

## HCP CS, 클라우드 워크로드에 최적화된 ‘오브젝트 스토리지’의 간판 주자

2021년 IDC가 진행한 오브젝트 스토리지 시장 활용 설문 결과를 보면, 응답자의 80%는 오브젝트 스토리지가 IoT, 보고, 분석을 포함한 가장 중요한 IT 이니셔티브를 지원할 수 있다고 답했다. 스토리지가 단순한 인프라가 아닌 기업의 전략을 지원하는 기반이 될 수 있다는 뜻이다. 그런 면에서 오브젝트 스토리지는 혁신의 지렛대이자 IT 관리자의 고민을 해결해 주는 중요한 열쇠가 될 것이다.

IT 전문 인력이 적거나 오브젝트 스토리지를 신속하게 도입하고자 하는 기업이라면 효성인포메이션시스템의 HCP CS 같은 상용 오브젝트 스토리지는 좋은 선택지가 될 수 있다.

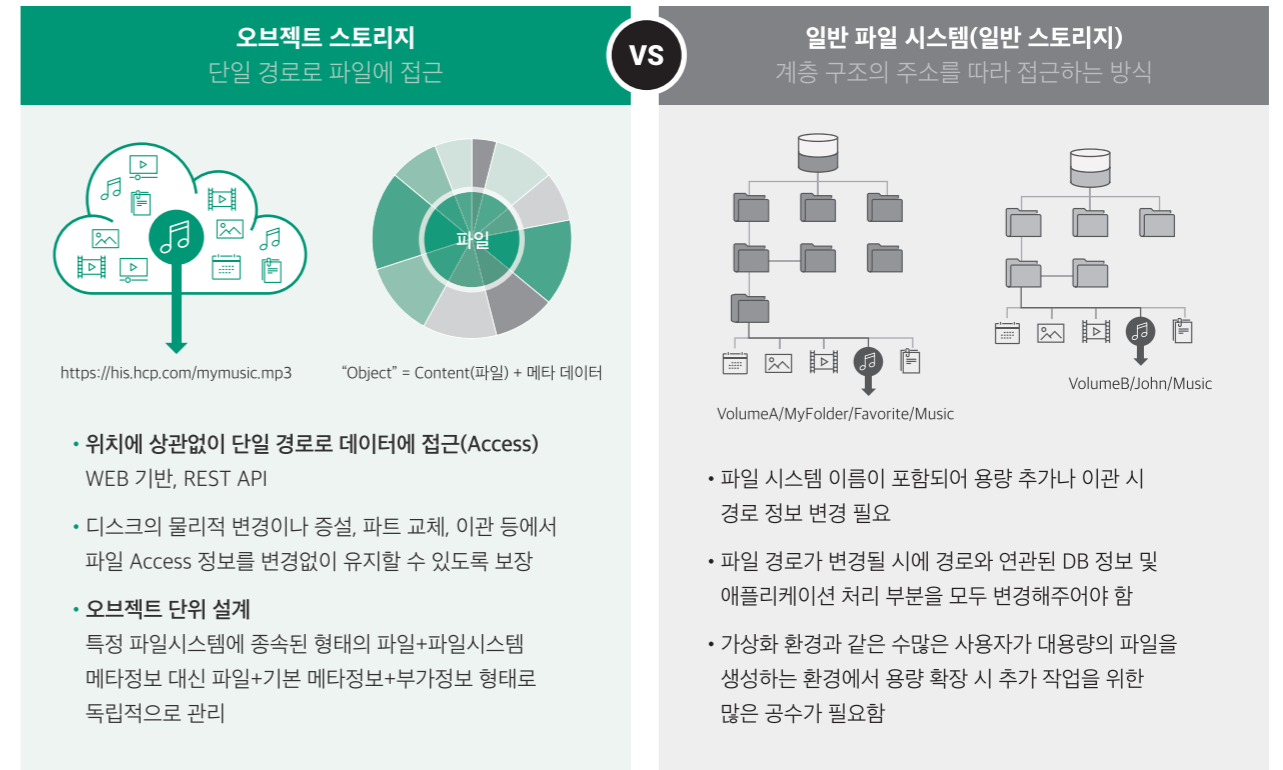
### 오브젝트 스토리지 대세, 이유가 있다

IT 미디어 네트워크월드에 따르면, ‘오브젝트 스토리지’는 최근 몇 년간 스토리지 트렌드를 이끄는 핵심으로 자리 잡고 있다. 여기에는 몇 가지 배경이 있다. 먼저 분석, 데이터 레이크, AI, 머신러닝 기술 등과 관련된 애플리케이션을 위해 오브젝트 스토리지를 주로 활용한다는 점이다. 또한 NVMe, SSD의 성능 개선과 가격 하락으로 플래시 디스크를 포함한 다양한 미디어를 채택한 오브젝트 스토리지 도입이 가속하고 있다. 오브젝트 스토리지가 주목받는 이유는 현대 IT 관리자가 직면한 스토리지 관련 과제를 통해 파악할 수 있다.

기업은 끊임없이 늘어나는 데이터를 처리하기 위해 자주 사용하지 않는 데이터는 조금 더 저렴한 디스크로 옮겨 놓는다. 엑스레이 사진이나 보험 계약서처럼 법적인 요소나 사규 때문에 장기간 원본 데이터를 보관해야 할 때도 있다. 여기에 중요한 데이터는 한 번 저장하는 것에 그치지 않고 백업을 해야 한다. 백업한 뒤에도 복구 시간을 빠르게 하거나 재해에 대비하기 위해 이중화나 재해복구(DR)도 신경 써야 한다.

