

서처 사이즈믹(Searcher Seismic)은 퓨어스토리지의 플래시블레이드(FlashBlade™)를 기반으로 데이터 허브를 구축함으로써 석유 가스 업계 기업 고객이 탐사 데이터에서 주요 인사이트를 더욱 신속하게 얻을 수 있도록 해주는 새로운 고부가가치 서비스를 제공할 수 있게 되었다.



비즈니스 혁신

고객들은 비즈니스에 매우 중요한 탄성파 탐사 자료를 몇 주가 아닌 단 몇 분만에 얻을 수 있으며, 동시에 서처 사이즈믹은 강력한 다목적 데이터 허브를 플랫폼 삼아 다양한 이니셔티브를 통해 새로운 수입원을 창출할 수 있다.

지역

전 세계

산업

석유, 가스

“플래시블레이드 데이터 허브를 기반으로 구축한 기술들은 우리 비즈니스를 혁신할 것이다.”

조슈아 토프, 지질탐사 부서 매니저

새로운 비즈니스 이니셔티브의 속도를 높이기 위하여 퓨어스토리지의 플래시블레이드를 선택한 서처 사이즈믹

매장된 석유와 가스에 대한 탐사는 전 세계적으로 끊임없이 지속되고 있으며, 어떠한 산업보다 더 많은 지출을 요하고 있어 해당 업계는 높은 부담과 위험을 안고 있다. 따라서 위험과 비용을 줄이고, 결과를 더욱 빠르게 도출해 내는 도구와 서비스가 매우 중요하다. 이러한 환경에서 서처 사이즈믹과 같은 기업은 매우 특별하고 가치 있는 역할을 수행하고 있다.

서처 사이즈믹은 전세계 석유 및 가스 기업들에게 비독점적인 데이터를 제공하는 독립 기업이다. 서처 사이즈믹은 탄성파 탐사 자료를 수집하고 처리하여 고객의 비용을 줄이고 탐사의 기회는 높인다.

최근 서처 사이즈믹은 경쟁력을 더욱 강화했다. 탄성파 탐사 자료를 고객에게 전달하는 시간을 줄여 고객이 인사이트를 얻는데 걸리는 시간(time to insight)을 단축한 것이다. 해당 업무는 서처 사이즈믹의 호주 퍼스 본사 지질탐사 부서의 조슈아 토프(Joshua Thorp) 매니저가 총괄하고 있다. “일반적으로 석유 및 가스 기업들은 많은 탄성파 탐사 자료를 입수하지만 이를 잘 활용하고 있지는 않다. 우리와 같은 멀티클라이언트 기업은 데이터를 수집하여 정리하고, 고객이 쓸 수 있는 형식으로 만들어 제공한다. 하지만 이는 시간도 오래 걸리고, 많은 인력이 다양한 프로세스를 수행해야 얻을 수 있다. 우리 목표는 해당 프로세스를 줄이고, 고객 경험을 개선하는 것이다.”

토프 매니저는 프로세스를 줄이는 것뿐만 아니라 석유, 가스 탐사 부문을 변화시킬 비전을 가지고 있었다. “최종적으로 우리가 원하는 것은 프로세스에서 모든 물리적인 매체를 없애는 것이다. 이를 위한 첫번째 단계가 바로 고객에게 직접 데이터베이스에서 데이터베이스로 데이터를 이전하는 것이다. 하지만 향후에는 선박이 바다 위에서 포착한 데이터가 곧바로 위성으로 전송되고, 답러닝과 빅데이터 분석을 통해 실시간으로 처리될 것이라 생각한다. 만약 결과를 얻기까지 9개월에서 18개월가량이 걸리지 않고, 선박이 바다 위에 있을 때 바로 탄화수소를 탐사할 수 있다면, 이 산업은 완전히 바뀔 것이다.”

2017년에 토프 매니저가 호튼웍스(Hortonworks)와 하둡 분산 파일 시스템(HDFS)을 사용하여 데이터 관리 개선 프로젝트를 시작하였을 때 처음으로 고려한 것은 주요 클라우드 기업의 서비스를 사용하는 것이었다. 하지만 6개월가량의 기술 검증(POC)을 통해 그는 클라우드 기반 솔루션에 몇 가지 단점이 존재한다는 것을 깨닫게 되었다. “우리는 클라우드에 플래시 스토리지를 직접 연결하여 사용하는 것이 매우 비싸다는 것을 알게 되었다. 그리고 속도가 빠르기는 했지만, 우리가 원하는 것만큼 빠르지는 않았다. 오버헤드가 많아 전송 시간이 꽤 느렸다. 우리 데이터 세트의 사이즈를 고려하였을 때 성능을 최대화할 필요가 있었다. 클라우드의 전송율에 기반하면 모든 데이터를 입수하는데 수 년이 걸리고, 저장하는 데는 수십만 달러가 소요된다는 계산이 나왔다. 따라서, 우리는 다른 대안을 찾기 시작했다.”

