 내부

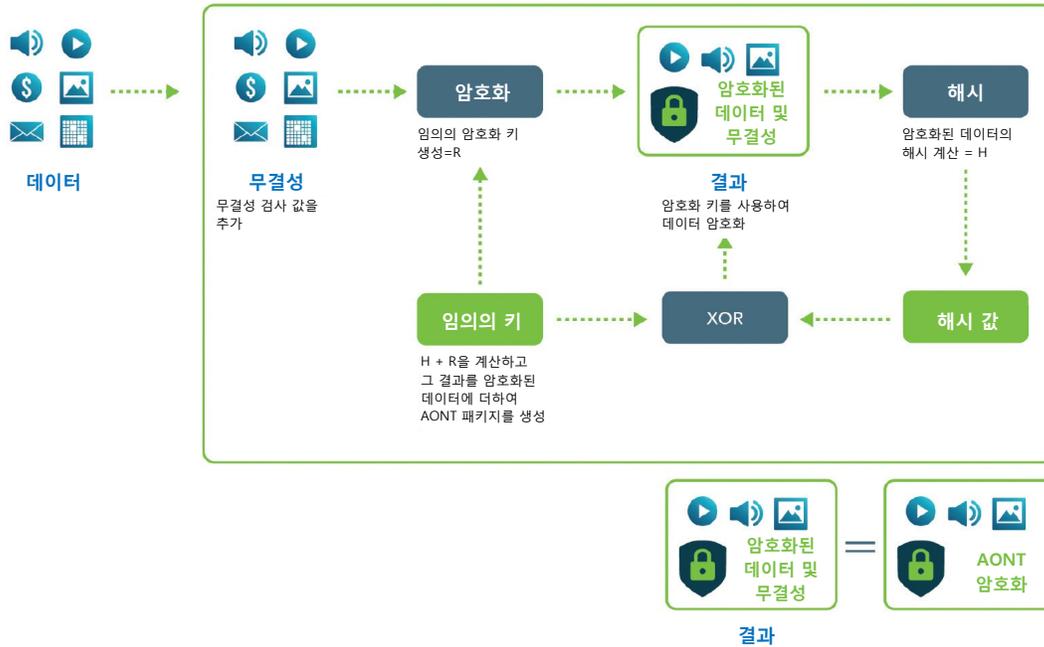


그림 11: IBM COS System에서 AONT(All-Or-Nothing Transform) 작업의 내부

데이터 세그먼트가 분산 스토리지 시스템에 저장되면 무결성 검사 값이 제일 먼저 데이터에 추가됩니다. 무결성 검사 값은 길이가 충분하다면, 잘 알려진 임의의 상수 값일 수 있습니다. 이 값은 손상이 발생하지 않았는지 확인하기 위해 복호화 후에 검사됩니다.

데이터 세그먼트의 재조합시 손상된 슬라이스가 있으면 무결성 검사도 손상되었을 가능성이 매우 높습니다. 분산 스토리지 시스템은 이러한 손상된 값을 인식하여 사용자에게 잘못된 값이 전달되지 않도록 합니다.

특정 슬라이스에서 무결성 검사 값이 손상된 경우, 분산 스토리지 시스템에서는 완전한 데이터 세그먼트를 검색하기 위해 슬라이스의 유효한 조합을 찾으려고 시도합니다.

네트워크 보안

어플라이언스에서 시작하거나 종료되는 모든 네트워크 트래픽 흐름은 TLS, SSL 또는 SNMPv3(AES)을 사용하여 암호화됩니다. 스토리지 노드는 복잡한 방화벽 또는 VPN 설정 없이 어디든 가상으로 배치될 수 있습니다.

IBM COS Accesser® Exposes HTTP REST API

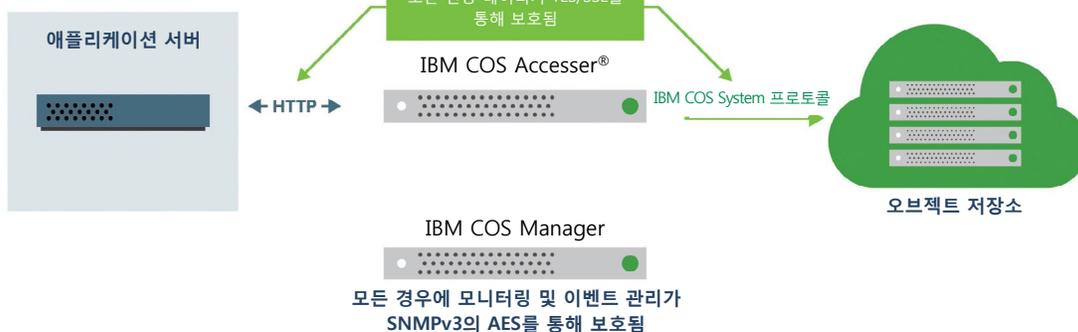
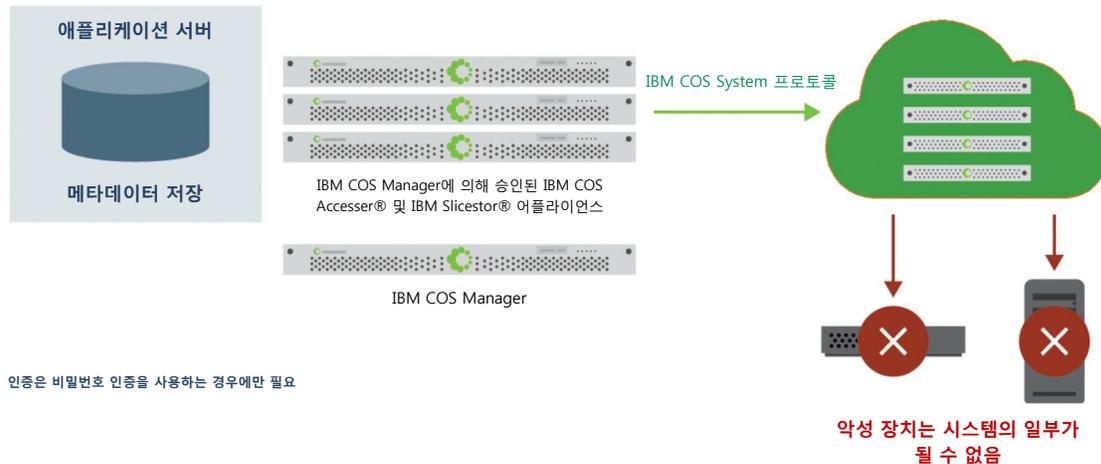


