



HPE Ezmeral

Data Fabric 소개

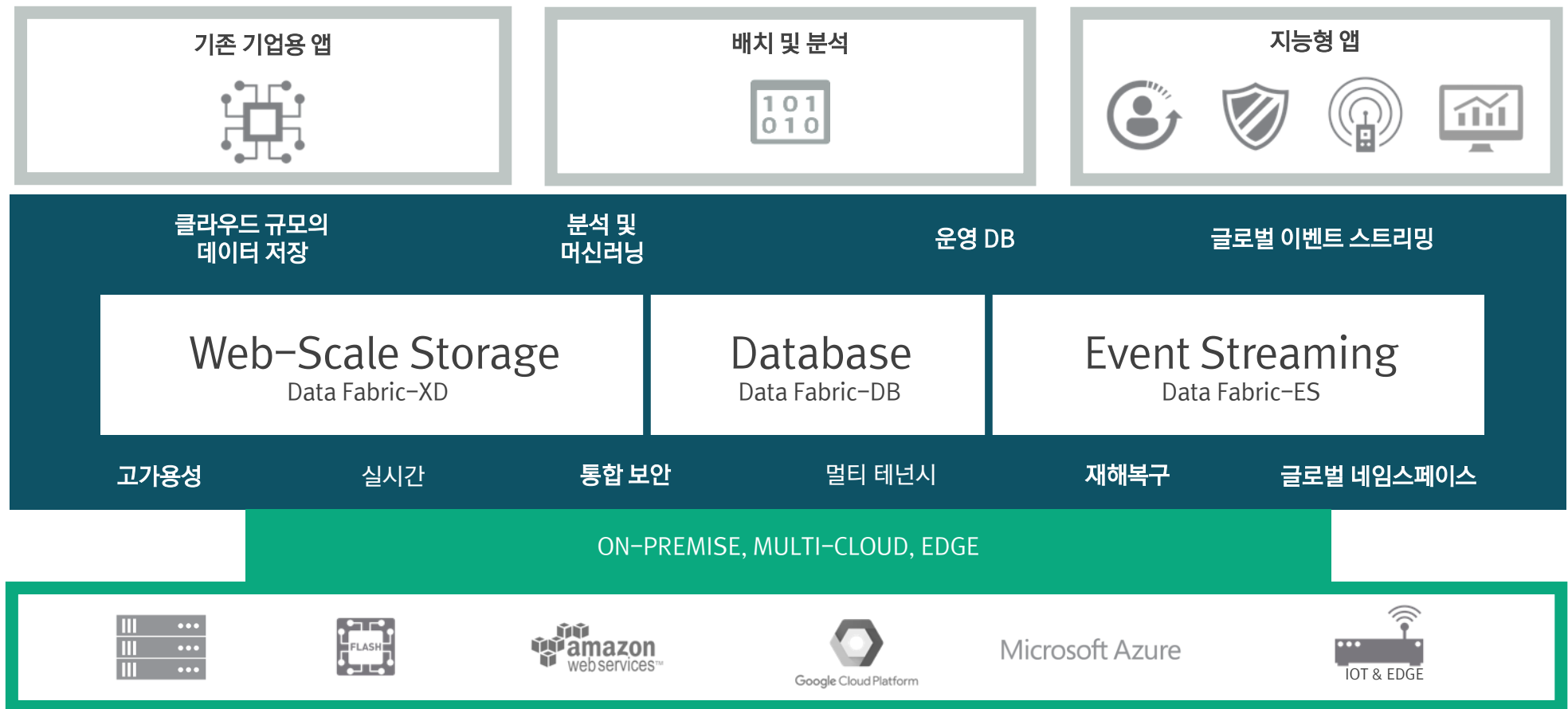
HPE EZMERAL이란?

기업이 Edge에서 Cloud에 이르기까지 조직 전체의 디지털 전환을
가속화하도록 설계된 새로운 브랜드 및 소프트웨어 포트폴리오

Maximizing Value from the HPE Ezmeral Data Fabric Data Platform “ According to IDC, HPE Ezmeral Data Fabric ROI is over **576%** “

HPE Ezmeral Data Fabric 이란?

HPE Ezmeral Data Fabric은 기업 요건이 반영되지 않은 단순한 Hadoop 오픈 소스 기반의 빅데이터 플랫폼이 아닌 엔터프라이즈 환경에서 필요로 하는 데이터의 안전성, 성능, 가용성을 강화한 **글로벌 최고의 유일 상용 데이터 플랫폼**입니다.



HPE Ezmeral Data Fabric 단일 데이터 플랫폼은 Hadoop, Apache Drill 및 Spark와 결합하여 데이터의 종류와 상관없이 하나의 단일 플랫폼에서 처리할 수 있도록 함으로써 고객들에게 가장 빠르고 저비용 고효율적인 인사이트를 제공합니다.

HPE Ezmeral Data Fabric 특징



대규모 복원력

엄격한 비즈니스 SLA를 충족하기 위해 자동 장애 조치, 스냅 샷, 파일 미러링 및 즉각적인 교차 DC 복제 기능을 갖춘 엔터프라이즈급 재해 복구 기능



글로벌 네임 스페이스

전 세계적으로 배포된 데이터에 대한 응용 프로그램 액세스를 단순화하고 관리자가 여러 개의 분산된 클러스터를 논리적 로컬 클러스터로 뷰잉.