기업

서치 사이즈믹(Searcher Seismic)
<https://www.searcherseismic.com/>

활용 사례:

- 탄성파 탐사 자료 프로세싱
- 고성능 컴퓨팅 - 딥러닝, 시각화, 하둡(Hadoop)
- 분석 - 호튼웍스 H베이스(Hortonworks® HBase), 나이파이(NiFi), 솔라(Solr), 스파크(Spark)
- 데이터 웨어하우스 - 호튼웍스 하이브(Hortonworks Hive)

도전과제:

- 클라우드 기반의 데이터 스토리지와 프로세싱은 속도가 느리고 유연하지 않으며, 비용이 높음.

IT 혁신:

- 플래시블레이드는 탄성파 탐사 자료의 입수 속도를 20~30배 향상
- 비즈니스 니즈에 따라 유연하게 확장 가능한 스토리지 용량
- 관리의 용이함 덕분에 IT 인적 자원을 최적화

“데이터 허브라는 개념은 바로 우리가 찾고 있던 것이었다. 퓨어스토리지는 우리와 같은 생각을 하고 있었다.”

조슈아 토프, 지질탐사 부서 매니저

로컬 SI인 MSS는 토프 매니저에게 퓨어스토리지를 제안했다. 당시는 퓨어스토리지가 플래시블레이드 데이터 허브를 출시한 지 얼마 되지 않은 시기였다. 플래시블레이드는 스케일 아웃 스토리지 시스템으로, 현대적 분석과 빅데이터 애플리케이션을 위하여 설계되었고, IOPS, 처리량, 레이턴시, 용량을 포함한 모든 측면에서 최고의 성능을 제공한다.

성능 및 용량의 비약적 향상을 가능하게 하는 플래시블레이드

토프 매니저는 퓨어스토리지의 플래시블레이드 데이터 허브에 대해 관심을 가지게 되었다. “데이터 허브라는 개념은 바로 우리가 찾고 있던 것이었다. 우리는 전사적으로 모든 데이터에 접근할 수 있기를 원했다. 퓨어스토리지가 같은 생각을 하고 있다는 사실이 고무적이었다.”

하지만 토프 매니저는 구매를 추진하기 전, “과연 자사의 소프트웨어가 플래시블레이드에서 구동 될지”에 대한 우려가 있었다. 따라서 퓨어스토리지와 기술 검증을 진행했다. “퓨어스토리지의 판매 담당 엔지니어가 코딩에 약간의 최적화를 수행하였고, 결과적으로 우리 소프트웨어가 플래시블레이드와 호환될 뿐만 아니라 전보다 훨씬 빠르다는 것을 확인할 수 있었다. 그리고 우리 소프트웨어를 개선할 수 있는 몇 가지 제안을 기반으로 성능이 훨씬 더 개선되었다.”

기술 검증이 마무리될 때쯤, “우리는 플래시블레이드가 주된 문제였던 입수 속도를 20~30배 향상시킨다는 것을 확인했다. 엄격한 테스트를 통하여 하루에 300TB를 입수할 수 있다는 것을 검증했으며, 이는 우리의 필요를 충족하는 것을 넘어 클라우드에서 시도하였던 어떤 방법보다 더 빠른 속도를 제공했다.”

기술 검증에서 매우 인상깊었던 결과를 바탕으로 토프 매니저는 2PB 이상의 사용가능 용량으로 플래시블레이드를 주문했다.

가까운 시일 내에 플래시블레이드는 각 탐사 때마다 수백 GB에 이를 수 있는 서치 사이즈믹의 프로세스된 데이터를 관리하게 될 것이다. 그리고 토프 매니저는 현재 최대 15PB로 점점 더 증가하고 있는 미가공 데이터를 결국에는 전부 데이터 허브로 옮길 것으로 예상하고 있다.

플래시블레이드의 적용 후 즉각적으로 나타난 효과는 데이터 준비 직원들의 생산성이 높아졌다는 것이다. 토프 매니저는 “과거 우리는 데이터를 준비하고 고객에게 제공하기 위하여 데이터 관리팀 전체가 풀타임으로 일했다. 하지만 플래시블레이드의 간편함 덕분에 우리는 20가지 단계와 많은 인력이 필요했던 프로세스를 단 한 번의 클릭으로 대체하게 되었다. 이는 데이터 전달 시간을 대폭 줄이며, 오류 가능성을 낮추고, 직원들을 보다 중요한 프로젝트에 배정할 수 있도록 한다. 고객에게는 더욱 좋은 경험을 선사하는 동시에 우리 회사는 효율성을 높일 수 있다.”라고 설명했다.