그림 12: IBM COS System의 네트워크 보안

장치 등록 - 인증 기관(CA) 신뢰

단순히 장치가 TLS 또는 SSL을 사용한다고 하는 것으로는 부족합니다. 이러한 프로토콜은 연결이 인증되지 않는 한 "중간자"(man in the middle) 공격을 예방하지 않습니다. TLS/SSL을 이용하는 인증을 위해서는 디지털 인증서 사용이 필요하고, 이러한 인증서는 스토리지 네트워크에서 유효한 노드에 속하는 것으로 확인될 수 있어야 합니다. 이렇게 하기 위해 노드가 스토리지 네트워크로 승인될 때 노드에 서명된 디지털 인증서가 제공됩니다.

이러한 승인을 위해서는 관리 인터페이스에 로그인하여 요청을 확인하고 승인할 관리자가 필요합니다. 관리자는 요청을 생성하는 장치의 IP 주소, MAC 주소 및 지문을 볼 수 있고, 장치를 시스템에 허용하기 전에 각각이 올바른지 확인할 수 있습니다. 승인이 이루어지면 스토리지 네트워크의 인증 기관(CA, Certificate Authority)에서 서명한 인증서가 노드에 부여됩니다. 스토리지 네트워크의 모든 장치는 이 인증 기관을 신뢰하며, 따라서 이 CA에서 서명한 유효한 인증서를 소유한 노드도 신뢰합니다. 어플라이언스는 미래의 어느 시점에 폐기되거나 손상될 수 있습니다. 이 시점에 IBM COS Manager는 장치의 인증서를 인증서 폐기 목록(CRL, Certificate Revocation List)에 추가하여 폐기할 수 있습니다. 이 목록은 시스템의 모든 노드에 의해 주기적으로 등록됩니다.



- 인증 단계:**
1. 장치가 IBM COS Manager 인증 기관(CA)에 등록됨
 2. 인증서가 장치로 다시 전송됨
 3. IBM COS Manager CA에 대하여 확인된 각 장치에 대하여 별도의 보안 연결이 생성됨
 4. 사용자 이름/비밀번호 또는 PKI를 사용하여 클라이언트를 인증
 5. 비밀번호를 인증 서비스를 통해 검사(*)

그림 13. IBM COS System의 장치 등록 및 인증

가용성 기능

분산 스토리지 시스템의 가용성 기능은 지속적으로 오류를 감지하고 수정하여 비트 단위까지 완벽한 데이터 가용성을 확보하도록 지원합니다.

모든 슬라이스 및 파일의 무결성 검사

분산 스토리지 시스템은 사전 대응적으로 오류를 스캔하여 수정하고, 데이터 슬라이스의 무결성을 스캔하고, 손상된

슬라이스를 재작성하고, 제공 전에 슬라이스 무결성과 파일 데이터 무결성을 점검하는 지능형 백그라운드 프로세스를 통해 데이터 무결성을 검사합니다. 이는 비트 오류의 사전 대응식 수정과 정상적인 읽기/쓰기 작업 중에 발생할 수 있는 잠재적인 소프트 오류의 수정을 통해 비트 단위까지 완벽한 데이터 제공을 보장하는 데 도움이 됩니다. 또한 데이터를 승인 없이 수정할 수 없도록 하고 악성 위협을 감지하는 데도 도움을 줍니다.



그림 14. 비트 단위까지 완벽한 데이터 가용성을 위한 IBM COS System의 데이터 무결성 검사

지속적인 오류 수정

슬라이스가 손상되었음이 확인되면(무결성 검사 값이 유효하지 않으면) IBM COS Slicestor®에서 분산 리빌더 기술을 시작하여 해당 슬라이스를 유효한 슬라이스로 대체합니다. 지속적으로 오류를 수정하면 오류를 감지하기 위해 데이터를 읽을 때까지 기다리지 않으므로 시스템 가용성이 향상됩니다.

이는 정보가 자주 읽혀지지 않는 장기 아카이브 및 대량 디지털 저장소에서 매우 중요한 기능입니다. 분산 리빌더 모델은 예측 가능성을 제공하며, 리빌더가 적절한 속도로 "항상 실행"되므로 I/O 성능의 예측 가능성이 크게 향상될 뿐만 아니라 리빌더가 스토리지와 함께 성장하므로 확장성도 높아집니다.