정책 기반 데이터 티어링

애플리케이션을 위해 핫, 웜 및 콜드 데이터 계층에 걸쳐 데이터 스토리지 및 관리를 자동화하여 성능, 비용 및 용량 트레이드 오프를 최적화



범용 데이터 서비스

복원력, 가용성, 복구 및 보안 기능이 내장된 파일, 데이터베이스 및 이벤트를 제공할 수 있는 단일 분산 데이터 서비스



HA 기본 제공 - 분산 메타데이터

메타 데이터 및 애플리케이션 데이터의 지능적이고 자동화된 복제로 즉시 사용 가능하고 가용성 (HA), 확장 가능 및 자체 치유 시스템 생성



서비스 관리 및 탐색

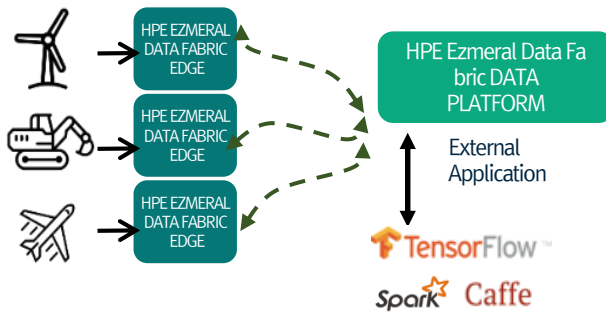
광범위한 애플리케이션에 필요한 필수 서비스를 지속적으로 발견하고 제공하기 위한 핵심 메커니즘 세트



공통 보안 및 거버넌스 모델 : 기본 제공 감사, 표현 권한 부여 및 모든 사용자 이름 / 암호 레지스트리 및 Kerberos를 지원하는 유연한 인증을 통해 모든 데이터에서 통합된 보안 플랫폼

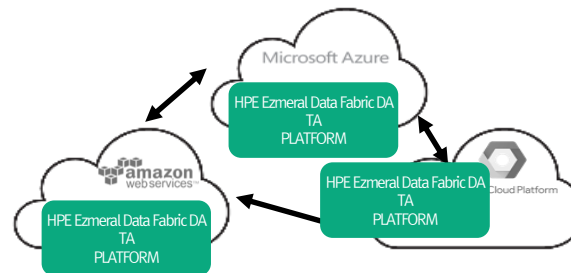
HPE Ezmeral Data Fabric 핵심 기능

IoT / Edge Analytics 지원



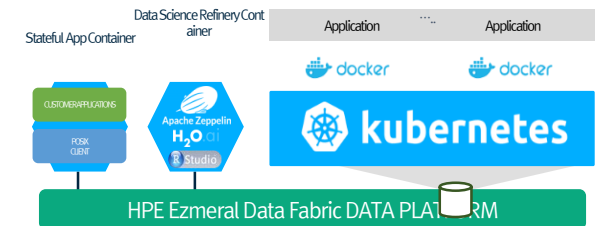
때때로 연결된 장치를 위한 작은 설치 공간,
로컬 처리 및 효율적인 통신 프로토콜

멀티 클라우드 지원



온프레미스(On-premise)와 클라우드(Cloud)의
일관된 데이터 관리를 통해 광범위한 분석 및
AI 기술을 통합

컨테이너 지속성 지원



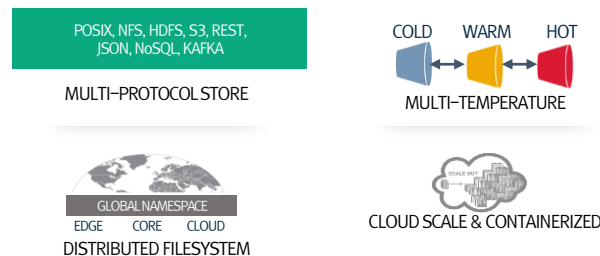
컨테이너 애플리케이션이 HPE Ezmeral Data Fabric
에서 데이터에 액세스하여 기업이
컴퓨팅/스토리지를 독립적 확장

AI / ML용 플랫폼 지원



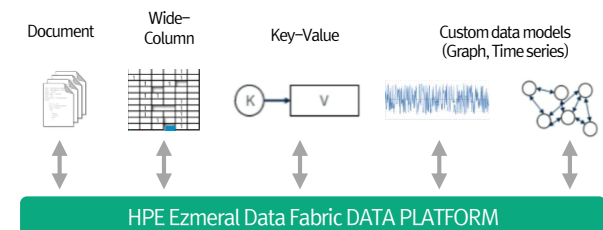
AI / ML 기술은 데이터 현지화로 GPU 최적화 된
동일한 클러스터의 POSIX 인터페이스를 통해
직접 실행

