

솔루션 개요

# 플래시블레이드(FlashBlade)로 경험하는 현대적 분석

빅데이터는 더 이상 정적이며 예측이 가능한 데이터가 아닙니다. 현대적 분석은 미지의 영역을 위해 설계된 인프라와 비정형 데이터에 클라우드와 같은 민첩성과 실시간으로 높은 성능을 제공할 수 있어야 합니다. 현대의 분석 워크로드를 위해 설계된 업계 최초의 동적 데이터 허브인 퓨어스토리지의 플래시블레이드(FlashBlade™)를 통해 기업에서 사용하는 하둡(Hadoop), 스파크(Spark) 및 데이터 분석을 현대화하세요.

## 과거: 레거시, 디스크 기반의 하둡

십년 전의 데이터는 지금과 많이 달랐습니다. 그 때의 하둡 파일 시스템(HDFS)과 분산형 직접 연결 스토리지(DDAS) 아키텍처는 정적이고 예측가능한 데이터를 위하여 설계되었습니다. 물론 순차적인 배치(batch) 워크로드 처리 작업에는 잘 활용되었으나, 아파치 스파크(Apache® Spark™), 머신러닝 등의 현대적 애플리케이션에서는 복잡한 애플리케이션 사일로로 인하여 요구되는 민첩성과 실시간 성능을 얻을 수 없었습니다.



## 모든 것을 바꾸는 현대적 데이터

업계는 새로운 시대에 직면해 있습니다. 십년 전만해도 상용화된 분석 툴은 하둡이 유일했지만, 오늘날의 데이터 과학자가 사용할 수 있는 도구는 너무나도 다양합니다. 스파크, 카프카(Kafka™), 엘라스틱(Elastic) 등의 도구들은 데이터 분석의 지평을 매우 다양한 방향으로 넓히고 있습니다. 또한, 머신러닝과 AI의 등장으로, 비정형의 랜덤 데이터를 기반으로 새로운 인사이트를 얻고, 한 때 공상과학으로만 여겨지던 혁신을 가속화할 수 있게 되었습니다.

## 십년 전의 데이터

### GFS & HDFS 데이터에 대한 추측:

- 일반적으로 큰 파일
- 순차적 접근
- 일상적인 하드웨어 고장
- 일괄 배치(batch) 처리되는 데이터
- 느린 네트워크

### 오늘 날의 데이터는 다릅니다:

- 작은 파일에서 큰 파일까지
- 랜덤 접근부터 순차적인 접근까지
- 실시간 혹은 배치(batch) 처리
- 빠르게 진화하는 애플리케이션과 데이터
- 탄력적인 인프라





## 고객 하이라이트

### 업계를 선도하는 투자기업 Man AHL이 도입한 10배 더 빠른 아파치 스파크

Man AHL은 체계적인 퀀트투자 기업으로 \$190억 이상의 자산을 관리하고 있습니다. Man AHL은 인간의 개입 없이 컴퓨터의 수학적 모델을 기반으로 투자 결정을 내립니다. 약 50명의 퀀트 연구진(quants)이 IT 도구에 의존하기 때문에 Man AHL은 플래시블레이드 기반의 인프라를 구축하기로 결정했습니다.

데이터 엔지니어링 부서장인 제임스 블랙번(James Blackburn)은 “많은 연구진들이 플래시블레이드의 도입 이후 스파크를 사용해 여러 시뮬레이션을 수행하는 것이 더욱 쉬워졌고, 심지어 그 중 한 명은 이전에 비해 스파크 워크로드의 처리량이 10배 20배 증가하는 것을 경험했습니다”라고 밝혔습니다. 이는 타임-투-마켓(time-to-market) 경쟁력을 높이며 시장 판도를 뒤엎고 있습니다.

## 현대적 데이터를 위하여 설계된 플래시블레이드

플래시블레이드는 업계 최초로 현대적 분석을 목적으로 설계된 동적 데이터 허브로, 클라우드와 같은 민첩성을 바탕으로 모든 비정형 데이터에 대한 실시간 대응을 가능케합니다. 플래시블레이드는 방대한 병렬 구조에 기반하여, 데이터 과학자와 엔지니어에게 전례 없는 성능과 간소성을 제공합니다. 작고 메타데이터가 많은 워크로드부터 랜덤하고 큰 파일에 대한 접근까지 모든 미지의 데이터를 위하여 설계되었으며, 기업의 어떠한 빅데이터 이니셔티브도 지원할 수 있습니다.

- 현대적 비정형 워크로드를 위한 실시간 성능 제공
- 클라우드와 같은 민첩한 인프라를 제공하기 위하여 컴퓨팅과 스토리지 분리
- 미지의 데이터를 위하여 설계된 데이터 플랫폼으로 미래 투자 보호

## 플래시블레이드: 오늘 날의 하둡과 스파크 워크로드를 위하여 구축된 데이터 플랫폼

DDAS 아키텍처는 과거의 분석에는 적절했지만, 현대의 데이터는 클라우드와 같은 민첩성과 간소성을 지닌 비정형 워크로드를 위한 데이터 허브 및 새로운 수준의 스케일-아웃 스토리지를 필요로 합니다. 플래시블레이드는 현대적 분석을 위하여 구축되었으며, 차세대 빅데이터를 위한 6가지 핵심 역량을 보유하고 있습니다.

### “모든 상황에 적절”



7.5M IOPS, 75GB/초 처리량으로, 작고, 크고, 랜덤하고, 순차적인 모든 데이터를 빠르게 처리 가능

### 실시간



즉각적인 워크로드를 위한 밀리세컨드 레이턴시(latency)

### 올플래시



소프트웨어에서 하드웨어까지 최적화된 플래시 기술

### 병렬 구조



방대한 병렬 구조 기반의 퓨리티 SW  
75개 블레이드까지 확장 가능한 스케일-아웃 패브릭

### 탄력적



각각의 블레이드로 선형적인 퍼포먼스 향상

### 간소함



에버그린(Evergreen)—TB를 재구매 하거나 수동 조율 불필요

[purestorage.com/kr](http://purestorage.com/kr)

+82 2 6001-3330

