



© Copyright IBM Corporation 2021

한국아이비엠주식회사

(07326) 서울시 영등포구 국제금융로10

서울국제금융센터 (Three IFC)

TEL.: (02) 3781-7114

www.ibm.com/kr

2021년 5월

Printed in Korea

All Rights Reserved

이 문서는 미국에서 제공되는 제품 및 서비스용으로 작성된 것입니다. IBM은 본 문서에서 검토한 제품이나 기능, 서비스를 다른 국가에서는 제공하지 않을 수도 있습니다. 본 정보는 통지 없이 변경될 수 있습니다. 귀하가 사용할 수 있는 제품, 기능 및 서비스에 대한 정보는 한국 IBM 비즈니스 담당자에게 문의하십시오.

IBM이 제시하는 방향 및 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다. 이는 SOD에 의해 명시됩니다.

IBM, IBM 로고, Active Memory, AIX, BladeCenter, EnergyScale, Power, POWER, POWER8, PowerHA, PowerVM, POWER SYSTEMS 및 POWER SYSTEMS Software는 미국 또는 기타 국가에서 사용되는 International Business Machines Corporation의 상표 또는 등록상표입니다. IBM이 소유한 미국 상표의 전체 목록은 다음 사이트에 있습니다.
ibm.com/legal/copytrade.shtml

고성능 AI 및 하이브리드 클라우드 환경을 위한 차세대 스토리지

IBM ESS 3200



IBM Spectrum Scale



업계 최고 성능

+80 GB/s
NVMe,
PCIe G4 & HDR

검증된 솔루션

+20 years
Evolution &
Verification

#1 in HPC
NVIDIA POD RA
GDS(GPU Direct Storage)
Support

AI & 빅데이터 최적화

+170 PB
Various Use Cases
Professional Support

축적된 노하우

업계 최고 성능

AI & 빅데이터 워크로드에 최적화된 IBM ESS 3200 스토리지는 최신의 하드웨어 기술인 올플래시 NVMe, PCIe Gen4, HDR 그리고 IBM Spectrum Scale 솔루션의 탑재로 업계 최고의 성능을 제공하는 어플라이언스 스토리지입니다.

New ESS Gen3 Model
IBM ESS 3200



- ▶ 최신의 AMD CPU, PCIe v4, HDR IB
- ▶ 12 또는 24개 NVMe 플래시 드라이브 장착
- ▶ **80 GB/s per 2U** 시스템, 46TB ~ 368TB raw 용량

구분	ESS 3200 스토리지 상세 사양
파일시스템 / OS	• IBM Spectrum Scale Version 5 / Redhat Enterprise Linux 8
데이터 보호	• De-Clustered RAID(Erasure coding구조를 지원) : 3-way/4-way replication, 4+2P, 4+3P, 8+2P, 8+3P
컨트롤러	• 듀얼 active-active 컨트롤러/2 x AMD EPYC 7552 48 core Processor
드라이브 지원	• 12개 또는 24개 NVMe 플래시 드라이브 (3.84TB, 7.68TB, 15.4TB)
메모리	• 1 TB (컨트롤러 당 512 GB)
네트워크(어댑터)	• PCIe v4 HDR Infiniband 최대 8 port (HDR Card 4개) • 100G Ethernet 최대 8 port (100G Ethernet Card 4개), ROCE 지원
프로토콜	• POSIX, GPFS, NFS v4.0, SMB v3.0, Hadoop MapReduce, OpenStack Cinder (block) • OpenStack Swift (object), S3 (object) and CSI (Container Storage Interface)
성능	• Sequential Read /Write: 80GB/s
전원/공조	• 2 Power Supply (Hot swappable) • Input Voltage : 250V / 10A, Nominal Power : 1,350 W (Drive 미장착) ; 2,000 W(24개 Drive 장착) • Nominal Heat : 4,606 BTU/hr (Drive 미장착) ; 6,825 BTU/hr (24개 Drive 장착)
크기/무게	• 2 Rack Unit (H: 88 mm, W: 483 mm, D: 850 mm), 19인치 표준 Rack 장착 • 무게 : 38.5 kg (Drive 미장착) ; 46.6 kg (24개 Drive 장착)

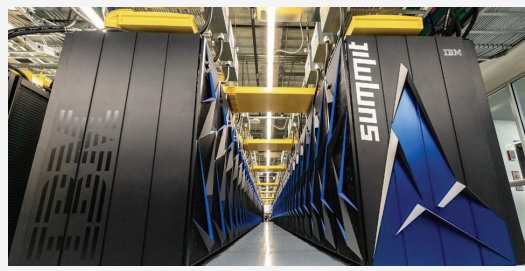


nvm

EXPRESS

IBM ESS 3200

“고성능 위한 하드웨어 컴포넌트 + 병렬 IO 처리 솔루션”



TOP 500

The List.