엔터프라이즈 프로덕션 규모 지원



엑사 바이트 규모의 분산 플랫폼; POSIX, NFS, HDFS,
S3, REST, JSON, KAFKA 인터페이스의 읽기 / 쓰기

다중 모델 데이터베이스 지원



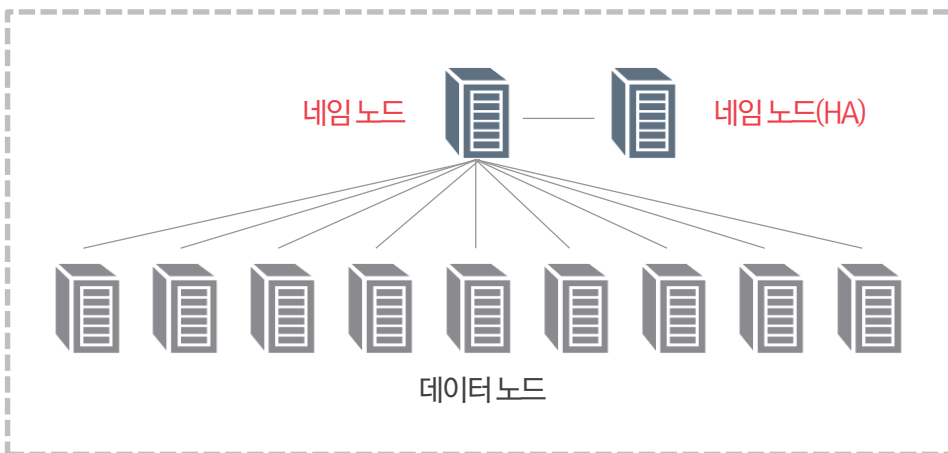
다양한 사용 사례에 대한 다중 모델 데이터베이스
구조, 부하 및 구조, 분석, AI/ML 및 데이터
웨어하우스 지원

HPE Ezmeral Data Fabric – 단일 장애점 제거

Apache Hadoop의 필수적인 요소인 네임노드는 분산환경에서 데이터 블록의 위치와 변경사항을 관리하는 핵심 요소이며, 장애와 번복현상으로 매우 불편하지만, HPE Ezmeral Data Fabric은 **네임노드를 제거한 아키텍처**로 고객의 고 품질의 서비스가 가능한 유일한 플랫폼입니다.