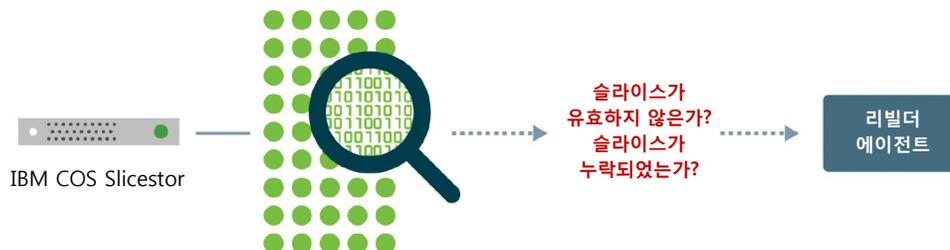


그림 15. IBM COS System의 IBM COS Slicestor®에 의한 지속적인 오류 수정

성능 최적화 기능

분산 스토리지는 IBM® COS SmartWrite™ 및 IBM® COS SmartRead™ 기술을 사용하여 슬라이스의 읽기 및 쓰기를 최적화하고, 따라서 처리량 및 효율성을 향상시킵니다.

n개 중의 k개 쓰기의 성능 향상

SmartWrite는 슬라이스의 전체 너비를 쓸 수 없는 경우(예를 들어 노드 또는 네트워크에 오류 상태가 존재하는 경우)에도 성공적인 쓰기 작업을 지원합니다. SmartWrite는 최적화된 모든 슬라이스의 쓰기 작업을 시도합니다. 슬라이드의 필수 쓰기 임계값에 도달하면 SmartWrite는 쓰기가 성공한 것으로 간주합니다. 나머지 슬라이스는 계속해서 비동기식으로 쓰기를 시도합니다. 슬라이스 쓰기 작업의 제한시간이 초과되면 해당 슬라이스가 감지되고 재작성됩니다.

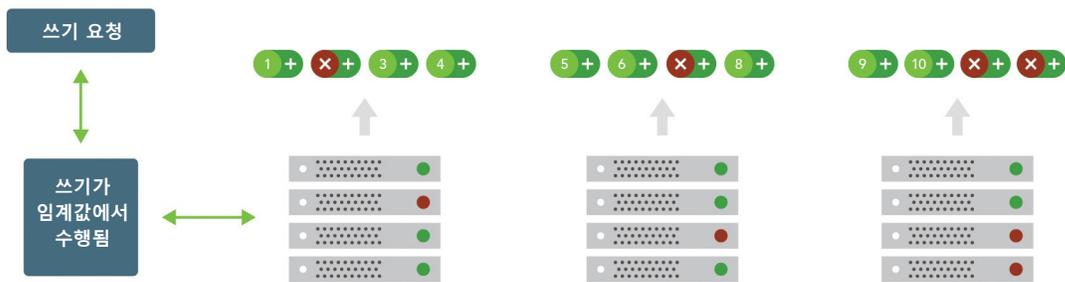


그림 16. IBM COS System의 SmartWrite 기술로 향상된 쓰기 성능

n개 중의 k개 읽기의 향상된 성능

SmartRead는 데이터를 보다 효율적으로 검색하기 위한 최적의 네트워크 경로 및 스토리지 노드를 예측합니다. 데이터는 세그먼트로 재조합되고, 최적의 제공 경로를 확인하기 위해 각 세그먼트마다 수천 개(또는 수백만 개)의 슬라이스의 조합이 검사됩니다. SmartRead는 온디맨드 성능을 기준으로 스토리지 노드의 순위를 지정하고 데이터를 재생성하기 위한 슬라이스의 최적의 조합을 요청합니다. 슬라이스 요청이 제대로 수행되지 않으면 SmartRead가 다른 노드에서 슬라이스를 요청합니다.



그림 17. IBM COS System의 SmartRead 기술로 향상된 읽기 성능

3부

분산 스토리지의 이점

확장성의 이점

분산 스토리지는 관리 오버헤드를 크게 절감하여 탁월한 확장성을 제공합니다. 시스템을 테라바이트에서 페타바이트, 엑사바이트 수준으로 쉽게 확장할 수 있습니다.

확장성의 여러 가지 동인

데이터를 페타바이트, 엑사바이트 이상의 수준으로 저장하고 관리하려면 확장 가능한 아키텍처가 필요합니다. 중앙 집중식 서버 없이 분산 스토리지 시스템의 용량 및 성능을 독립적으로 확장할 수 있습니다. 분산 스토리지 시스템의 오브젝트 스토리지 하부구조는 무제한적인 규모의 스토리지를 위해 필요한 데이터 이동성, 확장성 및 스토리지 효율성을 지원합니다.

분산 스토리지는 모든 개별 스토리지 노드를 가상화하여 단일 관리 지점을 제공하는 주소 지정 가능한 단일 글로벌 네임스페이스를 제공합니다. 글로벌 네임스페이스 방식을 사용하는 경우의 또 다른 이점은 더 큰 규모의 작업 풀을 위한 스토리지 풀을 추가로 개방하고, 데이터를 투명하게 마이그레이션하고, 단일 환경에서 마운트 지점 및/또는 공유의 수를 줄일 수 있다는 점입니다.

각각 독립적으로 스토리지 노드에 액세스하는 애플리케이션 서버 또는 장치 내에서 분산 스토리지 프로토콜을 사용할 수 있습니다. 이는 대량의 병렬 쓰기 및 콘텐츠 분산을 달성하는 데 도움이 됩니다. 그리고 게이트웨이의 병목 지점을 회피하여 분산 환경의 성능을 향상시키는 데 도움이 됩니다.

가용성의 이점

분산 스토리지는 수백만 개의 물리적 비트 오류가 발생하거나 여러 드라이브, 서버, 컨테이너 또는 위치가 변경 혹은 대체되어도 거의 100%의 데이터 무결성을 유지합니다. 입력되는 정보가 출력되는 정보와 동일하고, 완벽하게 인증되며, 정보를 사실상 언제 어디서나 액세스할 수 있습니다. 동시 오류를 허용할 수 있는 아키텍처를 통해 데이터를 항상 사용할 수 있습니다.

