

The background features a series of overlapping, three-dimensional rectangular blocks in various shades of gray, creating a sense of depth and perspective. A solid teal vertical bar is positioned on the right side of the page.

# Hitachi Content Software for File

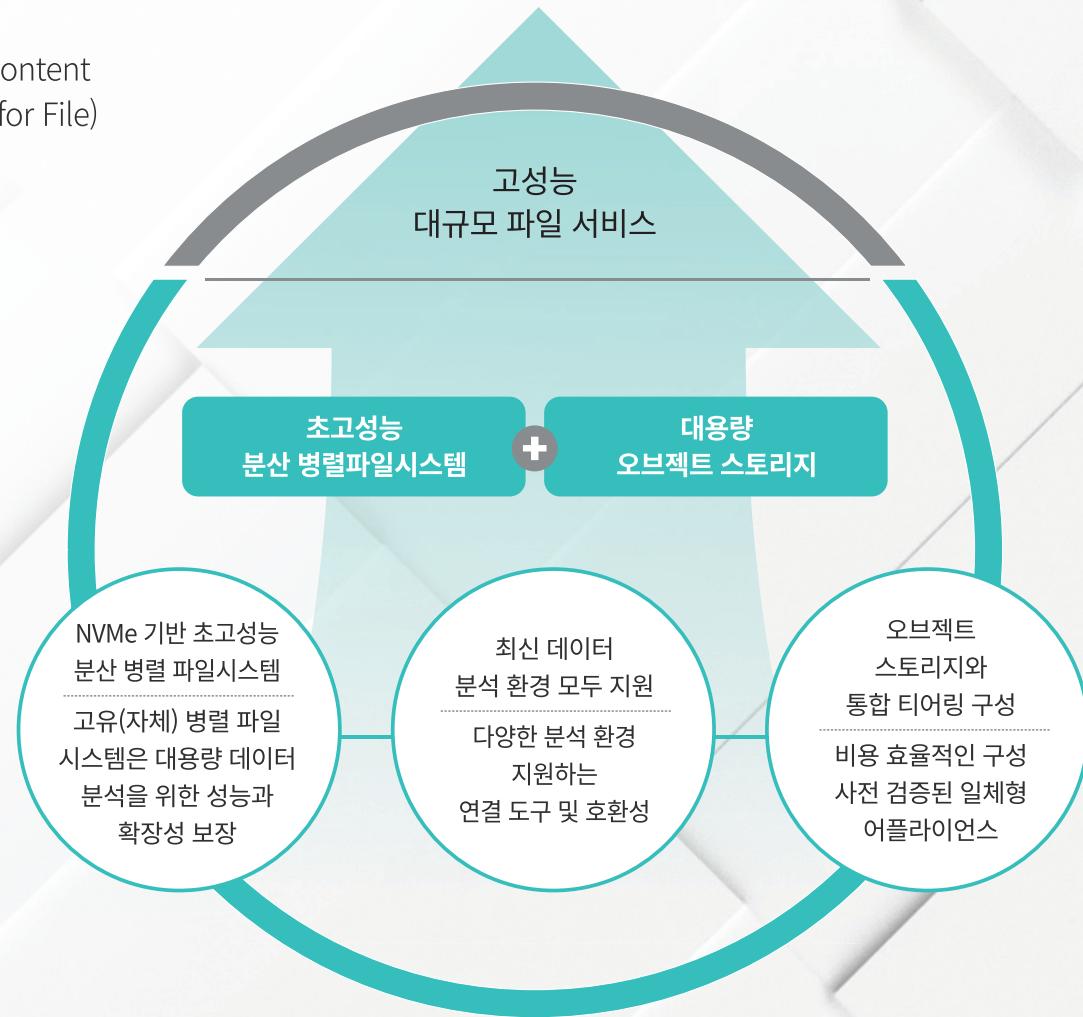
초고성능 파일 스토리지 솔루션

# HCSF

## Hitachi Content Software for File

세계에서 가장 빠른 상용 분산 병렬 파일시스템과  
업계 최상의 오브젝트 스토리지 HCP가 함께 제공됩니다

HCSF  
(Hitachi Content Software for File)  
특장점



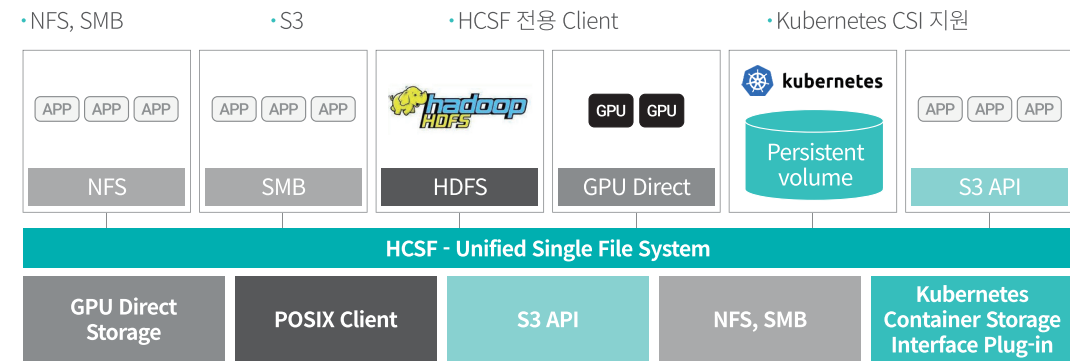
디지털 시대에 이전에 없었던 많은 데이터가 생성되고 있으며, 특히 이미지, IoT 같은 비정형 데이터의 고성능 AI/ML 분석이 요구되고 있습니다. AI/ML 분석은 데이터 '수집'부터 데이터 '스테이징 & 아카이브' 단계까지 작고 큰 파일 크기에 대한 성능 차이가 없어야 하며, 높은 처리량, 높은 IOPS, 매우 짧은 대기시간이 제공되어야 합니다