IBM ESS 로 구성된 Summit과 Sierra는 HPC 환경에 최고의 성능을 검증

4 | IBM ESS 3200

검증된 솔루션

IBM ESS 3200은 최신의 하드웨어와 함께 검증된 병렬 파일시스템 솔루션(Spectrum Scale)의 탑재로 다양한 기능과 성능 및 데이터의 안정성까지 확보된 어플라이언스 스토리지입니다.



Leader File/Object MQ : 2020-2016
#1 Ranked HPC and Analytics : 2020 - 2019



Leader Object-Based Storage MarketScape : 2019, 2018, 2016
Leader Scale-Out File-Based Storage MarketScape : 2019



40 Coolest Software-Defined Storage Vendors IBM : 2020



Leader Map for File Storage : 2020, 2019
Leader Map for Object Storage : 2019

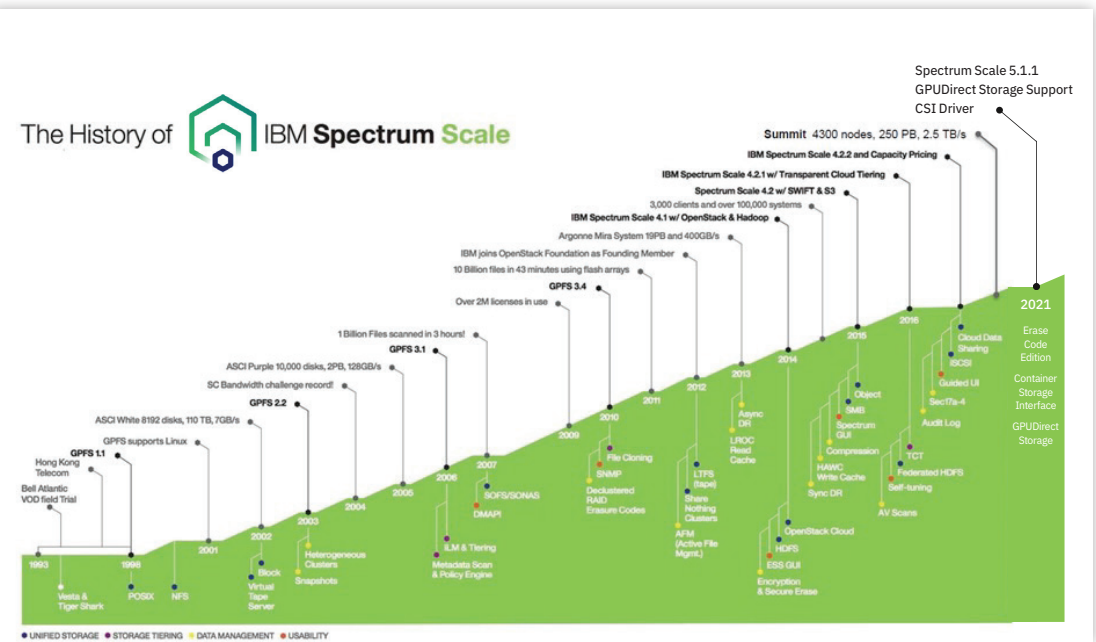


Best Distributed File and Object Storage : 2021



Winner Software-Defined Storage Vendor of the Year IBM : 2019

The History of IBM Spectrum Scale



- ▶ 20년 이상 지속적으로 최신 트렌드 기술(Cloud, Object, Container, GPUDirect 등)이 적용 및 진화된 **솔루션**
- ▶ 전 세계 포춘 TOP 100 기업 중 87개 기업의 사용으로 입증된 안정적인 병렬 파일시스템 **솔루션**
- ▶ 다양한 검증기관에서 최고 리더십을 발휘하는 **솔루션**

IBM ESS 3200 | 5

AI & 빅데이터 최적화

IBM ESS 3200은 AI에 최적화된 어플라이언스로 강력한 데이터 처리 성능을 슈퍼컴퓨터 Top 100의 Summit과 Sierra를 통해 검증받은 스토리지입니다.
또한 NVIDIA DGX GPU 시스템과의 검증된 파트너십으로 최고의 AI 인프라를 제공합니다.



✓ HPC TOP 3 중 2대의 슈퍼컴퓨터 (Summit & Sierra)를 통해 검증된 데이터 처리 성능

* 400PB IBM ESS + 44,000 NVIDIA GPUs

✓ AI 워크로드를 위한 데이터 Feeding을 가속화하여 GPU 활용률을 더욱 향상

✓ ESS 3200의 HDR 지원으로 대용량의 빅데이터 처리 성능 향상

IBM Spectrum Scale

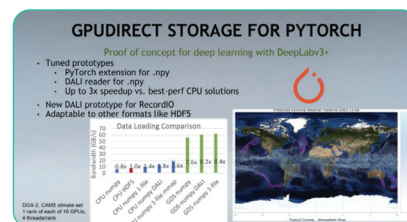


✓ AI 성능 벤치마크가 게시된 NVIDIA DGX SuperPOD 레퍼런스 아키텍처*

* #23 in HPC TOP500, 9.5 PFLOPs, 144 PFLOPs via Tensor Cores

✓ GPUDirect Storage의 지원으로 평균 3배 이상 성능 향상

* GDS for pytorch with DeepLabv3+



* <https://developer.nvidia.com/blog/dgx-superpod-world-record-supercomputing-enterprise/>

축적된 노하우

IBM은 2020년 이후 ESS(Elastic Storage System)를 통해 +170PB, 50여 국내 고객사에 AI & 빅데이터 환경 구축 레퍼런스를 확보하였으며 대표적으로 우주전파연구소, 성균관대학교, GC녹십자, 아프리카TV 등이 있습니다.