구성 가능한 가용성과 다운타임 없는 업그레이드

분산 기술은 탁월한 데이터 보호 및 가용성을 제공합니다. 분산에서는 다른 구현의 단점을 극복하기 위해 데이터를 복제하지 않습니다. 다른 스토리지 솔루션과 달리 분산 솔루션은, 단일 하드 드라이브에 대해 몇 시간 심지어 며칠이 소요될 수도 있는 재작성 프로세스 중에 경험할 수 있는 심각한 데이터 손실 위험이 없으므로 훨씬 우수합니다. 또한, 다른 여러 스토리지 솔루션에서 사용되는 RAID 5(N개 중의 1개) 및 RAID 6(n개 중의 2개)과 비교하여 훨씬 높은 수준의 오류(n개 중의 k개)를 허용하도록 구성할 수 있습니다. IBM COS는 IDA를 사용하고 그 결과로 생성되는 슬라이스를 단일 사이트 또는 지리적으로 분산된 독립적인 하드웨어에 저장하여, 복제 없이 신뢰성과 가용성을 달성하도록 지원합니다.

IBM COS System은 기업이 전체 사이트 오류를 허용하고 값비싼 복사본 없이 데이터에 원활하게 액세스할 수 있도록 설계되었습니다. 따라서 일상적인 유지보수를 위해 데이터 센터를 오프라인으로 설정하더라도 가용성은 변하지 않습니다. 분산된 스토리지를 사용하여 제로 다운타임 업그레이드가 가능합니다. 즉 롤링 업그레이드를 통해 프로세스 전반에서 데이터를 액세스할 수 있는 상태로 시스템이 유지됩니다. 별도로 예정된 유지보수 시간이 필요하지 않습니다.

보안의 이점

분산은 여러 장치, 서버, 컨테이너 또는 위치가 손상되는 경우에도 데이터 기밀성을 보장하는 데 도움이 됩니다. 전송 데이터와 저장 데이터가 암호화되어 완전히 인식 불가능하고 근본적으로 보호되는 상태가 되어 보안 침해 가능성을 제거합니다.

저장 데이터 및 전송 데이터의 뛰어난 보안

분산 스토리지 기술은 암호화, AONT(All-or-nothing Transformation), 무결성 검사 값 및 IDA를 사용하여 데이터를 근본적으로 안전한 슬라이스로 효과적으로 분할합니다. 각 슬라이스가 암호화되지만 외부의 키 관리가 필요하지 않습니다. 복호화를 위한 슬라이스의 "임계값"을 사용할 수 있는 한, 모든 데이터는 수학적으로 안전합니다. 그 결과 데이터를 나타내는 부분을 포함하지 않고 데이터를 비트 단위까지 완벽하게 재생성하기 위해 임계값 개수가 필요한 슬라이스가 생성됩니다.

이러한 슬라이스는 독립적인 하드웨어에 저장되므로, 어떠한 스토리지 볼륨에도 데이터의 완벽한 복사본이 존재하지 않습니다. 슬라이스를 포함하는 개별 서버는 물리적으로 다른 여러 위치에서 구해야 하는 임계값 개수만큼의 슬라이스를 확보하지 않는 한 소용이 없습니다. 이는 보안 침해 가능성이 크게 감소하고 경우에 따라 완전히 제거됨을 의미합니다. 뿐만 아니라 슬라이스는 네트워크를 통과하기 전에 생성되므로, 전송 과정의 보안 침해로부터 보호됩니다. 그 결과 네트워크로 전송되는 데이터와 저장된 데이터 모두에 대하여 뛰어난 데이터 보안이 구현됩니다.

경제적 이점

IBM COS System은 비용이 많이 드는 복제 및 관련 증분 비용을 상당히 절감하고 많은 경우 아예 제거하여 페타바이트 수준 및 그 이상의 스토리지 시스템의 총소유비용을 크게 감소시킵니다. 하드웨어, 전기 요금, 설치 공간, 지원 및 관리 비용도 절감됩니다.

복사본 없이 비용 절감

분산 기술은 비용이 높은 복제의 필요성을 없애줍니다. IBM COS System은 최대 4개의 복제된 데이터 복사본과 동일한 가용성을 제공하며 스토리지 요구사항을 기존의 방식과 비교하여 최대 5배까지 감소시킵니다.

그림 18과 같이, IBM COS는 일반적으로 사용되는 스토리지 파일 서비스와 비교하여 요구사항이 훨씬 적습니다.

- 데이터 분산을 위한 원시 스토리지가 3분의 1 미만
- 전력 및 냉각 비용이 거의 3분의 1
- 설치 공간이 거의 4분의 1
- 하드웨어, 소프트웨어 및 테이프 사본의 추가 비용이 없음

4부 분산 스토리지 솔루션

이용 사례

IBM COS System은 대규모의 비정형 데이터를 높은 가용성을 유지하면서 안전하게 저장해야 하고, 지연 시간이 중요한 고려 사항이 아닌 기업에게 적합한 솔루션입니다.

비정형 데이터 제공

새로운 디지털 정보의 전례 없는 증가와 함께, 기업이 무제한적인 데이터를 저장하고 배포할 수 있게 된 이용 사례가 등장하고 있습니다. 분산된 비중앙식 스토리지 아키텍처와 오브젝트 스토리지 인터페이스는 기업이 전례없이 전 세계의 사용자에게 데이터를 제공할 수 있도록 지원합니다. 이러한 이용 사례로는 콘텐츠 저장소, SaaS(Storage-as-a-service), 기업내 협업, 백업 및 아카이브 등이 있습니다.

콘텐츠 저장소 스토리지

IBM COS System은 비즈니스 크리티컬 데이터를 위해 가장 안정성이 높고 확장성이 뛰어난 플랫폼 중 하나를 제공합니다.

중요한 콘텐츠를 효과적으로 저장 및 보호

사용자들이 전 세계의 여러 위치에서 콘텐츠에 액세스하므로, 대체할 수 없는 원본을 대규모로 보호하는 일이 비즈니스 우선 순위가 되었습니다. IBM COS의 콘텐츠 저장소 솔루션은 페타바이트 및 그 이상의 확장성을 기반으로 데이터 가용성을 제공합니다. 쉽게 확장 가능한 IBM COS System은 콘텐츠를 지리적으로 분산하기 전에 원본 콘텐츠의 단일 사본에 캐리어급의 보안을 제공합니다. IBM COS 기술은 처음부터 끝까지 데이터 무결성을 보장합니다.