### 적용 분야

제조/통신	공공/바이오	금융 서비스
<ul style="list-style-type: none"> <li>· 자율주행 자동차</li> <li>· 스마트 팩토리 프로젝트</li> <li>· 실시간 이미지/음성인식 분석</li> <li>· 미디어/엔터테인먼트 렌더링</li> <li>· 고속 송출 서비스</li> </ul>	<ul style="list-style-type: none"> <li>· 디지털 포렌식</li> <li>· 해양/지리정보 조사</li> <li>· 기상 데이터 예측</li> <li>· 생명과학 연구 및 유전자 분석 프로젝트(이미지 및 영상 데이터 분석에 최적화)</li> </ul>	<ul style="list-style-type: none"> <li>· 실시간 고객데이터 분석</li> <li>· 금융 포트폴리오 평가시스템 구축</li> <li>· GDPR, IFRS 17 등 각종규제 대응, 리스크 관리</li> </ul>

### 멀티프로토콜 지원

HCSF는 AI/ML 분석 워크로드에 적합한 고성능 스토리지로 다양한 멀티프로토콜을 지원 및 HPC 환경 어플리케이션 용도에 맞춰 활용 가능

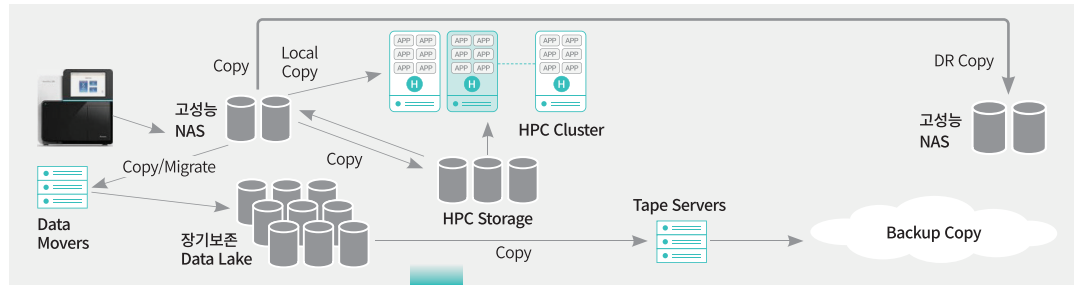


### 상세 스펙

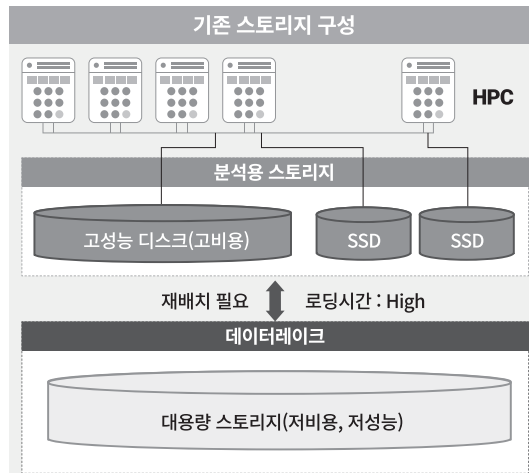
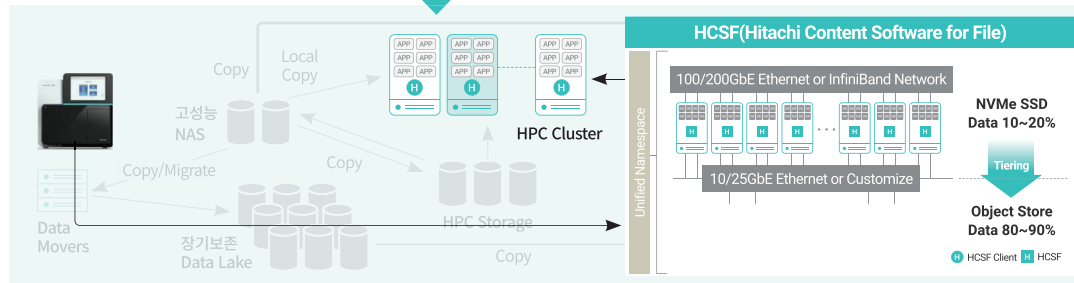
구분	상세 내용			
	HCSF30110	HCSF20220	HCSF10448	
Node Unit	1U	2U	4U	
Min. Node	6	6	8	
Max. Node	무제한	무제한	24	
Node당 Spec	CPU	32core	32core / 64core	24core
	Memory	384GB	512GB	256GB
	Disk	Data 수: 최대 NVMe SSD * 10개 Data 타입: 3.84TB / 7.68TB / 15.36TB / 30.72TB	최대 NVMe SSD * 20개	최대 NVMe SSD * 6개
Network	Ethernet : 10GbE/ 25GbE/ 100GbE/ 200GbE InfiniBand : EDR / HDR			
Protocol	POSIX, GDS, NFS, SMB, S3, K8S CSI			
Tiering	Object Storage HCP 또는 S3 Compatible Public Cloud			