체크포인트 #1: 다양한 구축 사례 확보

HPC, 온라인 강의, 유전체 분석, 실시간 스트리밍 등 수십 PB 이상의 대용량과 수십 GB/s 이상의 고성능 업무에 최적의 인프라로 채택되었으며, ML/DL과 제조, 공공, 금융 분야에서도 다양한 국내, 외 레퍼런스를 확보



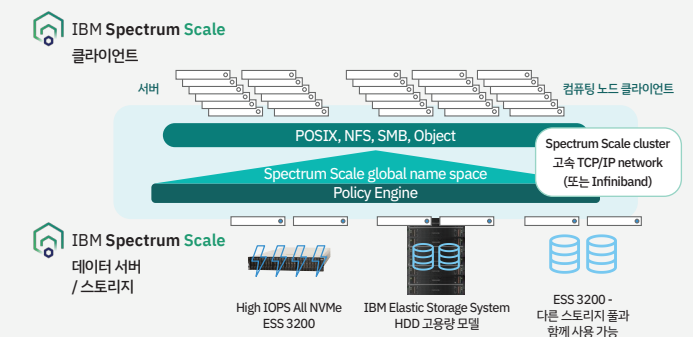
체크포인트 #2: 전문 기술 지원 조직 확보

- 수많은 경험을 통한 컨설팅 / 기술 지원 역량 확보
- 글로벌 Hot-line 지원 조직과 활발한 기술 정보 공유
- IBM Korea ESS 전문 Lab Service 지원팀 구성
- ESS 전문 파트너(챔피언 클럽) 커뮤니티 활성화



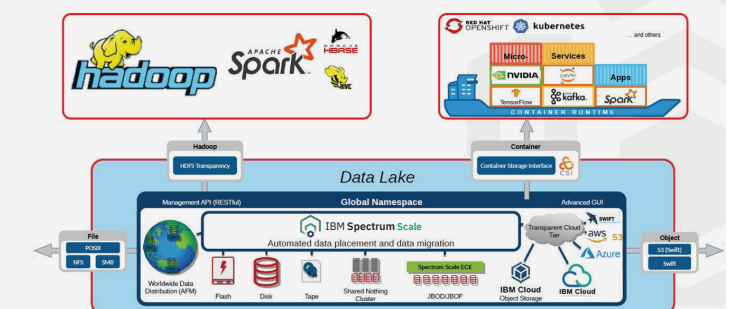
Key Use Case #1:

- (대용량 파일시스템 구성) 기존 병렬파일시스템 클러스터와 NVMe 기반 고성능 어플라이언스 ESS 3200의 확장 및 DISK 기반 어플라이언스와의 연계
- ILM 활용으로 데이터 Tier 구성



Key Use Case #2:

AI 개발 환경에 적합한 최신 컨테이너 기술과의 연동(Container Storage Interface 지원) 및 데이터 레이크 환경 구축





IBM Elastic Storage System 3000
고객 사례

GC녹십자지놈 GC녹십자지놈

스토리지 인프라 고도화로 유전체 분석 시스템 혁신

2013년 설립된 GC녹십자지놈은 GC녹십자의 가족사로 임상 유전체 분석에 특화된 기업입니다. GC녹십자지놈은 질병 진단, 예측, 맞춤형 치료를 위해 유전체 분석 정보를 활용하고 있으며, 연구 목적 기관과 달리 병원의 임상 데이터 분석에 집중하면서 임상 유전체 분석 분야를 선도하고 있습니다. 또한 GC녹십자 지놈은 차세대 염기서열 분석기(NGS) 등 첨단 장비를 활용해 다양한 프로젝트에 참여하고 있는데, K-DNA 프로젝트의 일환으로 추진되는 1차 시범 사업 수주를 계기로 첨단 분석을 위한 인프라 고도화에 나섰습니다. 그 첫 시작으로 IBM ESS를 중심으로 올플래시, 디스크, 테이프로 스토리지 계층화 기반을 마련했습니다.

고객 상황

GC녹십자지놈은 시장의 요구에 맞는 유전체 분석 서비스를 제공합니다. 이를 위해 실제 협력 병원에서 수집한 임상 데이터를 사용해 분석을 수행하고, 이를 바탕으로 시장이 요구하는 다양한 유전체 분석 검사 항목을 출시하고 있습니다. 더불어 중요 국가 과제에도 적극적으로 참여하고 있습니다. 관련해 GC녹십자지놈은 K-DNA 사업에 앞서 시행되는 한국인 5천 명 유전체 분석 시범 사업을 수주하였습니다. 이 사업은 수주 여부를 떠나 여러모로 GC녹십자지놈에게 의미가 큼니다. 본격적인 바이오 시장 성장이 곧 시작된다는 신호탄이기 때문입니다.