기업에는 인프라 전반에 분산되어 콘텐츠를 효과적으로 저장하고 배포하기 위한 콘텐츠 스토리지가 필요합니다. IBM COS는고가용성 환경, 장기적인 파일 무결성 및 액세스, 인증 수행 방식을 제공합니다. 기업에서 사용자들이 협업하는 콘텐츠가 덜 활성화되어 있든, 고정되어 있든, 또는 자주 액세스되든 관계없이, IBM COS는 강력하게 보호되고 신뢰할 수 있으며 비용 효율적인 방식을 제공합니다. 기업 내부 또는 외부의 사용자가 공유 콘텐츠 스토리지 저장소를 안전한 방식으로 액세스하여 지리적 위치에 관계없이 협업할 수 있습니다.

IBM COS vs. 일반 스토리지 파일 시스템



그림 18. IBM COS vs. 일반 스토리지 파일 서비스(\$/TB 비교)

사례 연구: 어느 메이저 리그 야구 팀

문제: 많은 메이저 리그 야구 팀과 마찬가지로, 동영상은 매우 중요합니다. 동영상 촬영자는 연중 모든 게임에서 모든 선수의 모든 투구와 모든 안타를 캡처합니다. 이러한 녹화 자료는 코치에게 중요합니다. 따라서 야구 팀은 많은 양의 비정형 동영상을 수집하여 저장하고, 보호하고, 코치와 직원들에게 제공합니다.

해결책: IBM COS는 클라우드 스토리지, 특히 아카이빙과 백업을 위한 이상적인 솔루션입니다. 그 이점으로는 뛰어난 확장성, 지리적 독립성, 멀티 테넌트 기능, 그리고 추가적인 비용 절감을 지원하는 즉시 사용 가능한(off-the-shelf) 상용 기술의 사용이 있습니다. 이 메이저 리그 야구 팀은 시스템의 백엔드에 IBM COS 소프트웨어를 배치하고, 원활한 데이터 액세스를 위해 글로벌 파일 시스템에 대한 로컬 인터페이스로 컨트롤러를 설정했습니다. 그 결과 IT 담당자에게 미치는 효과를 즉시 확인할 수 있었습니다. IBM COS 솔루션은 IT 전문가들이 백업 관리자로서 업무를 수행하고 모든 복제 루틴을 조정하는 데 소요되는 시간을 최대 30%까지 감소시켰습니다.

결과: IBM COS 기술은 이 메이저 리그 야구 팀에게 보안성을 저해하지 않고 각 데이터 사이트의 데이터에 대하여 교차 사이트 액세스를 구현하는 기능을 제공했습니다. IT 보안 부서장은 모든 사이트에서 모든 데이터를 최신 상태로 유지하고, 백업하고 복제할 수 있습니다. 코치, 스카우트 및 트레이너는 클라우드를 통해 로그인하고 필요한 대상을 검색하여 콘텐츠에 액세스할 수 있습니다. 이들은 향상된 데이터 보안과 원활한 데이터 배치 외에도, 사용자에 대한 시스템의 투명성을 높이 평가합니다. 이 팀은 IBM COS의 기술을 통해 경험한 성공을 기반으로 클라우드 스토리지 시스템을 마이너 리그의 계열 팀으로 확대하고 있습니다.

SaaS(Storage-as-a-service)

IBM COS 소프트웨어는 기업의 사용자 기반에 새로운 수준의 스토리지 용량 및 가용성과 함께 캐리어급의 데이터 보안을 제공합니다.

스토리지 용량 및 가용성 제공

중앙집중식 인프라에서 고객에게 용량을 판매할 수 있는 능력은 서비스 제공업체 및 대규모 엔터프라이즈가 반드시 갖춰야 할 사항입니다. IBM COS는 이러한 기업이 사용자와 소비자를 단일 플랫폼으로 통합하는 SaaS(Storage-as-a-service) 솔루션을 구현하는 데 도움이 됩니다. IBM COS는 관리를 간소화하고 스토리지를 요구사항에 맞게 효율적으로 확장할 수 있게 해줍니다. IT 부서는 보안 멀티 테넌시, 제로 터치 암호화 및 강력한 관리 API를 기반으로 확장이 가능하고 신뢰할 수 있으며 관리하기 쉽고 비용 효율적인 스토리지 오퍼링을 구축할 수 있습니다.

사례 연구: 데이터 스토리지 통합업체

문제: 업계 선두의 데이터 스토리지 통합업체가 관리되고 호스팅되는 데이터 스토리지 서비스를 확장하기로 결정했습니다. 이 업체는 잠재 고객의 대부분이 급속하게 성장하는 환경에서 비즈니스를 운영한다는 점을 고려할 때 자신이 재판매하는 플랫폼이 부족하다는 사실을 인식했습니다. 이러한 고객을 응대하고 비즈니스를 성장시키기 위해서 새로운 SaaS(Storage-as-a-service) 오퍼링이 필요했습니다.

해결책: IBM COS는 대량의 데이터를 비용 효율적으로 신뢰할 수 있게 저장하여 클라우드에 적합하게 만듭니다. 이 데이터 스토리지 통합업체는 철저한 평가 후에 IBM COS System을 선정하여, 퍼블릭 클라우드 제공업체보다 더 비용 효과적이고 배포하기 쉬운 경쟁력 있는 데이터 SaaS(Storage-as-a-service) 솔루션을 제공하게 되었습니다.