### AI/ML 분석 워크로드에 적합한 고성능 스토리지

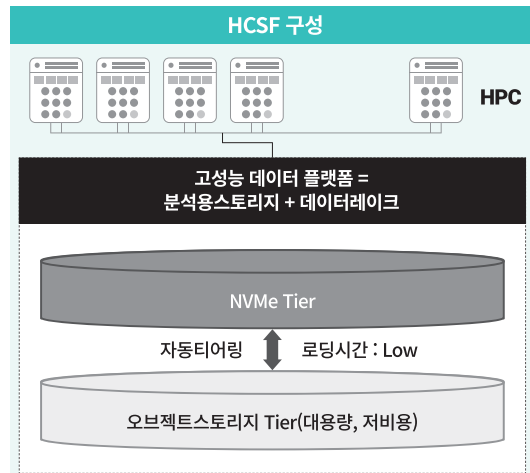
- 고성능 파일 시스템과 무한 확장성을 갖는 오브젝트 스토리지를 하나의 어플라이언스 형태로 제공하여 복잡한 스토리지 환경을 단순화 하고, 데이터 중복과 데이터 사일로를 제거해 손쉬운 관리 가능



### Zero-Copy 아키텍처



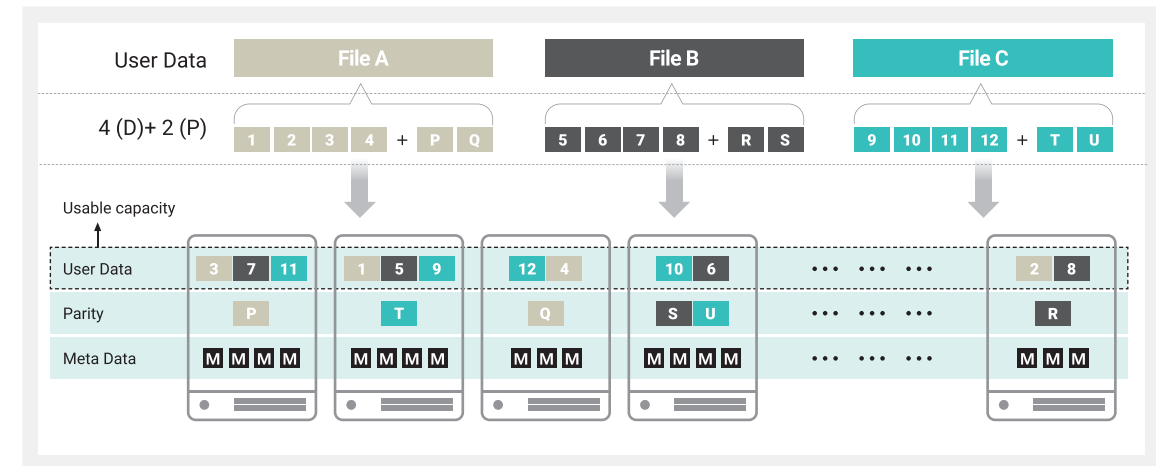
- #### 기존 스토리지 한계점
- 데이터를 GPU 분석 영역으로 로딩하는데 생각보다 많은 시간 소요
  - 성능 때문에 로컬 SSD 사용 시 여러 서버(GPU 서버/ CPU 서버)에서 동일 데이터 동시 접근 처리 어려움(데이터 사일로)
  - 데이터레이크가 너무 느려서 원하는 데이터를 적시에 찾기 어려움



- #### HCSF
- 하나의 플랫폼에서 분석용 고성능 스토리지와 데이터레이크를 같이 사용
  - 데이터 로딩 시간 감소(거의 zero) 여러 서버에서 동시에 동일 데이터 접근 처리 가능
  - 여러대의 GPU/CPU 서버에서 동시 처리 가능하므로 성능향상 있음

### 데이터 & 메타데이터 분산 저장, 처리를 통한 고성능 보장

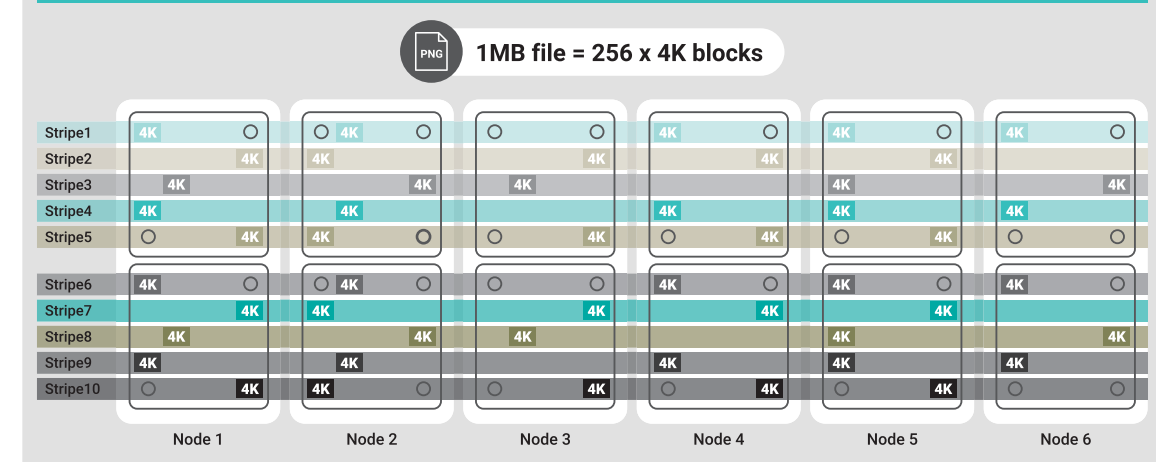
- 완전한 분산 병렬 파일시스템으로 데이터를 Extent 단위(4k Block의 결합)로 분산하여 완벽하게 처리
  - 데이터는 Erasure Coding 기반(N+2 또는 N+4 단위)으로 모든 노드와 디스크에 분산 처리데이터 보호
  - 메타데이터도 모든 노드에 분산 처리하며, 메타데이터가 별도 구성되었을 때 관리로 인한 이슈 차단