업무 지원 환경도 달라지고 있다. 인프라 담당자는 분석 업무를 맡은 타 팀으로부터 데이터 접근성을 높여 달라는 요구를 빈번히 받는다. 이를 위해서는 특정 업무나 기종에 제한을 두지 않고, 기업 내부나 원격지에 상관없이 데이터에 빠르게 접근할 수 있어야 한다. 또한 파일 원본 그대로 유지되는 환경도 마련해야 한다.

↓ 오브젝트 스토리지의 구조



시스템 운영 과정에서는 기술적 부담이나 관리 비용이 증가하고 있다는 점이다. 분산된 스토리지를 통합 관리하는 방법은 무엇인지, 지울 수 없거나 중복으로 저장되는 데이터를 어떻게 처리할지에 대해서도 해결책이 필요하다.

오브젝트 스토리지는 블록이나 파일 단위가 아닌 ‘오브젝트’라는 개별 단위로 데이터를 저장한다. 오브젝트에는 텍스트, 음성, 영상 같은 다양한 데이터가 포함되며, 파일에 대한 추가 정보를 담은 메타데이터도 있다. 또한 일반적인 파일 탐색기에서 볼 수 있는 중첩되는 계층 구조가 아닌 평면 구조로 데이터가 저장된다.

이 때문에 오브젝트 스토리지는 대량의 데이터를 안전하게 보관하면서 데이터에 빠르게 접근할 수 있다. 파일 스토리지는 대량 파일이 저장되거나 디렉토리 수가 많아질수록 성능이 저하된다. 오브젝트 스토리지는 내부 분산 DB를 활용하기 때문에 데이터 양이 늘어나도 데이터 접근 및 검색 속도가 빠르다.

### 컴플라이언스·콘텐츠 데이터 저장에 유용

오브젝트 스토리지가 필요한 환경으로는 디지털 아카이빙, 즉 콘텐츠를 통합하고, 보관·관리하는 곳을 들 수 있다. 내부 규범이나 법적 규제를 받아서 장기간 반드시 데이터를 보존해야 할 때도 유용하다. 이런 유형의 데이터는 시간이 지나거나 사업이 확장될수록 늘어나는데, 오브젝트 스토리지는 메타데이터를 활용해 대량의 파일을 체계적으로 관리할 수 있어 편리하게 운영할 수 있다.