유전체 분석 분야는 거대한 시장 기회를 맞이하고 있습니다. 빅데이터3법 시행으로 규제 수위가 낮아져 바이오 빅 데이터의 비즈니스 가치가 커질 것이란 기대와 함께 디지털 뉴딜 정책의 일환으로 바이오 빅 데이터 구축이 본격화되는 시기가 성큼 다가오고 있습니다. 이런 분위기 속에서 K-DNA의 일환으로 추진되는 시범 사업 수주는 앞으로 더욱 커질 시장 기회에 앞서 무엇을 개선하고 강화해야 하는지를 알려주는 일종의 나침반과 같았습니다.

GC녹십자지놈의 눈에 들어온 개선 대상은 바로 IT 인프라 운영

전략이었습니다. GC녹십자지놈의 IT 환경은 다른 유전체분석 기관과 크게 다르지 않았습니다. 일루미나 시퀀싱 장비에서 나온 데이터를 HPC 인프라를 통해 분석하는 환경이었습니다. 고성능 HPC와 스토리지를 핵심 인프라 요소로 관리하던 것이 기존 전략이었습니다. GC녹십자지놈은 유전체 분석 시장 성장세에 맞춰 가려면 IT 인프라가 아니라 데이터 센터라는 더 큰 개념을 전제로 전략을 수립해야 한다고 판단하고 자체 데이터센터를 구축하였습니다. 이 과정에서 기존 사업과 과제를 병행하다 보니 확장 요구가 있었고, 이에 대응하는 가운데 예상치 못한 문제가 불거졌습니다. 바로 스토리지 I/O 병목이 데이터센터 확장의 발목을 잡았던 것입니다.



솔루션

GC녹십자지놈은 차세대 데이터센터 전략에 맞는 스토리지 인프라의 중요성을 체감하고 곧 새로운 솔루션 도입 검토에 나섰습니다. 이에 고려하기 시작한 솔루션이 분산 파일 시스템이었습니다. 초기에는 Lustre와 같은 오픈소스를 포함해 여러 솔루션을 살펴보았지만, 핵심 역량이 필요한 주요 업무에 고성능과 안정성이 절대적으로 요구되었던 GC녹십자지놈은 IBM ESS 스토리지 도입을 결정했습니다.

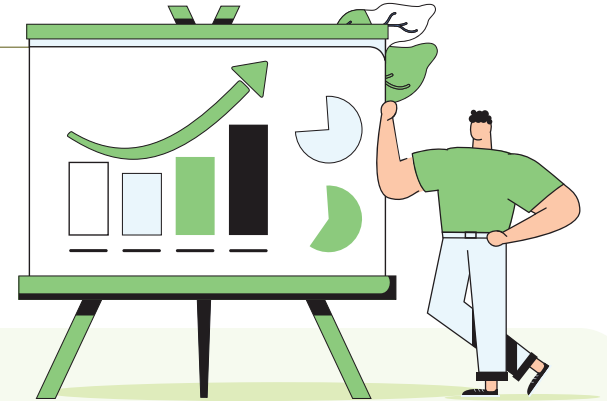
오픈 소스 기술인 Lustre를 범용 스토리지에 적용하는 것은 아무래도 안정적인 기술 지원에 대한 우려가 있었습니다. 이와 달리 IBM ESS 스토리지는 시장에서 충분히 검증된 장비에 IBM Spectrum Scale(이전 General Parallel File System(GPFS))의 이점까지 누릴 수 있었습니다. 여기에 성능도 GC녹십자지놈이 원하는 수준 이상을 보여 주었습니다. 관련해 GC녹십자지놈 이태현 박사는 “Lustre 같은 환경이 필요했는데, 검토한 결과 Spectrum Scale이 우리 기준에 잘 맞았습니다.”라며 “성능의 경우 40~50Gbps 수준의 I/O가 나오길 원했는데, IBM ESS는 기대를 넘어서는 성능을 보여 주었습니다.”라고 말했습니다.

한편 GC녹십자지놈은 IBM ESS GL2S와 ESS 3000 스토리지를 들여오면서 스토리지 계층화 인프라를 완성하였습니다. 바로 IBM ESS 스토리지, TS4500 테이프 라이브러리, Spectrum Archive 솔루션을 통하여 핫 데이터, 콜드 데이터에 대한 장기 아카이빙을 위한 체계를 정립한 것입니다. 이에 따라 GC녹십자지놈은 높은 성능이 요구되는 유전체 데이터 분석부터 안정적으로 장기 보관이 필요한 아카이빙 데이터까지 일관성을 갖고 ILM(Information Lifecycle Management)을 수행할 기초를 마련하였습니다. 관련해 GC녹십자지놈 이태현 박사는 “성능이 빠른 스토리지 도입이 원래 목표였지만, IBM ESS를 도입하고 나서 보니 성능을 넘어 스토리지 계층화를 통한 운영 효율성까지 높이는 효과를 거두었습니다.”라고 말했습니다.