기존 Hadoop File System

- 네임 노드 장애시 전체 클러스터가 중지되는 단일 장애점 발생
- 대기(Stand-by) 네임노드는 분산 처리가 아닌 장애 대비
- 대량의 노드가 구성된 클러스터 환경에서 네임노드 부하시 전체 클러스터의 성능이 저하되는 치명적인 구조



HPE Ezmeral Data Fabric File System

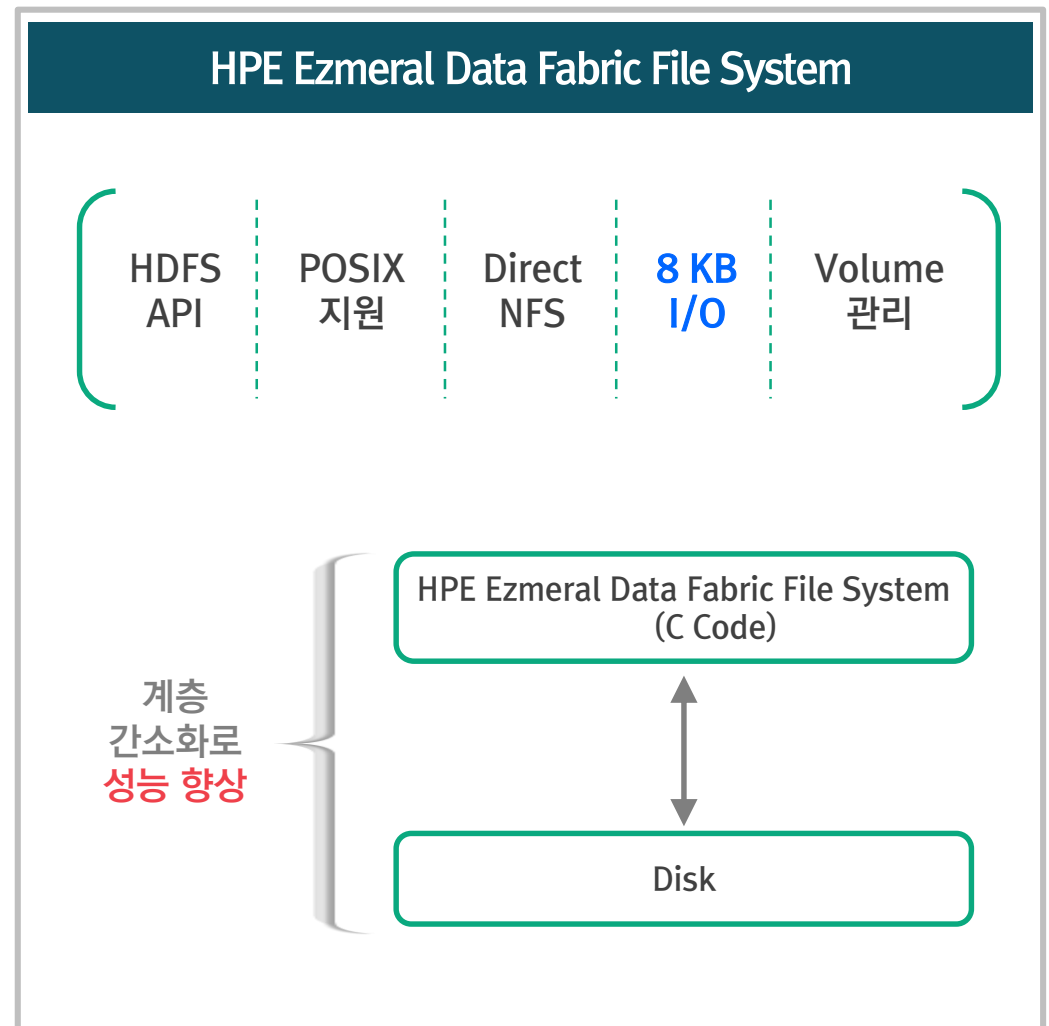
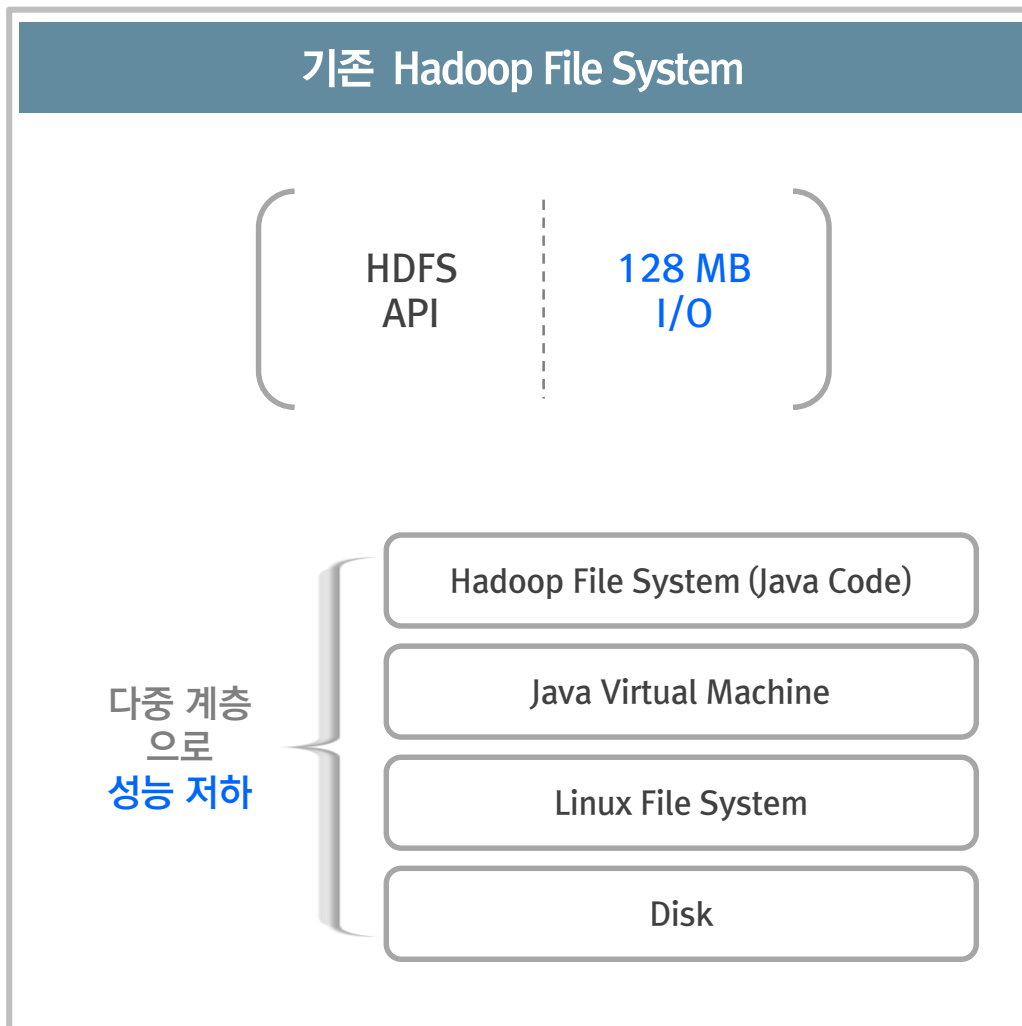
- 네임노드를 제거하여 **단일 장애점(SPOF) 제거**
- CLDB(Container Location Database) 기술기반 운영
- 대량의 노드가 구성된 클러스터 환경에서도 컨테이너 위치와 변경사항을 분산처리 하여 **성능에 영향을 미치지 않는 구조**