두 번째, 다양한 콘텐츠를 보관해야 할 때다. 페이스북, 트위터, 카카오톡 등의 서비스는 시스템이 아닌 개인 사용자 단에서 데이터를 실시간으로 생성하고 복제한다. 또한 IoT가 일반화되면서 기업 내의 다양한 기기나 장치들로부터 파일 데이터가 생성되고 있다. 이렇게 만들어진 데이터는 내부 인프라 어딘가에 저장해야 한다. 모든 데이터를 처음 그 상태로 보존하면 좋겠지만, 이를 위해 스토리지를 무한정 구매할 수는 없는 노릇이다. 즉, 비용 대비 효과를 높이기 위해 오브젝트 스토리지에 데이터를 저장할 수밖에 없는 상황이다.

### 클라우드의 익숙함과 온프레미스 관리 용이성이 만드는 시너지

오브젝트 스토리지가 대중화하기까지 아마존 S3 같은 클라우드 서비스가 큰 역할을 했다. 클라우드는 초기 구축 비용이 적고 접근하기 쉽다. 하지만 기업 환경에서는 클라우드 기반 오브젝트 스토리지를 활용할 수 없는 환경이 있다. 대표적으로 컴플라이언스 영향을 받을 때나 규모가 큰 데이터를 저장할 때다. 전자는 데이터 주권 등의 이유로 클라우드에 올리지 못하는 데이터가 있어 이용하지 못하는 것이고, 후자는 비용 때문이다. 클라우드 서비스는 데이터 저장 외에도 데이터를 전송하거나 호출하는 등 활용 과정에서도 비용을 부과한다. 이런 구조라면 대규모 프로젝트를 운영해야 할 경우 비용 책정 방식이 복잡해지고 지불해야 하는 금액도 높아진다.

하지만 사용자 입장에서 클라우드의 장점을 선뜻 버리기 힘들다. 이미 데이터 프로젝트를 클라우드 서비스에 맞춰 진행했다면 익숙한 개발 환경을 포기하기가 쉽지 않다. 이에 효성인포메이션시스템은 클라우드의 장점과 온프레미스 오브젝트 스토리지의 장점을 모두 취할 수 있도록 HCP 클라우드 스케일(Hitachi Content Platform Cloud Scale, 이하 HCP CS)을 내놓았다.

### 효성인포메이션시스템의 야심작, HCP CS

HCP CS는 프라이빗 클라우드 스토리지 소프트웨어 솔루션이다. 소프트웨어 정의 오브젝트 스토리지가기에 특정 하드웨어에 종속되지 않고 어떤 플랫폼이든 설치 가능하다. 값비싼 전용 스토리지 장비 없이 x86 계열 서버가 있다면 설치할 수 있고, 리눅스 종류에 상관없이 도커 버전만 맞으면 마이크로 서비스 구조로 운영할 수 있다. 수백 테라에서 페타 단위 데이터를 계속 저장할 수 있어 단일 클러스터 기준으로 버킷, 용량, 파일이 무제한으로 늘어나도 대응할 수 있다.

HCP CS의 또 다른 특징은 AWS의 아마존 S3 인터페이스 및 S3 API를 지원하는 온프레미스 오브젝트 스토리지라는 점이다. 사용자는 아마존 S3의 데이터를 쉽게 온프레미스 기반 스토리지로 옮기거나 하이브리드 클라우드 형태로 같이 쓸 수 있다. 많은 기업이 아마존 S3의 인터페이스에 익숙하고 데이터를 S3에 저장해 놓았다는 점에 주목해 이런 기능을 마련한 것이다.

만약 내부 규정 때문에 특정 데이터를 퍼블릭 클라우드에 업로드할 수 없거나 한국 외 타 국가에 데이터가 저장되는 상황을 피하고 싶다면, 프라이빗 오브젝트 스토리지를 이용하면 좋다. 이미 S3와 같은 클라우드 스토리지 구조에 맞춰 데이터 프로젝트를 진행했다고 해도 익숙한 인터페이스를 통해 빠르게 데이터를 전환할 수 있다.