그리고 GC녹십자지놈은 IBM ESS 스토리지 도입을 통해 단일 네임스페이스 상에서 초고속 데이터 분석 스토리지를 확장해 나갈 수 있는 유연성도 확보하였습니다. 이로써 새로운 서비스나 공공 프로젝트 추진 등에 따른 성능 확보가 요구될 때 신속한 대응이 가능해졌습니다.



도입효과



GC녹십자지놈은 스토리지 인프라 고도화를 통해 유전체 분석 역량을 한층 더 강화하였습니다. 기존 스토리지와 비교할 때 단일 분석 노드당 I/O 성능은 3배 이상 높아졌습니다. 스토리지 I/O 성능 향상은 분석 서버의 자원 효율성도 높였습니다. 기존에는 스토리지 병목으로 서버 CPU 자원이 제 성능을 발휘하지 못했습니다. 평균 CPU 사용률이 40~50%에 그쳤을 정도입니다. 그러던 것이 IBM ESS 스토리지 도입 후부터는 CPU 사용률이 크게 높아졌습니다. 자연스럽게 유전체 분석 워크플로우는 더욱더 빨라졌습니다. GC녹십자지놈은 워크플로우 가속을 위해 향후 유전체 분석에 최적화된 FPGA 기반 하드웨어 도입도 검토 중입니다.

한편, 이번 스토리지 인프라 고도화는 GC녹십자지놈이 중장기 목표 아래 데이터 자산화와 데이터 거버넌스 체계 구축의 기초를 다지는 효과까지 이어집니다. GC녹십자지놈 이태현 박사는 “우리가 다루는 유전체 데이터는 양이 큰 편입니다. 한 사람당 1TB 정도 데이터가 나오는데, 일년에 15만 건 정도를 분석합니다. 이들 데이터는 모두 보관을 해야 합니다. 따라서 분석 단계에서는 더 빠른 성능의 스토리지가 필요하고, 대용량 데이터를 경제적이고 안정적으로 장기 보관할 수 있는 장치도 있어야 합니다. IBM ESS 스토리지, TS4500 테이프 라이브러리, Spectrum Archive 솔루션을 통해 우리는 성능과 용량 모든 측면에서 스케일아웃이 유리한 최신 스토리지 인프라를 갖출 수 있게 되었습니다” 라고 말했습니다.



라이브 방송과 VOD 서비스 역량 강화 위해 스토리지 세대교체

아프리카TV는 ‘내가 주인인 미디어’를 누구나 쉽게 실시간 방송을 할 수 있는 유무선 통합 소셜 미디어 플랫폼입니다. 실시간 방송으로 1인 미디어 시대를 열었고, 2017년 VOD 서비스를 시작하여 동영상 플랫폼 서비스를 강화하며 견조한 성장을 이어가고 있습니다. 이런 성장세에 발맞춰 아프리카TV는 2020년 스토리지 인프라 교체를 추진했습니다. 이번 프로젝트는 VOD 서비스 개시 후 급증하는 스토리지 수요와 데이터 관리에 대한 근본적인 해결책을 찾는 것이 목표였습니다. 아프리카TV는 IBM Elastic Storage System(이하 ESS)에서 답을 찾았고, 그 결과 실시간 방송과 VOD 수요가 빠르게 늘어난 비대면 시대에도 스토리지 용량과 성능 걱정 없이 안정적인 서비스를 제공하고 있습니다.

고객 상황

아프리카TV는 오픈 소스 기술을 매우 효율적으로 사용하는 곳으로 유명합니다. 대규모 분산 컴퓨팅 환경을 구축하고 오픈 소스 소프트웨어를 중심으로 서비스를 운영해왔습니다. 스토리지 인프라 역시 x86과 리눅스 기반 클러스터 상에 글러스터 파일시스템(Gluster Filesystem)을 적용해 분산 스토리지로 사용했습니다. 아프리카TV의 분산 스토리지 환경은 비용 효율성이 높았고, 확장도 유리했습니다. 다만, VOD 서비스 사업 헷수가 쌓이면서 저장 규모가 테라바이트 수준을 넘어서자 관리 부담이 커졌고, 향후 데이터 규모가 더 커질 경우 사용자 경험 저하로 이어질 잠재적 가능성이 있었습니다. 이에 아프리카TV는 단일 글로벌 네임스페이스 상에서 모든 스토리지 자원을 통합할 수 있는 고성능 파일시스템 찾아 나섰습니다.