HPE Ezmeral Data Fabric – File System

빅데이터 플랫폼을 통한 선제적 의사결정 환경을 구현하기 위해서는 적시성을 확보하는 것이 핵심입니다.

HPE Ezmeral Data Fabric Data Platform은 기존 Hadoop의 아키텍처를 완전히 혁신하여 실시간 업무와 Mission-critical한 운영 환경을 제공합니다.



HPE Ezmeral Data Fabric – Drill (매우 빠른 데이터 쿼리 엔진)

HPE Ezmeral Data Fabric 데이터 플랫폼 내의 모든 정형, 비정형, 반정형 등의 데이터를 조회할 수 있는 Drill을 제공하며 별도 스키마를 생성하지 않아 매우 빠릅니다.

아파치 드릴(Apache Drill) 제공

- 아파치 드릴을 기본적으로 제공하며, 데이터 검색을 위해서 별도 스키마(Schema)를 생성하지 않기 때문에 매우 빠른 데이터 검색 성능을 보장.



Raw data

• JSON, CSV, ...



“Optimized” data

• Parquet, ...



Centrally-structured data

• Schemas in Hive Metastore



Relational data

• Highly-structured data

Exploration
(known and unknown questions)

Oracle, Teradata

Hive, Impala, Spark SQL



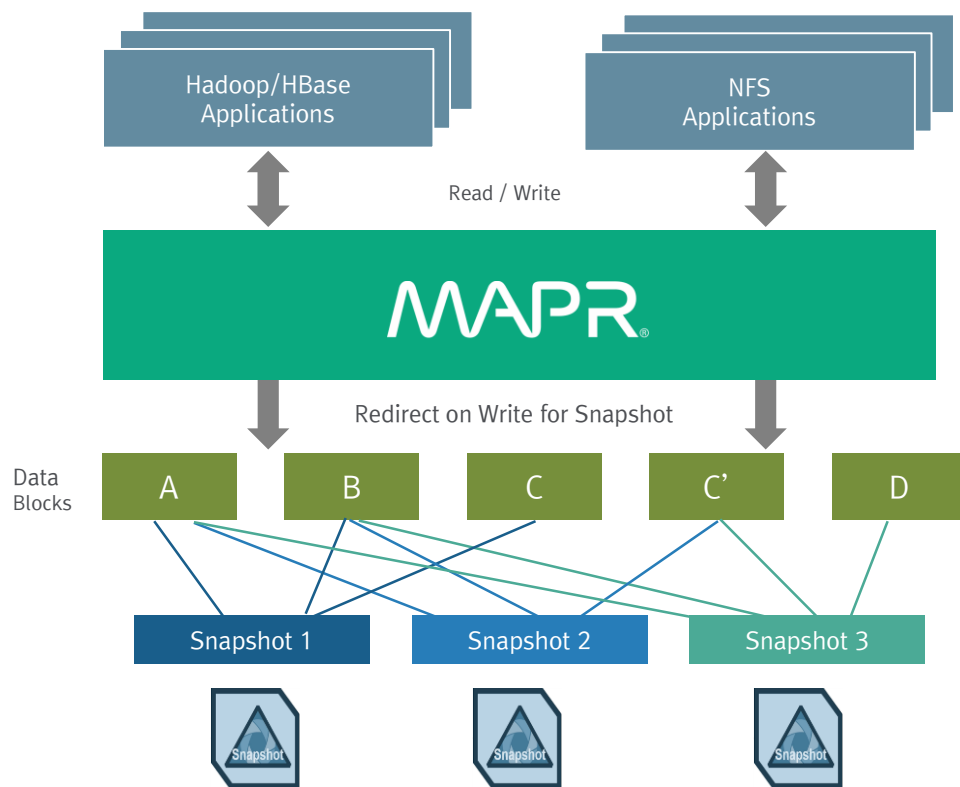
기존 POXIS와 100% 호환

- 기존 POXIS와 100% 호환으로 별도 Script 수정이 불 필요.
- HDFS API, POSIX, NFS, Hbase API, JSON API, Kafka API 등 다양한 프로토콜을 지원.

HPE Ezmeral Data Fabric – Snapshot(데이터 보호)

데이터 볼륨 스냅샷(Snapshot) 기능을 활용하여 완벽한 데이터 백업 및 복구 환경을 구현합니다.

스냅샷을 통한 데이터 보호



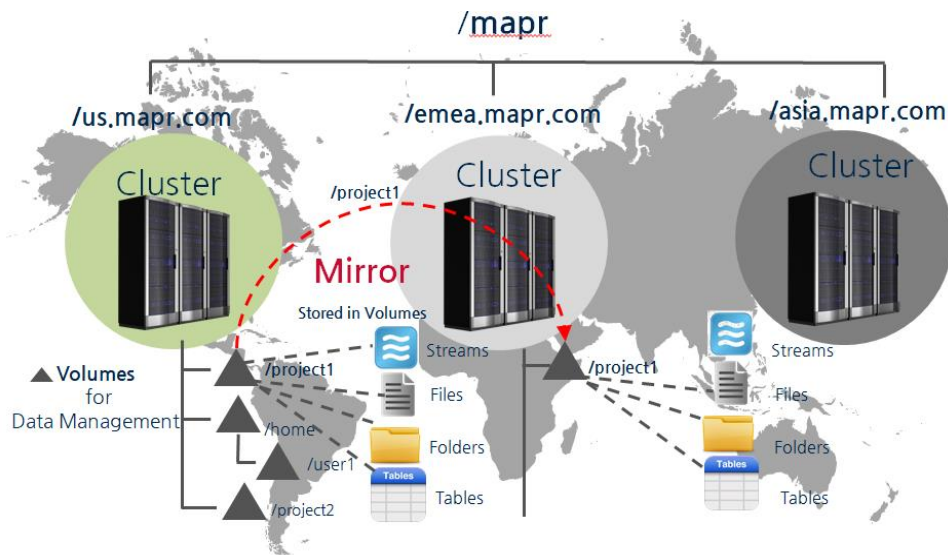
구성 특징

- 데이터 보호를 위한 볼륨 스냅샷 기능 제공
- 스냅샷 생성 및 관리 기능은 엔터프라이즈급 스토리지에서 데이터 보호를 위해서 필수적으로 사용.
- HPE Ezmeral Data Fabric XD에서는 특정 시점의 볼륨에 대한 읽기 전용 스냅샷(이미지)를 생성하여 데이터를 보호.
 - GUI(HPE Ezmeral Data Fabric Control System) 또는 명령어로 쉽고 빠르게 스냅샷 수행
 - 스케줄러 설정으로 데이터 보호 자동화 구현
 - 생성된 시점에서 변경사항에 대하여 증분 저장 방식으로 중복된 데이터는 존재하지 않음

HPE Ezmeral Data Fabric – DR(재해복구)

HPE Ezmeral Data Fabric 데이터 플랫폼은 백업 및 미러(Mirroring) 기능이 포함되어 손쉽게 재해복구환경을 구현할 수 있습니다.