### Small IO 처리에 특화된 아키텍처

- 데이터는 완전 분산되고 전체 NVMe SSD pool에 걸쳐 4K Block으로 Write 수행 - no 'Hot Spots'
  - 파일을 1MB 단위의 Extent의 체인으로 처리하며 하나의 Extent는 4k Block의 집합으로 구성
  - 4KB 블록 크기는 NVMe 미디어 4K 블록과 일치하여 NVMe의 완벽한 성능을 제공(타사의 경우 64K-1MB)
- 파일크기에 관계없이 아주 작은 크기의 파일 뿐만 아니라 대용량의 파일 처리에도 처리 가능

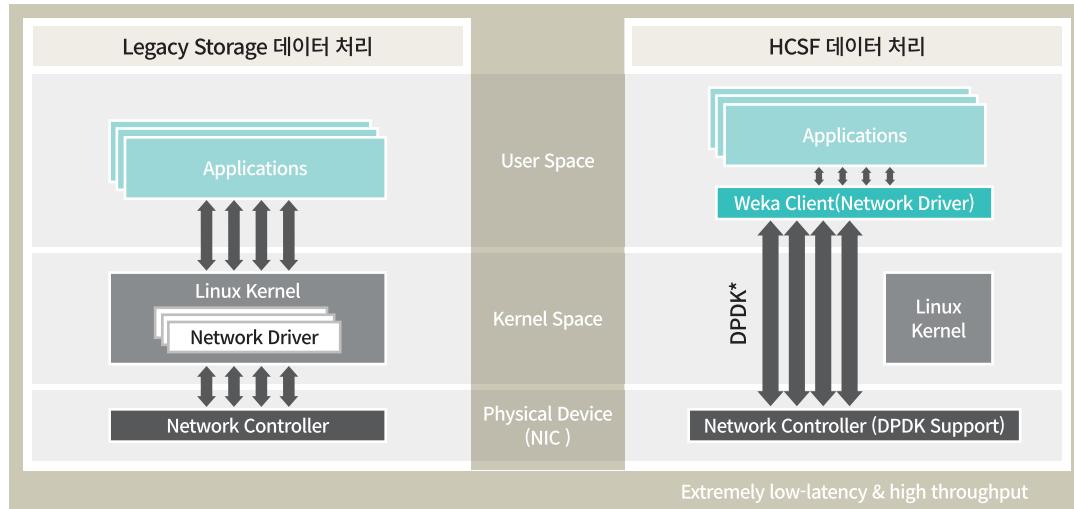
### HCSF는 Throughput, IOPS, Metadata, 그리고 Latency 부분에서 최고의 성능을 제공





### 고속 처리 기술 : DDPK

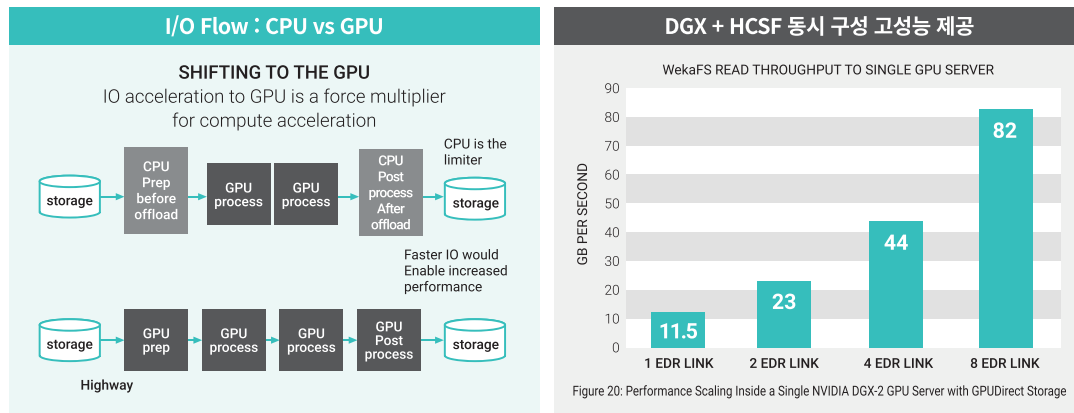
- HCSF는 Linux 커널을 거치지 않고 NIC와 직접 통신하는 기술을 채택하여 기존 데이터 처리 방식의 한계를 벗어나 Low Latency와 최고의 Throughput을 제공하며 Container 환경에도 적합한 고성능 스토리지 인프라 제공 가능