↓ HCP CS의 특징점



### 모니터링부터 파일 보호까지 가능

효성인포메이션시스템은 시중에 있는 인프라 기술을 버전별로 테스트해 HCP CS가 정상적으로 작동하는 설치 환경 가이드라인을 제공한다. 어플라이언스 형태로도 지원하고 있다. 데이터 규모가 가파르게 증가하는 상황에서 내부적으로 가용성 설계와 검증에 어려움을 겪고 있다면 어플라이언스가 도움이 될 수 있다. HCP CS의 세부적인 특징은 다음과 같다.

### 멀티테넌시 아키텍처

물리적인 오브젝트 스토리지 안에서 논리적으로 버킷을 생성해 서비스할 수 있으며, 버킷을 사용자, 용도, 부서 등에 따라 분할하고 배분할 수 있다. 각 사용자는 버킷 안에서 파일을 공유할 때 서브 폴더를 만들어 전달할 수 있다. 논리적으로 분리된 버킷은 각각 사용량, 제한 용량, 인증 등을 S3 API 환경에서 상세하게 모니터링 할 수 있다.

### 분산 DB 구조

파일 및 용량이 급격히 늘어날 때 관리하기 용이하도록 분산 DB 구조를 채택하고 있다. 일반적인 파일 시스템은 용량이 커질수록 파일 메타 정보도 커지면서 검색 성능이 저하될 수 있다. HCP CS는 메타데이터 파티셔닝 기법을 이용해 모든 메타 정보가 분산되어 효율적으로 관리 및 운용되도록 설계되었다.

### 파일 보호 지원·암호화 및 무결성 검증

오브젝트 스토리지는 컴플라이언스 규제 때문에 오래 보관하는 데이터를 많이 저장한다. HCP CS에서는 지정한 기간에는 데이터 삭제와 훼손이 안 되도록 설정하거나 파일을 정상적으로 올렸는지, 깨진 파일은 없는지 확인할 수 있다. 파일을 5년 동안 보관해야 한다면 일종의 태그를 붙여 5년 뒤에 지울 수 있도록 설정하는 등 스토리지 레벨에서 유용한 기능도 지원한다.

### 실무자를 돕는 최적의 운영 관리 기능

클라우드 기반 오브젝트 스토리지에서는 버킷을 만들고 배포하고 용량을 증설하는 등의 일은 클라우드 서비스 벤더가 처리한다. 하지만 온프레미스 스토리지 환경을 구축하면 해당 업무는 온전히 인프라 담당자의 몫이 된다. 이 과정에 필요한 것은 당연히 모니터링 도구다. HCP CS에서는 별도의 프로그램을 설치하지 않아도 GUI 환경에서 기본 관리 기능을 제공하며, 오브젝트 스토리지 구성 요소를 관리하거나 계정 관리, 보안, 알림, 모니터링도 가능하다.

### 빅데이터 시대 IT 이니셔티브를 이끄는 주춧돌

무제한급의 프라이빗 오브젝트 스토리지 솔루션은 어디서 빛을 발할 수 있을까? 최근에는 수많은 데이터를 보유하고 있으나 데이터 보호 정책을 확립하기 쉽지 않은 제조 환경에서 관심을 보인다. 일례로, 국내 한 제조사는 운영 중인 제조 라인의 공정 데이터 수집 인프라의 1차 저장소로 오브젝트 스토리지를 선택했다. 이 기업은 기존 스토리지 인프라 운영 비용 절감과 대용량 데이터 운영의 효율성을 개선하기 위해 페타바이트 규모의 대형 오브젝트 스토리지를 도입했다. 이를 통해 궁극적으로 딥러닝 및 빅데이터 분석에 필요한 장기간 아카이빙 인프라를 확보할 수 있었다.

해외에서는 대용량 하둡파일시스템이나 병렬파일시스템의 데이터 오프로드 구성을 위해 오브젝트 스토리지를 도입하는 사례가 늘고 있다. 하둡은 3개의 사본을 두는 것이 기본 정책이기 때문에 하나의 데이터를 저장할 때 3배의 데이터 용량을 차지한다. 이때 티어링 또는 아카이빙 용도로 오브젝트 스토리지를 이용하면 용량 대비 비용 효율적으로 데이터 분석 인프라를 구성할 수 있다.