아프리카TV는 PoC를 통해 주요 엔터프라이즈 스토리지 공급 기업의 장비 사전 검증을 진행했습니다. 이 과정에서 아프리카TV가 중점적으로 살핀 것은 단일 글로벌 네임스페이스 기반 관리, 스토리지 계층화, 생명 주기 기반 데이터 관리 그리고 무중단 이전이었습니다. 기술 검토를 통해 아프리카TV는 IBM ESS가 최선의 선택이란 결론을 내렸습니다. IBM이 제안한 고성능 글로벌 병렬 파일시스템인 스펙트럼 스케일(Spectrum Scale)은 아프리카TV의 요구 사항을 기대 이상으로 충족했습니다. 단일 글로벌 네임스페이스 지원의 경우 단일 마운트 포인트로 프로덕션부터 아카이빙 환경에 있는 데이터까지 모두 접근이 가능합니다. 또한, IBM은 다른 업체의 제안과 달리 ILM(Information Lifecycle Management) 기반으로 계층화된 스토리지를 관리할 방안도 제공합니다. 데이터센터 이전의 경우 IBM은 액티브/액티브 방식으로 서비스 중단 없이 신규 센터 오픈이 가능합니다.



솔루션

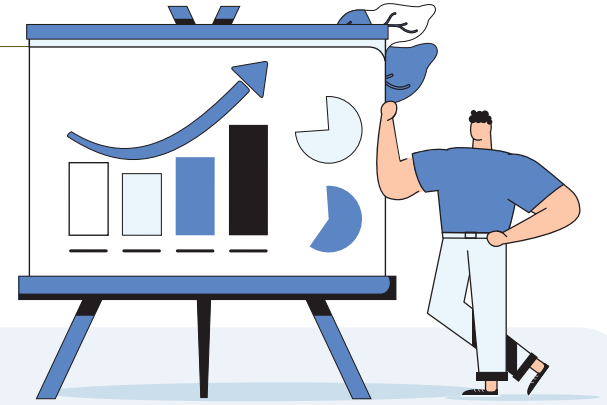
아프리카TV는 2020년 16페타바이트 규모의 신규 스토리지 인프라를 구축했습니다. 아프리카TV는 신규 스토리지 도입과 함께 스토리지 환경을 계층화했습니다. 신규로 도입하는 IBM ESS 3000과 GL5S 스토리지 어플라이언스는 빈번하게 사용되는 데이터 저장 위치에 배치했습니다. NVMe 기반 플래시 스토리지와 디스크 기반 스토리지를 프로덕션 환경에 두고 운영하는 가운데 장기 보관이 필요한 데이터는 기존에 사용하던 x86 기반 분산 스토리지를 재활용하는 쪽으로 방향을 잡았습니다. 참고로 기존 장비의 경우 글러스터 파일시스템을 유지해도 별도의 리스토어 작업 없이 단일 글로벌 네임스페이스 상에서 접근할 수 있어 재활용이 수월합니다.

한편, 아프리카TV는 스토리지 인프라를 구성하면서 IBM ESS의 높은 성능 이점을 체감했습니다. 작업 초기 아프리카TV는 캐싱 영역과 저장 영역으로 스토리지를 나누어 구성하는 것을 고려했습니다. 읽기와 쓰기 작업 분리를 고민한 이유는 아프리카TV 서비스 특성상 라이브 방송을 송출할 때 VOD 용도의 데이터가 동시에 저장되기 때문입니다. 하지만 우려와 달리 사전 테스트 결과 굳이 읽기와 쓰기를 따로 할 필요가 없었습니다. ESS 3000 성능이면 방송 송출과 저장을 동시에 처리해도 서비스 지연 걱정이 전혀 없다는 것을 테스트를 통해 알 수 있었기 때문입니다.

아프리카TV는 고성능 NVMe 기반 플래시 스토리지인 ESS 3000으로 방송 송출과 저장에 대한 읽기와 쓰기 작업을 모두 처리한 후 2시간 뒤에 디스크 기반 스토리지인 GL5S에 저장하도록 설정했습니다. 별도의 복제 작업을 할 필요가 없이 정책 기반으로 데이터를 처리하도록 설정한 것인데, 만약 다른 업체의 스토리지를 도입했다면 스토리지 간 데이터를 옮기는 번거로움이 있었을 것으로 아프리카TV는 보고 있습니다.



도입효과



아프리카TV는 IBM ESS 3000과 GL5S 도입으로 스토리지 관리 방식에 일대 혁신을 가져왔습니다. 가장 두드러진 변화는 관리 포인트입니다. 기존 글러스터 파일시스템 기반 분산 스토리지를 계속 운영했다면 데이터 증가에 따른 관리 포인트가 많아질 수밖에 없었을 것입니다. 특히 비대면 시대가 장기화되면서 실시간 방송과 VOD 수요가 많이 늘어난 것을 생각해 보면 IBM ESS 3000과 GL5S 도입이 없었더라면 관리 부담이 매우 크게 다가왔을 것입니다. 아프리카TV는 단일 네임스페이스 기반 관리 체계 정립으로 스토리지 관리 방식을 간소화했습니다.