서치 사이즈믹은 플래시블레이드를 사용하여 가치 있는 ‘웹 기반 시각화’ 서비스를 새롭게 제공할 수 있게 되었다. “서치 사이즈믹은 이제 데이터가 고객들에게 전달되기 전에 고객들이 로그인하여 데이터를 확인할 수 있는 웹사이트를 보유하고 있다. 최근 이 기능을 고객에게 선보이자 좋은 반응을 얻었다. 고객들은 탄성파 탐사 자료가 고객의 기업에 도달하여 시스템에 통합되기까지 6주에서 3개월이 걸린다고 말했다. 플래시블레이드의 속도와 용량을 이용한다면 고객사의 연구진들이 접근할 수 없었던 데이터를 살펴볼 수 있는 기회를 제공하는 것이다.”

플래시블레이드, 전사적으로 사용되는 데이터 허브

토프 매니저는 플래시블레이드 데이터 허브를 사용하여 더욱 새롭고 강화된 서비스를 구축하고 있다. “우리 모든 데이터는 플래시블레이드에 저장되어 더욱 빠른 고객 데이터 전달, 시각화, 딥러닝, 고도화된 분석 등 다양하게 활용될 것이다.”

탄성파 탐사 자료는 기존에 업계 표준인 SEGY 파일 포맷으로 저장되어 있어 고객이 사용하기 전에 해석 소프트웨어(interpretation software)를 돌려야 했다. 서처 사이즈믹은 해석 소프트웨어 공급 업체와 함께 소프트웨어가 플래시블레이드와 잘 통합될 수 있도록 했다. “우리는 이를 통해 모두가 데이터에 직접 접근할 수 있도록 하여, 복사본이 여러 개 만들어지지 않도록 했다.”라며 토프 매니저는 “이것이 의미하는 바는 기업 내부 사람들이 소프트웨어 개발자와 연구진을 좋아하게 되었다는 것이다. 하지만 나중에는 우리 고객들이 직접 데이터에 접근할 수 있도록 할 것이며, 퍼블릭 도메인 데이터에는 자유롭게 접근할 수 있도록 할 것이다. 플래시블레이드의 속도는 이를 가능하게 할 만큼 빠르다.”라고 설명했다.

토프 매니저에 따르면 플래시블레이드의 가장 큰 매력은 장기적인 관점에서 유연하고 다용도라는 것이다. “우리는 구매를 하기 전에 타당성 실험을 수행하고 독립 컨설턴트에게 제안서를 검토 받았다. CFO와 이사진 또한 이 프로세스에 적극적으로 참여했다. 모두가 한 목소리로 우리가 어떤 소프트웨어를 쓰든지 플래시블레이드를 사용할 수 있다고 결론지었다. 플래시블레이드는 빠르고 유연하여 우리가 필요한 어떠한 데이터베이스 아키텍처도 다룰 수 있기 때문이다.”

서처 사이즈믹은 최근 업계에서 전례 없는 속도로 더욱 강력해진 시각화와 분석 서비스를 제공하는 사이즈믹(Saismic)을 도입했다. 그리고 플래시블레이드는 사이즈믹을 가능케하는 핵심적인 요소이다. “플래시블레이드 데이터 허브와 함께 구축한 기술은 우리 비즈니스에 변화를 불러와 새로운 수익원에 대한 가능성을 열어주며, 우리가 비즈니스를 하는 모든 방식에 영향을 주게 될 것이다. USB에 저장하여 데이터를 물리적으로 고객에게 제공하는 방식에서, 시각화와 구독 모델을 기반으로 고객이 실시간으로 처리 데이터에 접근할 수 있게 하는 방식으로 변할 것이다.”

다양한 이점을 제시하는 플래시블레이드

토프 매니저는 플래시블레이드 데이터 허브를 도입하였을 때의 다양한 이점을 다음과 같이 설명했다.

성능. “즉석에서 바로 사용할 수 있어 매우 빠르다. 딥러닝과 시각화는 모두 대역폭과 레이턴시에 상당히 민감할 수 있지만, 플래시블레이드는 이 두 가지 측면에서 뛰어난 성능을 발휘한다. 클라우드 서비스를 사용할 당시 우리는 500MB의 대역폭이 최대였지만, 플래시블레이드는 초당 17GB를 구가한다. 이는 30 배 이상 높은 수준이다.”

손쉬운 배치. “플래시블레이드의 설치는 매우 쉬웠다. 우리 IT 부서는 매우 규모가 작기 때문에, 수백개의 하드 드라이브를 설치하거나 복잡한 설정을 하지 않아도 되는 것이 매우 중요했다. 스토리지의 프로비저닝도 매우 쉬웠다. 작업 영역을 설정하고, 확장하거나 완전히 삭제할 수 있으며, 이러한 변화를 바로 적용할 수 있다.”