재해복구 구현



성능에 영향없이 사이트간
Mirror 구성

End-to-End Checksum 통한
데이터 일관성 보장

구성 특징

- 단일 데이터 플랫폼 구현을 위한 **글로벌 네임스페이스(Global NameSpace)** 및 다중 원격지 미러 구성
 - 주센터(Master) 에서 재해복구센터(DR 또는 타 사이트)로 연결한 다음 Mirroring 지원
 - 볼륨 스냅샷 기능을 활용하여 백업과 복제의 두가지 방법으로 DR 환경 구현
- 파일, 폴더, 테이블 포함하여 지정된 타 사이트(타 클러스터)로 복사하여 서비스 환경 구현

재난복구 절차	설명
1. 읽기 권한 부여	원격지 마스터 데이터 방지로 읽기 권한 부여
2. 클러스터 활성화	재난 발생시 원격지 클러스터 활성화
3. 쓰기 권한 부여	Promotable Mirrors 기능으로 읽기/쓰기 상태 활성화로 운영서비스 수행

HPE Ezmeral Data Fabric – 경쟁사비교

작은 규모에서 대규모까지 모든 데이터 영역의 데이터 처리를 위한 빅데이터 플랫폼 요구 기능을 100% 충족 합니다.

기능 요구	HPE HPE Ezmeral Data Fabric-XD	Hadoop	IBM GPFS	SDS New Entrants	Legacy NAS (EMC Isilon)	Public Cloud
Software Offering On Standard Servers	Yes	Yes	-	Yes	-	-
Global Namespace	Yes	-	Yes	Yes	-	-
High Speed Client/Ingestors	Yes	-	Yes	-	-	-
Trillions Files	Yes	-	-	May be	-	-
Inline Compression	Yes	Optional	Yes	-	-	Yes
NFS (RW)	Yes	-	Yes	Yes	Yes	Yes
Analytics Module (Hadoop, Spark, Drill, more)	Yes	Yes	-	-	-	-
Massive Scale (100s Nodes & PBs)	Yes	Yes	Yes	Yes	-	-
Object(S3)	Yes	-	Yes	Yes	Yes	yes
Erasur Coding	Yes	Roadmap	Yes	Yes	Yes	Yes
Flash Optimized File system	Yes	-	-	-	Yes	-
Docker Container support	Yes	-	-	-	-	Yes
Faster Rebuilds on 1-2 Disk Failures	Yes	-	-	-	-	-
Fast Rebalancing on Node Outage	Yes	-	-	-	-	-
Billions of Snapshots	Yes	-	-	-	-	Yes
Rack Awareness	Yes	Yes	-	Yes	-	-
Self Healing	Yes	-	-	Yes	-	-
Fine Grain Multi Tenant ACEs	Yes	ACLs	-	-	-	-

HPE Ezmeral Data Fabric – IDC 자료(2020년)

17%의 성공률을 가진 경쟁사의 빅데이터 프로젝트와 달리 HPE Ezmeral Data Fabric 고객들은 85% 이상의 높은 성공 경험 보유

ROI

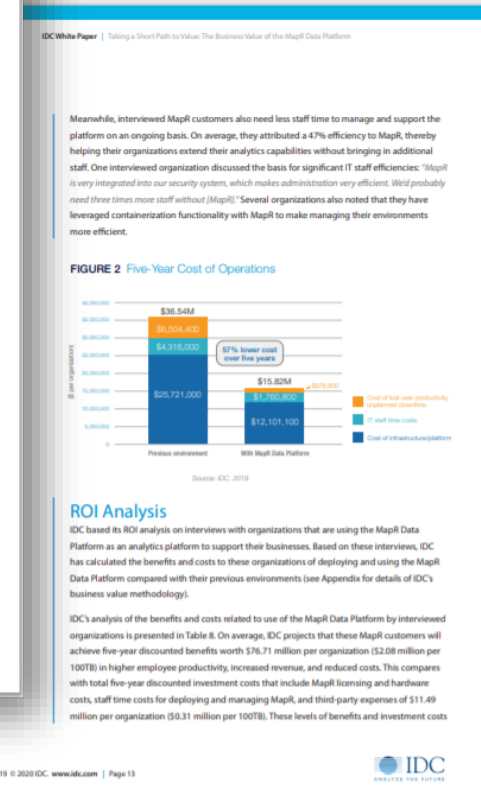
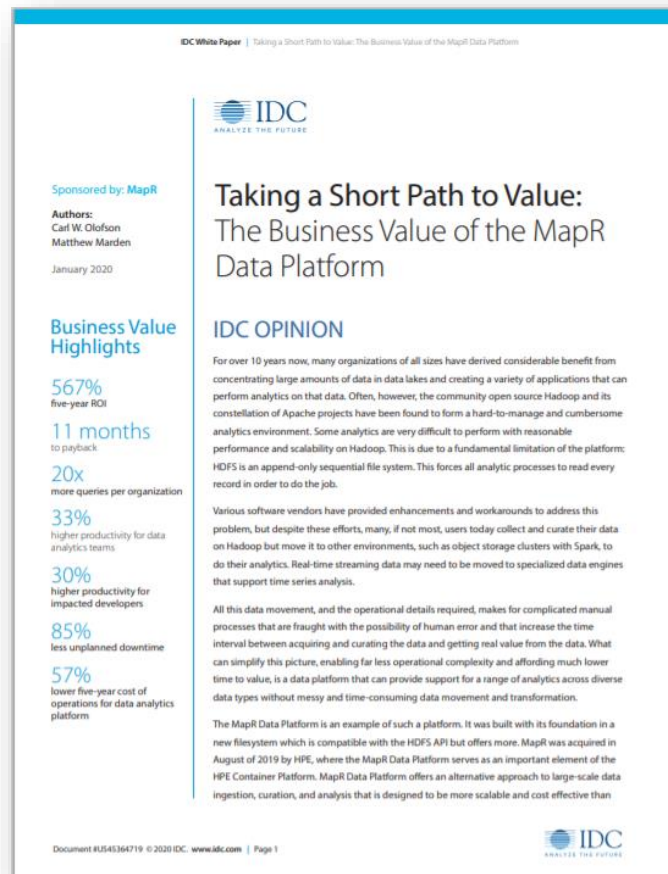
6배