두 번째 혁신은 사용자 경험입니다. 최신 스토리지와 고성능 병렬 파일 시스템의 조합은 실시간 방송과 VOD 관련 사용자 경험을 크게 높였습니다. 시청자의 경험 측면에서 스토리지 인프라 세대교체 후 지연 걱정 없는 안정성이 강화되었습니다. 기존 스토리지 환경은 순간적인 사용자 접속 폭증 시 서비스 지연 우려가 있었습니다. 반면에 신규 환경은 순간적인 접속 폭증 상황에도 스토리지 IO가 안정적으로 유지되어 서비스 지연 걱정이 없습니다. BJ 사용자 경험도 개선되었습니다. 기존에는 스토리지 문제로 VOD용 영상이 저장되지 않을 경우 BJ가 자신의 데스크톱에 저장한 데이터를 다시 올리는 수고를 해야 했지만, 이제는 그럴 일이 없습니다. 고성능 병렬 파일시스템 기반 스토리지 자원을 단일 글로벌 네임스페이스 상에서 관리하는 구조이기 때문입니다.

한편, 아프리카TV는 스토리지 교체로 무중단 데이터센터 이전을 위한 만반의 대비를 하였을 뿐만 아니라 자연재해나 사이버 공격으로 인한 시스템 중단 걱정도 크게 덜어내었습니다. 별도의 재해복구 솔루션 없이도 프로덕션 환경에 있는 데이터를 같은 용량의 백업 어플라이언스로 복제하는 것이 매우 간단해졌기 때문입니다.



온라인 강의 인프라 확충으로 코로나19에 선제적 대응

성균관대학교는 코로나19 사태 초기인 3월 이미 1학기 1주 차 수업 90% 이상의 녹화 작업을 마치는 등 발 빠르게 움직였습니다. 이후 코로나19 사태 장기화에 대비해 온라인 강의 콘텐츠 저장 공간 확대와 이용자 폭주에 따른 성능 병목 해결을 위해 고성능 올 플래시 스토리지인 IBM Elastic Storage System 3000을 도입했습니다. 서버, 회선 증설과 함께 최신 스토리지 도입으로 성균관대학교는 차질 없이 원격 수업 수요 폭증에 효과적으로 대응하였습니다.

고객 상황

성균관대학교는 코로나19 사태로 인해 2020년 초 1학기 전 강좌에 대해 온라인 강의 운영 체계로 전환했으며, 2학기에는 사회적 거리두기 단계에 따른 온오프라인 수업을 병행하였습니다. 성균관대학교는 코로나19 사태 이전에도 실시간 화상 수업 방식과 온라인 수업 플랫폼인 아이캠퍼스(i-Campus)를 병행 운영하며 언제, 어디서나 수업을 진행할 수 있는 환경을 제공해왔습니다.

이 중 아이캠퍼스는 사전 녹화 제작한 강의 영상을 CMS(Contents Management System)에 올려 수업을 진행하는 방식이며, 실시간 화상 수업 방식의 강의도 복습 가능하도록 녹화 영상을 제공합니다. 예전에도 온라인 수업을 활발히 이용하였지만, 코로나19로 인한 온라인 강의 운영 결정은 기존 아이캠퍼스 운영에 큰 영향을 끼쳤습니다. 성균관대학교는 일단 급한 대로 3월 초 1학기 1주 차 수업 90% 이상의 녹화를 마치고 온라인 수업에 대비하였습니다. 이를 진행하면서 다른 한편으로 아이캠퍼스 인프라 증설을 준비하였습니다.

성균관대학교는 서버, 회선을 증설하고 스토리지는 교체하는 쪽으로 인프라 투자 계획을 세웠습니다.

이중 가장 큰 규모로 진행된 투자가 스토리지 교체입니다.성균관대학교는 온라인 강의 동시 접속 처리의 경우 서버와 회선 증설로 어느 정도 처리가 가능하다고 판단했습니다. 하지만 대용량의 동영상 스트리밍 서비스가 필요한 온라인 강의에 대한 안정적인 품질을 유지하려면 스토리지 성능이 뒷받침되어야 한다고 보았습니다.

온라인 강의는 특정 시간대에 동시 접속이 몰려 큰 부하가 걸립니다. 여기에 교육 콘텐츠의 특성상 난해한 부분을 반복해 보는 등 교육 콘텐츠를 저장해 둔 스토리지 I/O 성능에 영향을 끼치는 특징적인 요소가 많습니다.