DDPK\* Data Plane Development Kit 의 약어로서, 네트워크 패킷 고속 처리 기술

### 고속 처리 기술 : GDS\*(GPU Direct Storage)

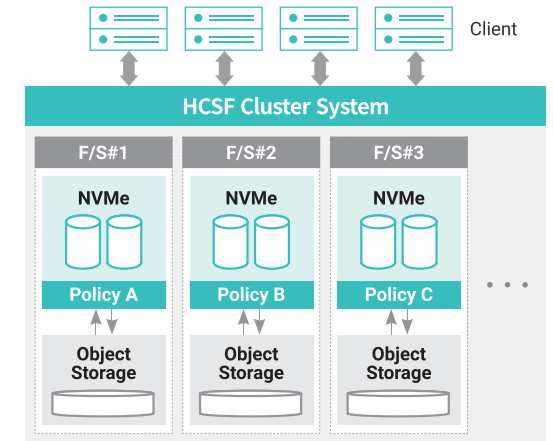
- GPU 코어와 스토리지를 직접 연결하는 GPU Direct Storage 기술을 적용하여 보다 빠르고 많은 대역폭의 I/O 처리 가능
- HPC AI/ML 분석 환경에서 GPU 서버는 필수 사항이며, 현재 GPU 서버 기술은 Nvidia사에서 주도하고 있음



GDS\* NVIDIA GPU Direct Storage 의 약어. GPU - Storage 병목 구간 고속 처리 기술  
 EDR 인피니밴드 100Gb/s (InfiniBand speeds: FDR / EDR / HDR)  
 NVIDIA DGX Nvidia DGX는 GPGPU(General-Purpose computing on Graphics Processing Units)를 사용하여 딥 러닝 애플리케이션을 가속화하는 데 특화된 Nvidia 서버

### 정책기반 Tiered 파일시스템

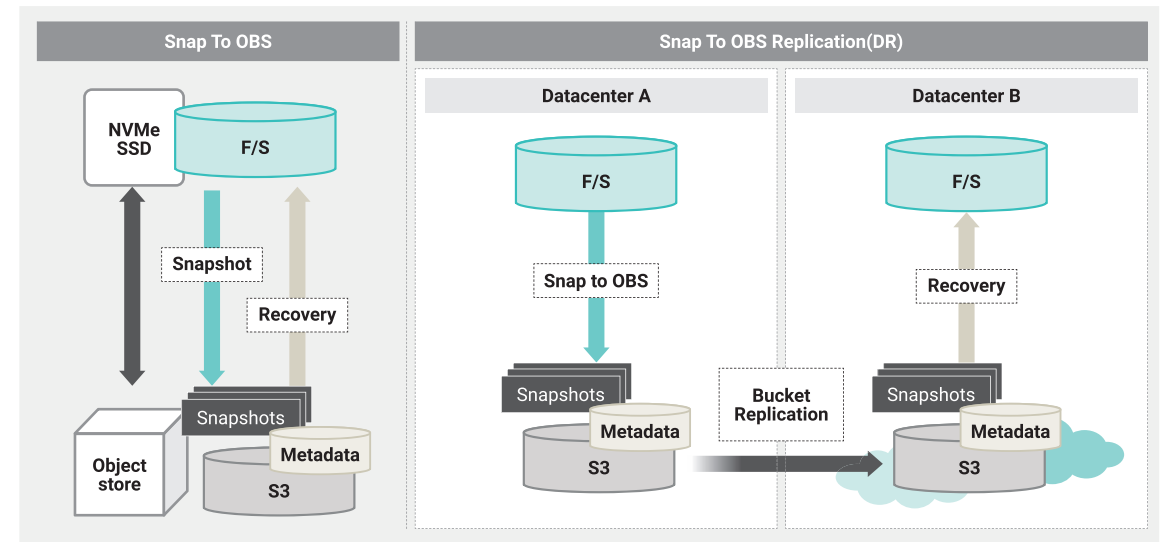
- 정책 기반 Tiering 파일시스템 운영으로 비용 대비 높은 성능, 용량 구성 효율성 제공
- Client는 Tier 관련없이 동일한 파일시스템 접근 사용
- 메타데이터 기반 Tiering Policy 운영
  - 최신 데이터는 HOT Tier(Tier1)에 저장/운영
  - 정책에 의해 Cold data는 오브젝트 스토리지 (Object Storage)(Tier2)로 Auto Tiering : Hot Tier 사용량에 따른 Tiering 정책 운영 가능하며, Hot Tier 사용량 임계치 도달 시 최근 사용되지 않은 데이터부터 오브젝트 스토리지로 이동
- 각 Tier 운영에 대한 데이터 관리, 모니터링 지원



### 시스템 복구/DR을 위한 Snap to OBS 기능

#### 스냅샷(Snapshot)

- Snap to OBS Replication
  - 오브젝트 스토리지에 저장된 스냅샷을 원격지로 복제하여 HCSF 전체 시스템에 대해 DR 시스템 구현 가능
  - 오브젝트 스토리지 복제 기능 활용
  - DR S3 저장소는 상용 S3 오브젝트 스토리지 뿐만 아니라, Public Cloud S3 저장소도 활용 가능
  - 노후화된 시스템에서 신규 시스템으로 데이터 마이그레이션 활용 가능