원금 회수 소요
개월 수

11개월

개발자 생산성

85%증대



다운타임

85% 감소

조직 간의 업무 능력

20X

향상

데이터 사이언티스트
생산성

33%증대

HPE Ezmeral Data Fabric은 전 세계 최고의 고객이 선정한 최고의 플랫폼



Enterprise-wide AI, ML, Analytic Data Service

HPE Ezmeral Data Fabric Data Platform Approach Benefits

- 손쉬운 수집 및 통합 : 데이터 이동 없이 모든 워크플로우를 지원하는 많은 API 및 플러그인: HDFS, S3, K8s volume plugin, POSIX
- 분석: 모든 분산 컴퓨팅 알고리즘은 HPE Ezmeral Data Fabric 클러스터와 함께 실행되어 POSIX 또는 NFS를 통해 데이터에 액세스: TensorFlow, Caffe and more
- 보안: 툴 및 애플리케이션은 Data Platform의 HA, DR, 보안 및 안정성을 상속함
- 컨테이너의 stateful 앱을 지원합니다. HPE Ezmeral Data Fabric volume plugin for K8s를 통해 다른 벤더가 할 수 없는 ML 워크플로우 지원
- 까다로운 규제 환경에서 여러 테넌트 및 비즈니스 단위에 대한 안전한 액세스



UnitedHealth Group

- 50개 이상의 사업부를 위한 서비스형 빅 데이터 플랫폼
- 다양한 건강 기록, 청구, 병원, 회원 정보, 콜센터, 웹 트래픽을위한 데이터 레이크
- 데이터 사이언티스트를 위한 협업 환경
- 환자와 프로바이더의 네트워크 그래프 분석
- 클레임 사기 방지 (초기 사용 사례에서 10% 사기 적발 / 청구 건수 180만 건 /일 이상)
- 추가 사용 사례에 따라 계속 증가

비즈니스 가치: 의료 부정 행위 감소, 고객 서비스 향상, 혁신적인 신규 서비스 생성 비용 절감

Why HPE Ezmeral Data Fabric Wins

✓ 수 많은 API, 소스로부터 데이터 수집 및 통합	✓ 분석, ML, AI를 위한 광범위한 처리	✓ 보안을 위한 멀티테넌트, 데이터 유형	✓ 거버넌스, 액세스, 카탈로그, 제어	✓ 운영 애플리케이션에 결과 제공 및 사용
-------------------------------	--------------------------	------------------------	-----------------------	-------------------------

Enhanced Security Using Data Science

- 보안 분석은 데이터 과학과 AI에 의해 주도되어 배치, 마이크로 배치 및 딥러닝 AI가 동시에 필요함
- 분석: HPE Ezmeral Data Fabric Data Platform은 고도로 진화하는 시장에서 필요한 속도와 민첩성으로 대응할 수 있는 모든 기능을 단일 플랫폼에 통합
- 간편한 수집 및 통합: 높은 수준의 빠른 수집 I/O 및 실시간 처리가 필수적임
- 신속한 로컬 응답을 위한 Edge-first 로컬 프로세싱 및 다운스트림 애플리케이션에 결과 제공



Why HPE Ezmeral Data Fabric Wins

✓ 수 많은 API, 소스로 부터 데이터 수집 및 통합	✓ 분석, ML, AI를 위한 광범위한 처리	✓ 보안을 위한 멀티테넌트, 데이터 유형	✓ 거버넌스, 액세스, 카탈로그, 제어	✓ 운영 애플리케이션에 결과 제공 및 사용
--------------------------------	--------------------------	------------------------	-----------------------	-------------------------

Autonomous Automobile Program

목표는 차량 개발의 전체 과정에서 발생하는 오류 및 결함을 줄여 출시 시간을 단축하고 전체 품질을 향상 시키는 것입니다.

- 손쉬운 수집 및 통합 : 수백만 킬로미터를 주행하는 실제 테스트 차량에서 데이터 수집 및 분석
- 광범위한 분석 : 수집된 데이터의 자동 분류, 분류 및 라벨링을 위한 머신 러닝
- 팀을 위한 안전하고 통제된 액세스 : 개발자에게 실제 및 가상 세계의 데이터를 제공하여 지속적으로 주행 기능을 최적화 할 수 있습니다
- 다운 스트림 비즈니스 애플리케이션에 데이터, 앱 및 결과의 글로벌 배치