솔루션

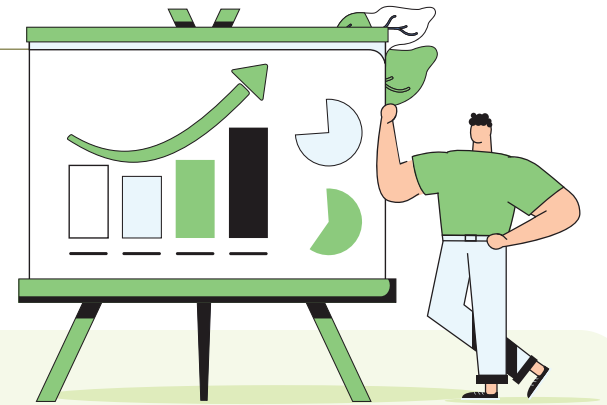
성균관대학교는 처음부터 스토리지 교체를 검토하지는 않았습니다. 기존에 사용하던 스토리지 용량 증설을 알아본 결과 예상보다 추가 비용이 너무 많이 들다 보니, 이 기회에 고성능 올플래시 스토리지로 바꾸는 쪽으로 계획을 변경하게 된 것입니다. 성균관대학교는 여러 시스템을 물망에 올리고 비용 대비 성능을 따져 보았습니다. 성능이 가장 중요한 평가 요소였지만, 가급적이면 용량 대비 비용 측면에서 더 경제적인 시스템을 찾고자 했습니다. 이 기준에 가장 잘 맞는 것은 IBM Elastic Storage System 3000(이하 IBM ESS 3000)이었습니다. 이와 관련해 성균관대학교 관계자는 “코로나19 대응 초기 기존 NAS 스토리지에 용량 20TB를 증설하는데 그 비용이 수천만 원에 이를 정도로 컸습니다”라며 “IBM ESS 3000을 보니 200TB 용량을 도입하는 데 드는 비용과 기존 시스템에 20TB를 증설하는 비용 차이가 생각보다 크지 않아 용량 증설 보다는 신규 장비 도입이 경제적이라고 판단했고 또한 코로나19 사태가 장기화될 것을 예상함에 따라 대용량 스토리지 공간이 필요하다고 판단해 최종적으로 스토리지 교체를 결정하게 되었습니다.”라고 말했습니다.

IBM ESS 3000은 AI와 데이터 분석에 특화된 올플래시 스토리지입니다. 높은 I/O 성능을 요구하는 AI, 대용량 데이터 처리 워크로드에 최적화된 스토리지로 2U 크기의 시스템당 40GB/s의 데이터 처리량 성능을 제공합니다. 성균관대학교가 IBM ESS 3000을 선택한 중요 이유 중 하나가 처리량입니다. 서버를 아무리 늘려도 스토리지의 데이터 처리량이 적으면 원활한 온라인 강의 서비스가 어렵습니다.

안정성 역시 높은 점수를 준 부분입니다. 교수진이 어렵게 만들어 올린 강의 영상이 시스템 장애나 해킹 등의 이유로 손실된다면? 이는 온라인 강의를 지속하는 데 있어 심각한 영향을 줄 수 있습니다. IBM ESS 3000은 소프트웨어 정의 이레이저 코딩(eraser coding)을 지원합니다. 이를 이용하면 데이터 복제 방식보다 적은 공간만 확보해도 데이터 보호 체계를 갖출 수 있습니다. 신속하고 편리한 복원도 지원하여 혹시 모를 장애 발생 시에도 며칠이 아니라 몇 분이면 서비스 재개가 가능합니다.



도입효과



IBM ESS 3000 도입으로 성균관대학교는 코로나19 사태 초기 수립한 온라인 강의 운영 계획을 차질 없이 실행하고 있습니다. 성균관대학교는 코로나19 상황이 해소될 때까지 온라인 강의를 운영할 계획이며 포스트 코로나 시대를 향하여 지속적으로 활용하게 될 것으로 예상하고 있습니다.

관련해 성균관대학교 관계자는 “코로나 19 사태를 맞이한 첫 학기는 온라인 강의 운영을 위해 실기 과목만 대면 강의를 하고, 실험과 실습 강의는 여름 학기에 집중 수업을 하는 등의 대책을 마련하였습니다.”라며 “코로나 19에 대응한 학사 운영 대책이 차질 없이 이루어지려면 온라인 강의 인프라가 탄탄히 뒤를 받쳐야 하는데, 서버와 회선 증설과 함께 IBM ESS 3000을 도입해 성능과 용량 걱정을 덜어냈습니다.”라고 말했습니다.

IBM ESS 3000 도입 후 가장 도드라진 변화는 성능입니다. 기존 SAN 스토리지 환경과 비교할 때 10배 이상의 성능 차이를 보이기 때문입니다. 이는 고스란히 온라인 강의에 대한 학생들의 사용자 경험으로 이어지고 있습니다. 동시 접속이 폭주하는 시간대에도 영상의 끊김이나 지연이 없었습니다. 더 놀라운 것은 1년 가까이 온라인 강의를 진행하는 가운데 스토리지에 부하가 거의 없다는 것입니다. 최대 부하가 20~30% 사이를 오갈 정도로 IBM ESS 3000은 넉넉한 성능을 보여주고 있습니다.

한편, 성균관대학교는 IBM ESS 3000이라는 고성능 스토리지를 통해 CDN 용량을 늘리지 않아도 되는 추가적인 효과도 거두었습니다. 보통 CDN 서버는 오리진 서버의 콘텐츠를 캐싱하여 사용자에게 제공합니다. 이를 통해 사용자가 체감하는 콘텐츠 이용 속도를 보장합니다. 고성능 스토리지를 배치하면 웹캐시에 저장되어 있지 않은 콘텐츠를 요청하더라도 원본 스토리지에서 콘텐츠를 검색하는 속도 자체가 매우 빠릅니다. 이런 이유로 성균관대학교는 CDN 용량 증설 없이 온라인 강의 운영을 이어가고 있습니다.