# Hitachi Content Platform

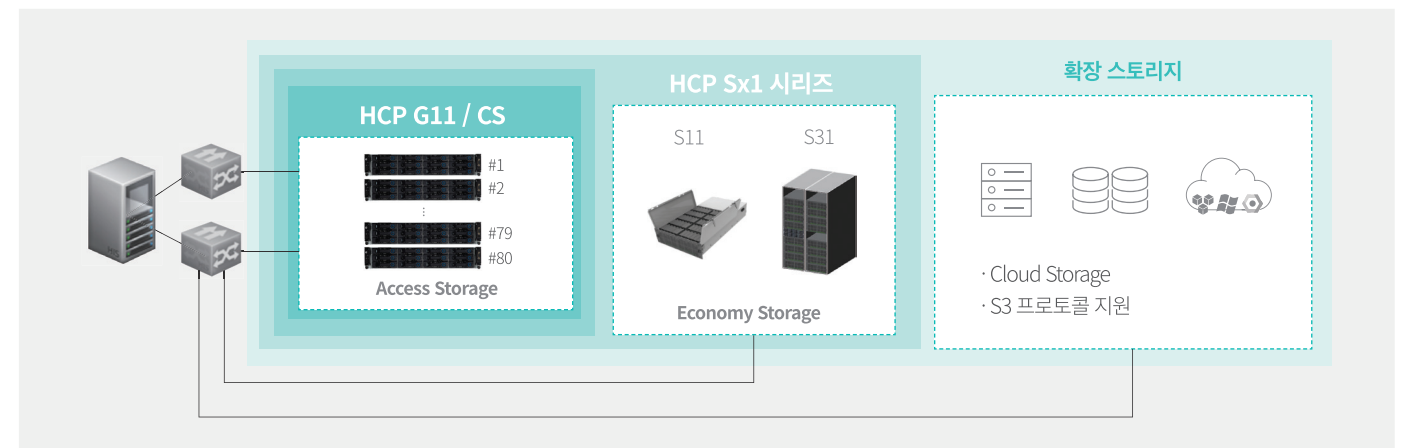
정책기반 Tiered 파일시스템,  
시스템 복구/DR을 위한 Snap to OBS 기능

폭발적으로 쏟아지는 데이터들 그리고 전통적인  
시스템에서 클라우드로 빠르게 바뀌고 있는 IT 트렌드

이러한 변화의 흐름에 신속하게 대응하고 새로운  
부가가치를 창출하는 것이 현 기업들이 당면한 과제입니다

Hitachi Content Platform(이하 HCP)은 기업 내 수 많은 데이터를 효율적으로 저장 및 관리하고, 데이터 분류, 검색, 빅데이터 분석 등 다양한 활용성을 제공함으로써, 급속도로 변화하는 IT 트렌드 흐름에 맞추어 데이터 기반의 인사이트를 도출할 수 있도록 지원합니다. 나아가 기존 IT 센터의 경계를 넘어 클라우드 기술과 연결하는, 유연하고 지능적인 차세대 클라우드 오브젝트 스토리지 플랫폼입니다.

## Architecture



### 데이터 서비스 관점

#### 클라우드 오브젝트 스토리지

- 가상화 및 멀티테넌시 구성
- 다양한 플랫폼을 수용할 수 있는 API 지원(REST, S3 등)
- 퍼블릭 클라우드와 연계한 하이브리드 인프라 구현

#### 비용 절감

- 원격지 분산 저장 아키텍처를 통한 TCO 절감 (Geo-EC)
- 압축 및 중복제거를 통한 저장공간 절약
- 고집적 저비용 저장장치를 통한 비용효율적인 구성

### 데이터 보호 관점

#### 안전한 데이터 보호

- 데이터 암호화 및 접근제어
- RAID + EC + SW미러(DPL)
- 파일 위변조 방지 (WORM)
- 다양한 원격 복제 토폴로지

#### 다양한 데이터 복구

- 버전별 데이터 보관 및 복구
- 데이터 자가 에러 체크 및 복구

### 데이터 인텔리전스 관점

#### 대규모 비정형 데이터 관리 & 응용

- 데이터에 커스텀 메타데이터를 추가하여 원하는 정보 삽입
- 메타데이터 인덱스 관리를 통한 빠른 검색
- 안정적인 성능으로 대용량 파일 처리

#### 데이터 활용을 위한 데이터 레이크

- 단일 시스템 내 1,000억개 이상 파일 저장 가능
- 단일 시스템 내 1,000 PB 이상 확장성 제공
- 대용량 저장장치로써 빅데이터 분석 및 IoT 활용 기반 조성

### 데이터 관리자동화

#### 데이터 가치에 따른 정책 적용

- 파일에 정책을 부여하여, 압축, 중복제거, 티어링, 자동삭제 등 데이터 관리 자동화
- 각 영역별 정책 부여 및 관리 가능

#### 비정형 데이터의 효율적 관리

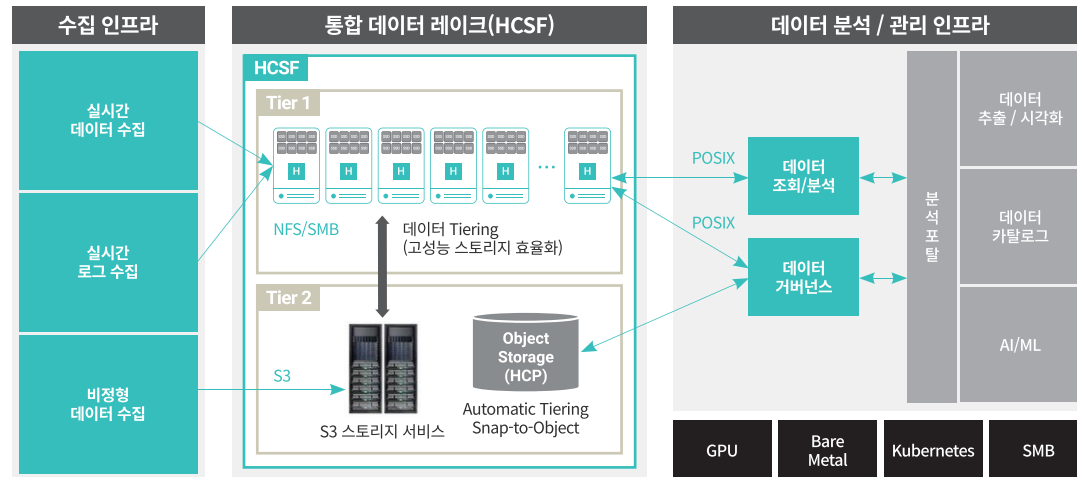
- 단일 시스템 내에서 업무 성격, 데이터 가치(사용빈도), 성능 등에 따라 저장매체 및 주요 정책 차등 적용