업그레이드 가능성. “플래시블레이드 아키텍처는 미래의 성장을 수용할 수 있도록 설계되어 있다. 용량을 확장하기 위하여 새로운 블레이드만 추가하며 되기 때문에 큰 규모의 IT 프로젝트를 계획하지 않아도 된다. 다수의 플래시블레이드 새시를 테더링할 수도 있으며, 병렬식의 이점도 얻을 수 있다. 현재의 환경 설정은 가공된 데이터를 위한 것이다. 그리고 이후에 우리가 미가공 현장 데이터를 입수할 준비가 되면, 기존 투자 비용 손실 없이 플래시블레이드를 확장할 수 있다.”

효율성. “우리가 필요한 것은 모든 데이터를 저장하고, 프로그램을 기반으로 접근할 수 있으며, 데이터를 클라우드나 FTP 사이트, USB 드라이브에 복사할 수 있도록 충분히 빠른 시스템이었다. 우리는 인간이 수행하던 20개 단계의 절차를 자동화하여 클릭 한 번으로 대체하고 싶었고, 그것이 바로 플래시블레이드로 인해 가능해졌다. 적은 시간과 자원으로 더 많은 일을 할 수 있게 된 것이다.”

“우리는 플래시블레이드가 주된 문제였던 입수 속도를 20~30배 향상시킨다는 것을 확인했다.”

조슈아 토프, 지질탐사 부서 매니저

“플래시블레이드를 통해 우리는 더 적은 시간과 자원으로 더 많은 일을 할 수 있게 되었다.”

조슈아 토프, 지질탐사 부서 매니저

총소유비용(TCO). “석유 가스 산업에서는 데이터가 너무 많기 때문에 클라우드로 기업의 모든 데이터를 옮겨야 한다는 믿음이 있었다. 하지만 플래시블레이드는 우리 기업의 모든 데이터를 저장하고 코로케이션 데이터 센터에 단 4U만 차지했다. 이는 큰 용량에 비하여 매우 적은 용량을 차지하는 것이다. 이런 질문을 하고 싶다. 플래시블레이드를 선택하지 않을 이유가 무엇인가. 기업의 모든 데이터가 저장되어 있는 확장성 있는 시스템을 바로 옆에 설치할 수 있는데, 왜 클라우드의 문제점과 비용을 감당해야 할까?”

“클라우드 벤더와의 기술 검증에서 스토리지와 컴퓨팅은 함께 묶여 있었으며, 둘 중 어떤 것을 원하던 간에 함께 구매해야 했다. 하지만 최신 소프트웨어 프로토콜을 기반으로 하는 플래시블레이드에서는 스토리지와 컴퓨팅을 분리할 수 있다. 이는 고객이 비용을 절약하고 비즈니스 니즈에 더욱 빠르게 대응할 수 있도록 해준다.”

다양한 워크로드의 통합. “앞으로 우리는 데이터를 새롭고 다른 방식으로 사용하고 싶다. 우리는 지형 배치에서 석유나 가스가 매립되어 있을 가능성이 있는지 없는지를 볼 때, 딥러닝을 통하여 더욱 빠르게 판단할 수 있는 가능성이 있다는 것을 확인했다. 우리는 엔비디아(NVIDIA) DGX와 HGX GPU(그래픽 처리 장치)의 가치가 높다는 것을 확인했다. 하지만 만약 GPU 팜(farm)에 투자를 한다면, 최고의 가치를 얻기 위해서는 지속적으로 데이터를 투입해야 한다. 이것이 바로 플래시블레이드가 하는 일이다. 플래시블레이드가 엔비디아 DGX-1와 거의 네이티브 수준의 통합을 이루는 것은 매우 인상깊었다. 이 둘은 함께 잘 구동된다.”

토프 매니저는 “아직까지 활용되지 않은 플래시블레이드의 가능성은 바로 미가공 필드 데이터에 관한 것이다. 워낙 필드 데이터의 용량이 크기 때문에 현재까지는 업계의 그 누구도 이 미가공 데이터를 활발하게 사용하지 못한다. 하지만 플래시블레이드는 빠른 속도와 다양한 기능을 기반으로 최초로 필드 데이터를 의미 있는 방식으로 사용할 가능성이 보인다.”고 말했다.

사이즈믹(Saismic) 서비스 및 여타 다른 이니셔티브는 서처 사이즈믹이 석유, 가스 산업의 판도를 뒤집고 선두를 달릴 수 있도록 설계되어 있다. 그리고 플래시블레이드는 이러한 노력의 중심에 서 있다. 토프 매니저는 “우리는 업계의 미래에 대한 비전을 가지고 이러한 노력을 시작했다. 플래시블레이드는 이러한 비전을 실현하는데 많은 도움을 주고 있다.”라고 덧붙였다.

purestorage.com/kr

+82 2 6001-3330