결과: IBM COS System은 이 데이터 스토리지 통합업체가 퍼블릭 클라우드 가격에 비해 경쟁력이 있으면서도 99.9999999% 이상의 데이터 안정성을 갖춘 스토리지 서비스를 제공할 수 있게 해주었습니다. 단순한 가격책정 모델은 특히 급속하게 성장하는 환경의 잠재 고객에게 매력적입니다. 또한 관리하기 쉬운 IBM COS System은 이 데이터 스토리지 통합업체가 고객에게 퍼블릭 클라우드 오퍼링에서 일반적으로 사용할 수 없는 성능 보고서, 용량 소비 보고서, 기술 업데이트와 같이 다양한 데이터 및 분석을 제공할 수 있게 해주었습니다.

기업내 협업

IBM COS는 가치 있는 콘텐츠에 대해 강력하게 보안되는 분산된 액세스를 제공하여, 전 세계에서 작업장 생산성 달성을 용이하게 합니다.

협업 및 생산성

오늘날의 인력은 끊임없이 이동하며 비즈니스의 범위는 전 세계에 걸쳐 있습니다. 성공을 위해서는 직원들이 사실상 언제 어디서나 미션 크리티컬 데이터에 원활하게 액세스할 수 있어야 합니다. IBM COS 솔루션은 기업이 데이터에 대한 글로벌 액세스를 제공할 수 있는 데이터 허브를 구축합니다. 제로 터치 암호화로 이를 온프레미스에서 동시에 보호하여, 기업내 협력을 달성하고 생산성을 향상시키는 강력한 보안의 분산 데이터 액세스를 제공합니다.

사례 연구: 주요 글로벌 마케팅 에이전시

문제: 업계 선두의 글로벌 마케팅 에이전시는 모든 미디어와 언어를 망라하여 고객을 위한 광고와 마케팅 커뮤니케이션을 제작합니다. 이 회사는 상당히 빠르게 성장하고 있었기 때문에 각 캠페인이 프로덕션 환경에서 상당한 규모의 스토리지 공간을 차지했습니다. IT 팀은 비정형 데이터의 증가하는 양을 비용 효과적으로 처리하는 스토리지 솔루션을 연구했습니다. 또한 이 팀은 해외 인력이 전 세계에서 그리고 중단 없이 협력할 수 있도록 지원하는 솔루션이 필요했습니다.

해결책: 데이터 중심의 많은 양의 콘텐츠에 최적화된 IBM COS 솔루션은 이 에이전시의 요구사항을 충족했으므로 모든 주요 스튜디오에 구현되었습니다. IBM COS는 주소 지정이 가능한 단일 글로벌 네임스페이스를 사용하여, 통합된 단일 관리 지점과 기존의 중앙 집중식 메타데이터 서버의 한계를 초월하여 확장 가능한 액세스를 제공합니다. 이제 이 에이전시의 모든 직원이 여기에 저장할 수 있고, 필요한 경우 엔드 오피스(end office) 중 어디서나 데이터를 다시 해당 시스템으로 가져올 수 있습니다. 프로덕션 스튜디오의 대다수가 서로 연결되어 있으므로 회사 전체에서 생산성이 향상되었습니다.

결과: IBM COS는 직원들이 훨씬 더 적극적으로 아카이빙하고, 자신이 사용하는 고비용 프로덕션 스토리지의 크기를 제한할 수 있게 합니다. 또한 IBM COS는 직원들이 오피스 간에 더 쉽게 협업할 수 있도록 지원하고, 고객에게는 필요한 자산에 대하여 더 쉽고, 더 안전하고, 더 안정성 있는 액세스를 제공하여 컴플라이언스를 지원합니다.

백업

IBM COS System은 신뢰할 수 있는 복구 및 보안을 위해 확장 가능한 백업과 항시적인 데이터 가용성을 최대 80% 낮은 인프라 비용으로 제공합니다.

비용 효과적이고 보안이 우수하며 액세스 가능한 스토리지

IT 부서에서 다양한 애플리케이션 서버와 사용자 시스템의 데이터를 수집하고 백업하는 일은 어려운 과제입니다. 이러한 데이터를 장기적으로 그리고 대규모로 저장하는 일은 더욱 어렵습니다. IBM COS Backup Solutions는 장기 데이터 보호를 위해 비용 효과적이고, 액세스하기 쉬운 스토리지 플랫폼을 제공합니다. SecureSlice 제로 터치 암호화는 데이터를 비용 효과적으로 여러 사이트에 분산하기 전에 보호하여 장기적으로 비트 단위까지 완벽한 보호를 대규모로 수행하도록 지원합니다. IBM COS System은 데이터가 일단 백업되면 데이터에 대한 더 빠른 액세스를 지원하여, 재해 발생 시 비즈니스 복구 시간을 단축합니다.

사례 연구: 주요 소매업체

문제: 주요 소매업체에서 동영상, 사진 등을 포함한 비정형 데이터의 지속적인 생산이 놀랄 만한 속도로 증가하고 있습니다. 이와 동시에, IT 스토리지 플랫폼을 과도한 가격 인상이나 안정성 저하 없이 확장하기 위해 노력하고 있었습니다.

해결책: IBM COS는 기업이 페타바이트 규모 이상의 데이터를 효율적으로 저장, 관리 및 액세스할 수 있게 해줍니다. IBM COS 솔루션은 정방향 오류 수정의 일종인 삭제 코딩을 사용하여 다른 스토리지 솔루션보다 훨씬 우수한 데이터 복원성을 제공하며, 일반적인 오브젝트 스토리지 솔루션에 비해 스토리지 용량 요구사항도 훨씬 적습니다. IBM COS는 오브젝트 소프트웨어 스토리지의 확장성, 효율성, 보안 및 단순성을 입증했으며 이 소매업체의 고위 경영진은 이를 신속하게 채택했습니다.