구분	HCP G11	HCP CS	HCP S11	HCP S31
노드 타입	Access Storage Node		Economy Storage Node	
지원 디스크	노드 당 최대 4TB HDD * 12개 최대 1.9TB/3.8TB SSD * 12개	Software Only (Docker 기반)	노드 당 10TB/14TB/16TB HDD * 200 개	노드 당 10TB/14TB/16TB HDD * 942 개
	RAID 6	-	Erasure Coding	
프로토콜	REST, HTTP(S), Amazon S3, NFS, CIFS, WebDAV, SMTP, NDMP	S3	-	-
클라우드 지원	Amazon S3, Microsoft Azure, Google Cloud	-	-	-
스토리지 옵션	HCP G11 internal, HCP S series, SAN, NFS, Cloud	HCP S Series	-	-
노드 확장성	4 - 80	Unlimited	1 - 80	
노드 당 Expansion Tray 확장성	-	-	1 Expansion	8 Expansion
최대 용량	노드 당 48TB	-	3.2PB	15PB
하드웨어 Rack Unit	노드 당 2RU	-	<b>(min)</b> 5RU <b>(max)</b> 9RU	<b>(min)</b> 5RU <b>(max)</b> 37RU
CPU & 메모리	노드 당 2 x 10 Core  64 - 768 GB	<b>(min)</b> 10 core <b>(권장)</b> 20 core	2 x 8 Cores	4 x 8 Cores
		<b>(min)</b> 128 GB <b>(권장)</b> 256 GB	128GB	512GB
SSD (내부 DB용)	2 x 1.9 TB	-	6 x 400GB	6 x 400GB
Networking	4x 10GbE SFP+ / 4x10GbE Base-T	-	8x 10GbE SFP+ / 8x 10GbE Base-T	8x 10GbE SFP+ / 8x 10GbE Base-T

### HCSF 구축사례

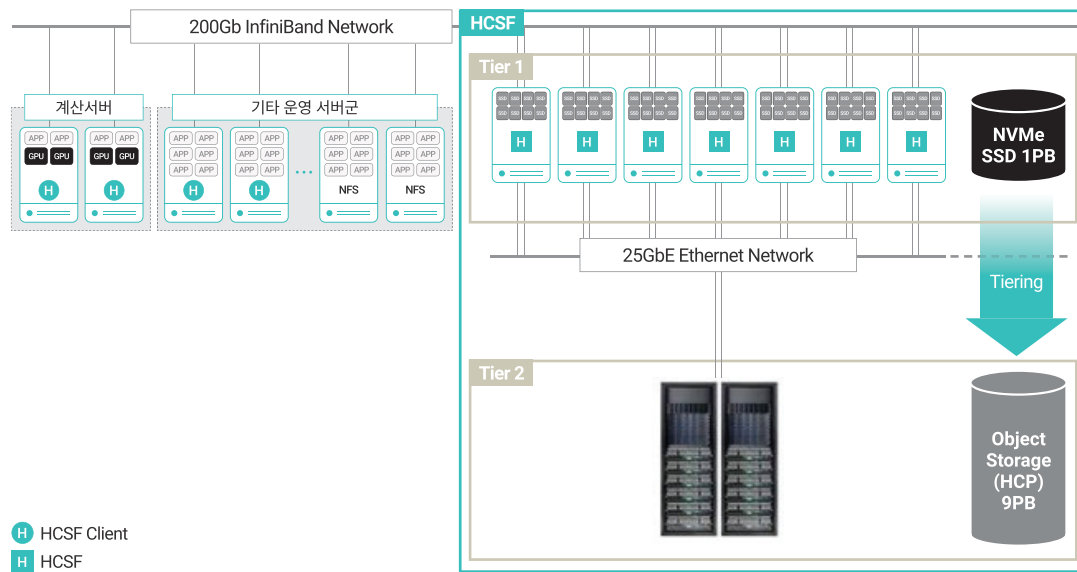
국내 제조기업 사례 - 멀티 프로토콜을 활용한 데이터레이크 스토리지

- 고객의 차세대 전자 데이터 분석 체계의 통합 저장소 구축 (초기 1xPB 규모)
- 고성능 데이터 분석 기반과 각기 다른 데이터 인터페이스 클라이언트 환경을 지원하는 통합 스토리지
- 오브젝트 스토리지는 HCSF 티어링 용도와 오브젝트 스토리지 서비스 용도 두 가지로 나누어 동시에 운영 중