Major European Auto Manufacturer

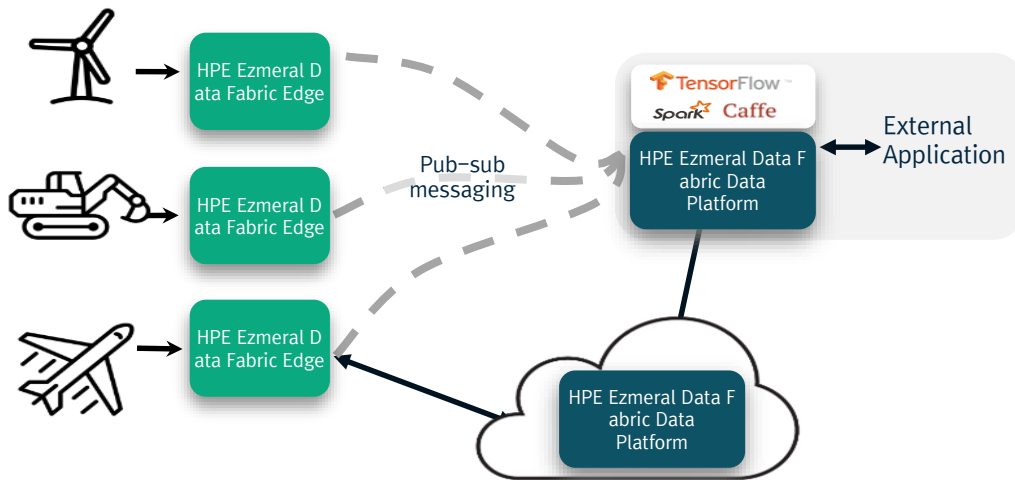
- HPE Ezmeral Data Fabric의 데이터 플랫폼은 자율 주행 기능의 개발 및 운영을 지원하는 데 사용될 것이며 향후 3년 동안이 최대 235PB의 데이터를 생성 할 것으로 예상

Why HPE Ezmeral Data Fabric Wins

✓ 수 많은 API, 소스로부터 데이터 수집 및 통합	✓ 분석, ML, AI를 위한 광범위한 처리	✓ 보안을 위한 멀티테넌트, 데이터 유형	✓ 거버넌스, 액세스, 카탈로그, 제어	✓ 운영 애플리케이션에 결과 제공 및 사용
-------------------------------	--------------------------	------------------------	-----------------------	-------------------------

Edge-First, IoT & Analytics With Hybrid Cloud

- HPE Ezmeral Data Fabric Edge는 에지 장치에서 로컬 분석 처리를 위한 작은 설치 공간 구현
- HPE Ezmeral Data Fabric Pub-sub 데이터 스트리밍은 중앙 집중식 분석을 위해 스트림이 유지되는 에지-클라우드 데이터 통신에 선호되는 솔루션
- 모든 에지는 클라우드와 연결된 더 큰 패브릭의 일부임
- 커넥티드 카, 의료 기기, 석유 굴착 장치



Anadarko

시추 플랫폼에서 실시간 데이터 수집

- 중앙 집중 데이터 레이크와 결합된 에지 처리
- 데이터레이크에 대한 GPU 가속화 분석

Tüpraş

정제 공정에서 가동 중단 시간 감소

- 파이프 라인 부식을 줄여 생산 최적화
- 200,000 개의 센서
- 30 초가 아닌 초당 데이터 수집
- 가동 중지 시간 감소

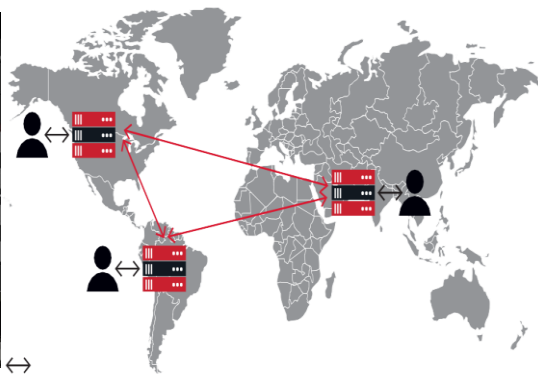
비즈니스 가치: 수익성 향상, 비즈니스 프로세스주기 시간 단축, 가동 중지 시간 감소

Why HPE Ezmeral Data Fabric Wins

✓ 수 많은 API, 소스로 부터 데이터 수집 및 통합	✓ 분석, ML, AI를 위한 광범위한 처리	✓ 보안을 위한 멀티테넌트, 데이터 유형	✓ 거버넌스, 액세스, 카탈로그, 제어	✓ 운영 애플리케이션에 결과 제공 및 사용
--------------------------------	--------------------------	------------------------	-----------------------	-------------------------

Edge-First, IoT & Analytics With Hybrid Cloud

- 기존 Hadoop 기반 솔루션 원백
 - 성능 한계 (비 일관된 응답 속도)
- 4개의 데이터센터 간 Clustering 구성 (Mirroring)
- 일관된 응답 속도(0.2초) 성능 보장
- 10개 지문, 2개 홍채, 디지털 사진
 - 매일 백만건 이상 등록 및 1경건 이상 조회



- 취약 계층을 위한 정부의 의료 지원 프로그램의 부정수급 방지를 위해 모든 국민을 대상으로 한 인도 국민인증 생체 데이터베이스



- 정부 보조 프로그램의 부정수급 비율 20% 감소
- 현재 약 10억 이상의 생체 인증 등록

Why HPE Ezmeral Data Fabric Wins

✓ 수 많은 API, 소스로부터 데이터 수집 및 통합

✓ 분석, ML, AI를 위한 광범위한 처리

✓ 보안을 위한 멀티테넌트, 데이터 유형

✓ 거버넌스, 액세스, 카탈로그, 제어

✓ 운영 애플리케이션에 결과 제공 및 사용

HPE EZMERAL

Thank you

Maximizing Value from the HPE Ezmeral Data Fabric Data Platform “ According to IDC, HPE Ezmeral Data Fabric ROI is over **576%** “