결과: 이 소매업체의 스토리지는 2015년 말에 약 0.5페타바이트(500테라바이트)에 달할 것으로 전망되었습니다. IBM COS Backup Solution은 방대한 양의 동영상 및 시각 자료를 더 쉽고 경제적으로 저장합니다. IBM을 선택한 가장 큰 동기는 비용 효율성이었으며 비용 절감 효과는 빠르게 나타났습니다. 이 소매업체는 이전의 스토리지 공급업체와 비교하여 (바이트당) 약 50%의 비용 절감을 실현했습니다. IBM COS의 데이터 슬라이싱과 구조화 방식은 중요한 정보를 손상시키거나 가로채기가 사실상 불가능하여 데이터 보호를 보장합니다.

활성 아카이브

IBM COS System은 확장 가능하고, 신뢰할 수 있고 보안이 강력한 장기 데이터 아카이브를 사용하여 콘텐츠를 액세스 가능한 상태로 유지합니다.

확장 가능하고 신뢰할 수 있는 강력한 보안의 아카이브 스토리지

많은 기업들이 사용자에게 데이터에 대한 신속한 액세스를 제공하는 아카이브 솔루션을 모색합니다. IBM은 사실상 무제한적인 가용성과 가장 높은 수준의 데이터 무결성 및 기밀성을 갖춘 아카이브 스토리지 솔루션을 제공합니다.

장기 아카이브의 정보는 그다지 자주 읽히지 않으므로, 장기 아카이브의 상태를 정상으로 유지하는 데는 사전 대응적인 오류 수정이 중요합니다. IBM COS는 스토리지 노드를 스캔하고 오류를 점검하여 수정하는 지능형 백그라운드 프로세스를 사용합니다.

IBM COS 솔루션은 여러 데이터 센터에 배포할 수 있습니다. 여러 오프사이트 위치에 분산된 아카이브는 데이터를 잠재적인 단일 위치 오류 또는 돌발적인 재해로부터 보호하는 데 도움이 되며, 장기 보존을 위해 더 안전하게 액세스할 수 있습니다. 특정 서버 또는 스토리지 장치에 지정되어 있지 않고, 새로운 스토리지 노드가 시스템에 설치될 때 데이터가 자동으로 재구성되기 때문입니다. IBM COS Archive Solution은 기업이 컴플라이언스 요구사항과 장기 보존 목표를 충족하도록 설계되었습니다.

사례 연구: 인터넷 기반 사진 게시 서비스

문제: 지속적으로 이동하고 증가하는 수십억 개의 사진으로 구성된 활성 아카이브를 가진 기업이 고객의 요구사항에 부응하고 성능, 가용성 및 안정성을 동일한 수준으로 유지해야 한다는 중대한 과제에 직면했습니다. 페타바이트 규모의 원시 스토리지와 두 자릿수의 성장을 때문에, 이러한 데이터를 저장하기 위한 비용 역시 급속하게 증가하고 있었습니다. 스토리지 비용을 절감하고 더 쉽게 관리할 수 있는 대안을 찾아야 했습니다.

해결책: 데이터 수집 및 저장을 위한 IBM COS 방식은 가장 중요한 기준 - 단일 장애 포인트를 제거하고 높은 수준의 오류 허용성을 제공한다는 기준을 근본적으로 해결했습니다. 또한 IBM COS 솔루션에는 이미지 아카이브의 관리를 훨씬 단순하게 만드는 여러 가지 특성이 있었습니다. 고급 삭제 코딩 기법은 개별 컴포넌트의 성능 및 안정성을 애플리케이션 수준의 성능 및 안정성으로부터 분리합니다. 따라서 이 회사는 데이터의 지속적인 가용성을 보장하게 되었으며, 스토리지 계층에서 잠재적인 하드웨어 및 소프트웨어 문제로부터 영향을 훨씬 덜 받게 되었습니다.

결과: 현재 이 회사는 프로덕션 환경에서 150PB 이상의 스토리지를 운영하고 있으며 확장 가능한 무제한 용량을 기반으로 급속하게 성장하고 있습니다. 또한 전사적으로 상당한 전력 소비 및 관리 비용 절감을 실현하고 있습니다. 이제 수십억 개의 오브젝트가 포함되고 150PB 이상의 용량을 갖춘 전체 스토리지 플랫폼을 파트타임 스토리지 관리자 3명만으로 관리할 수 있습니다.

5부 결론

많은 양의 비정형 데이터를 저장해야 하는 기업은 기존 스토리지 솔루션 외에 새로운 방식을 고려해야 합니다. 분산 스토리지는 보안, 가용성 및 안정성을 보장하는 데 도움이 되는 동시에 많은 양의 비정형 데이터를 비용 효율적으로 저장하기 위한 혁신적인 방식의 하나입니다.

분산 스토리지의 완벽 가이드는 분산 스토리지의 기능 및 이점을 다섯 가지 주요 영역으로 나누어 설명했습니다.

- **가용성** - 계획된 또는 계획되지 않은 다운타임에 관계없이 데이터를 항상 사용할 수 있습니다.
- **확장성** - 시스템을 테라바이트에서 페타바이트, 엑사바이트 규모로 쉽게 확장할 수 있습니다.
- **보안** - 여러 드라이브, 서버, 컨테이너 또는 위치가 손상되더라도 데이터 기밀성이 유지됩니다.
- **경제성** - 비용이 높은 복제의 필요성이 사라지므로, 페타바이트 수준 및 그 이상의 스토리지 시스템에 대한 총소유비용이 크게 감소했습니다.
- **효율성** - 관리자 한 명당 수십 페타바이트의 스토리지를 보다 쉽게 관리할 수 있습니다.

이러한 이유로, 분산 스토리지는 대량의 비정형 데이터를 저장해야 하고 지연 시간이 주요한 고려 사항이 아닌 엔터프라이즈에 이상적인 솔루션입니다. 콘텐츠 스토리지, 활성 아카이브 그리고/또는 콘텐츠 분산 요구사항이 있는 엔터프라이즈는 분산 스토리지를 또 다른 기술 옵션으로 고려해야 합니다.

분산 스토리지가 귀사에 적합한가?