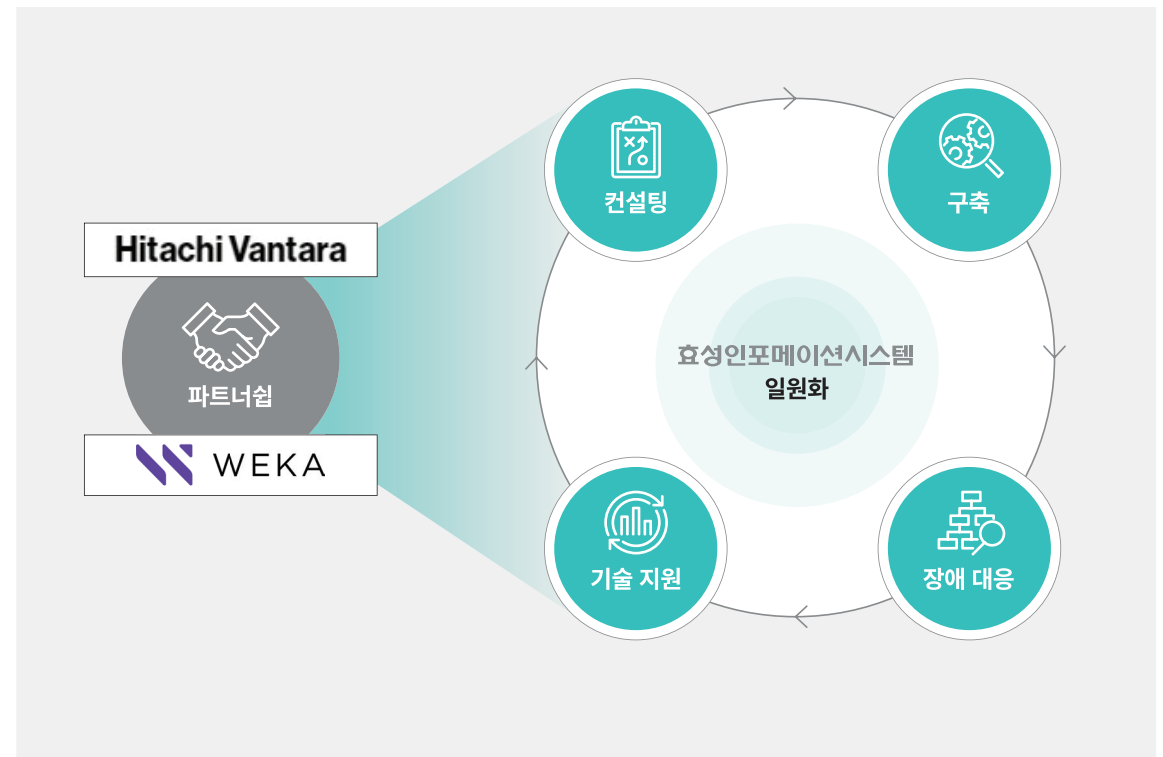
국내 공공 AI 연구센터 사례 - GDS 및 k8s CSI를 활용한 고성능 분석 스토리지

- 공용 인프라 저장소 HCSF 스토리지 (초기 1xPB 규모)
- GPU 서버로 구성된 컴퓨팅 자원과 고성능 스토리지를 연결한 AI 센터 통합 아키텍처 구축
- 서비스 네트워크는 200G InfiniBand HDR 인터페이스로 구성되어 분석데이터 고속 처리

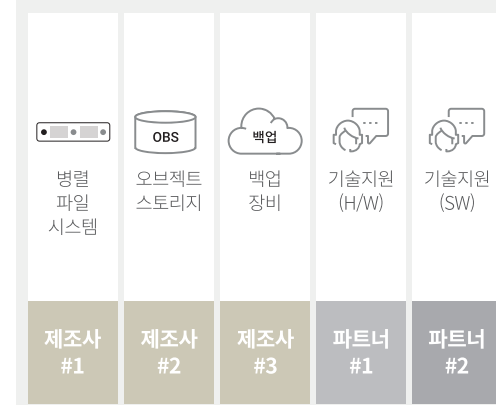


### A to Z 일원화 기술지원 제공

효성인포메이션시스템은 고성능 스토리지 시장에서 최고의 제품 구축 경험과 기술을 가지고 있으며, 컨설팅, 구축, 기술지원, 장애 대응 등 구축에 필요한 모든 요소를 A to Z 까지 일원화하여 지원



### 일반 병렬 파일시스템



VS 기술지원

### HCSF



- 효성인포메이션시스템은 히타치벤타라의 모든 H/W와 S/W 제품을 통합하여 제공

2023년 9월  
www.his21.co.kr

본 카탈로그에 수록된 솔루션 사양은 인쇄일을 기준으로  
사전 고지 없이 변경될 수 있으며, 최신 사양은 당사 영업대표 또는  
홈페이지를 통해 확인하시기 바랍니다.  
솔루션 관련 문의는 홈페이지의 <제품문의>를 통해 연락 부탁드립니다.

본사	서울특별시 강남구 도산대로 524 청담빌딩 5층	TEL 02-510-0300	FAX 02-547-9998
부산사무소	부산광역시 해운대구 센텀서로 30 KNN 타워 1303호	TEL 051-784-7811, 7813	FAX 051-463-7805
대구사무소	대구광역시 동구 회랑로 47 (신천동, 전문건설회관 3층)	TEL 053-426-9800	FAX 053-426-9830
서부사무소	대전광역시 서구 둔산서로 59 고운손빌딩 702호	TEL 042-485-4856	FAX 042-484-0366
광주사무소	광주광역시 서구 상무연하로 112 제갈량비즈타워 3층	TEL 062-385-2193	FAX 062-385-2194
수원사무소	경기도 수원시 영통구 삼성로 182-1 R7빌딩 3층	TEL 031-216-8717-8	FAX 031-216-8719