귀사가 분산 스토리지에서 이점을 얻을 수 있는지 확인하려면 다음과 같은 체크리스트를 사용하십시오.

- 데이터를 장기 보존해야 하는 애플리케이션이 있습니까?
- 데이터가 이미지, 동영상 및 문서와 같이 크기가 큰 비정형 데이터로 구성되어 있습니까?
- 이러한 데이터를 위해 500테라바이트 이상을 사용할 수 있습니까?
- 데이터 보안, 가용성, 확장성 및 비용 효율성을 위한 요구사항이 있습니까?
 - 네트워크 연결
 - 고품질 대역폭(지리적으로 분산된 경우)

IBM Cloud Object Storage에 대한 정보

IBM Cloud Object Storage는 오늘날 급속하게 성장하는 비정형 데이터를 하이브리드 클라우드 환경에서 저장, 관리 및 액세스하기 위해 필요한 유연성, 확장성 및 단순성을 제공합니다. 세계 최대의 저장소들을 기반으로 하는 IBM의 검증된 솔루션은 스토리지 비용을 절감하는 동시에, 기업용 모바일, 소셜, 분석 및 인지 컴퓨팅을 위한 기존의 워크로드와 클라우드에서 발생하는 새로운 워크로드를 모두 안정적으로 지원하여 스토리지 문제를 비즈니스 이점으로 전환합니다. IBM Cloud Object Storage는 IBM이 2015년에 인수한 오브젝트 스토리지 분야의 리더인 Cleversafe의 기술을 기반으로 합니다.

자세한 정보

(02)3781-6900으로 전화하거나, 이메일 mktg@kr.ibm.com으로 연락하거나, 웹사이트 <http://www.ibm.com/cloud-computing/kr-ko/infrastructure/object-storage/>을 방문하십시오.

IBM Cloud 컴퓨팅에 대한 자세한 내용을 보려면 <http://www.ibm.com/cloud-computing/kr/ko/>를 방문하십시오.



© Copyright IBM Corporation 2016

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
2016년 7월

IBM, IBM 로고, ibm.com 및 Cleversafe는 미국 또는 기타 국가에서 사용되는 International Business Machines Corporation의 상표 또는 등록상표입니다. 이와 함께 기타 IBM 상표가 기재된 용어가 상표 기호(®tm; 또는 &tm;)와 함께 이 정보에 처음 표시된 경우, 이와 같은 기호는 이 정보를 발행할 때 미국에서 IBM이 소유한 등록상표 또는 일반 법적 상표입니다. 또한 이러한 상표는 기타 국가에서 등록상표 또는 일반 법적 상표입니다. 현재 IBM 상표 목록은 웹 "저작권 및 상표 정보"(ibm.com/legal/copytrade.shtml)에 있습니다.

Apple, iPhone, iPad, iPod touch 및 iOS는 전세계 여러 국가에 등록된 International Business Machines Corporation의 상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

기타 회사, 제품 및 서비스 이름은 타사의 상표 또는 서비스표입니다.

이 문서는 최초 발행일을 기준으로 하며, 통지 없이 언제든지 변경될 수 있습니다. IBM이 영업하는 모든 국가에서 모든 오퍼링이 제공되는 것은 아닙니다.

인용된 성능 데이터와 고객 예제는 예시 용도로만 제공됩니다. 실제 성능 결과는 특정 구성과 운영 조건에 따라 다를 수 있습니다. 그러나 IBM 제품 및 프로그램과 함께 사용한 기타 다른 제품이나 프로그램의 운영에 대한 평가와 검증은 사용자의 책임입니다. 이 문서의 정보는 상품성, 특정 목적에의 적합성에 대한 보증 및 타인의 권리 비침해에 대한 보증 및 타인의 권리 비침해에 대한 보증이나 조건을 포함하여(단, 이에 한하지 않음) 명시적이든 묵시적이든 일체의 보증 없이 "현상태대로" 제공됩니다. IBM 제품에 대한 보증은 제품의 준거 계약 조항에 의거하여 제공됩니다.

우수 보안 관리제도에 대한 설명: IT 시스템 보안은 귀사 내/외부로부터의 부적절한 접근을 방지, 감지, 대응함으로써 시스템과 정보를 보호하는 일을 포함합니다. 부적절한 접근은 정보의 변경, 파괴 또는 유출을 초래하거나, 타 시스템에 대한 공격을 포함한 귀사 시스템에 대한 피해나 오용을 초래할 수 있습니다. 어떠한 IT 시스템이나 제품도 완벽하게 안전할 수 없으며, 단 하나의 제품이나 보안 조치만으로는 부적절한 접근을 완벽하게 방지하는 데 효과적이지 않을 수 있습니다. IBM 시스템과 제품은 종합적인 보안 접근방법의 일부로서 고안되며, 이러한 접근방법은 필연적으로 추가적인 실행절차를 수반하며 가장 효과적이기 위해서는 다른 시스템, 제품 또는 서비스가 필요할 수도 있습니다.

1 Nadkarni, A.; IDC MarketScape: Worldwide Object-Based Storage 2014 Vendor Assessment Report, 2014년 12월; IDC #253055e; https://www.cleversafe.com/documents/public/IDC-MarketScape-Worldwide-Object-Based-Storage-2014-Vendor-Assessment-Cleversafe.pdf?cm_mc_uid=88103752002214484629162&cm_mc_sid_50200000=1450709958

2 Ibid, 2014.

3 Reichman, A., Forrester Research: Prepare for Object Storage in the Enterprise, 2010년 11월 29일, <https://www.forrester.com/Prepare+For+Object+Storage+In+The+Enterprise/fulltext/-/E-res57953>

4 "Objectively better - leverage object-based storage to store petabytes of data efficiently," DMD Data Systems, Inc., 블로그, 2016년 4월, <https://dmddatasystems.com/blog/2016/04/leverage-object-based-storage/>



재활용하십시오