



**Hewlett Packard
Enterprise**

빅데이터 처리를 위한 **HPE EZMERAL UNIFIED ANALYTICS** 소개

서지영(Ezmeral Pre-Sales)

2022.2.23

1. Unified Analytics는 왜 필요한가?

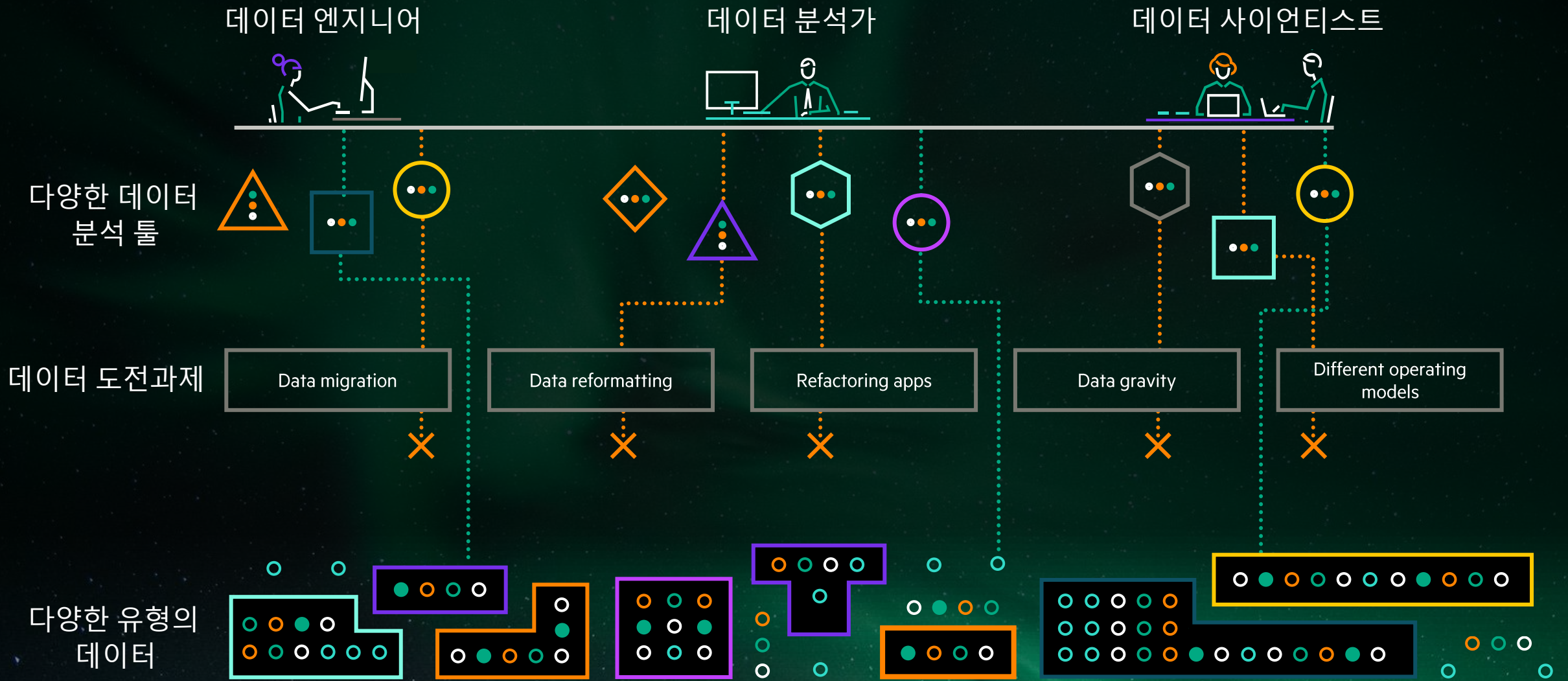
2. Unified Analytics란?

3. Ezmeral Unified Analytics란?

UNIFIED ANALYTICS는 왜 필요한가?



현재 기업의 과제



UNIFIED ANALYTICS란?



UNIFIED ANALYTICS란?



데이터센터



NAS



Lake



인프라스트럭처

Cloud

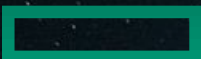


인프라스트럭처

Edge



인프라스트럭처



데이터 엔지니어



데이터 사용자



데이터 사이언티스트



데이터센터



NAS



Lake



인프라스트럭처

Cloud

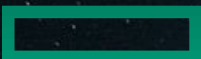


인프라스트럭처

Edge



인프라스트럭처



데이터 엔지니어



데이터 사용자



데이터 사이언티스트



셀프 서비스 앱 스토어

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처

데이터 엔지니어



데이터 사용자



데이터 사이언티스트



자동화

셀프 서비스 앱 스토어

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처

데이터 엔지니어



데이터 사용자



데이터 사이언티스트



자동화

셀프 서비스 앱 스토어

Delta Lake

Livy

Spark

Kubernetes

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처

데이터 엔지니어



데이터 사용자



데이터 사이언티스트



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Livy

Hive

Spark

Spark

Kubernetes

Kubernetes

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처

데이터 엔지니어



데이터 사용자



데이터 사이언티스트



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Jupyter NB

Livy

Hive

PyTorch

Spark

Spark

Spark

Kubeflow

Kubernetes

Kubernetes

Kubernetes

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처

데이터 엔지니어



데이터 사용자



데이터 사이언티스트



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Jupyter NB

Livy

Hive

PyTorch

Spark

Spark

Spark

Kubeflow

Kubernetes

Kubernetes

Kubernetes

파일/오브젝트 data fabric

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Jupyter NB

Livy

Hive

PyTorch

Spark

Spark

Spark

Kubeflow

Kubernetes

Kubernetes

Kubernetes

파일/오브젝트 data fabric

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Jupyter NB

Livy

Hive

PyTorch

Spark

Spark

Spark

Kubeflow

Kubernetes

Kubernetes

Kubernetes

HDFS

파일/오브젝트 data fabric

데이터센터



NAS



Lake



인프라스트럭처

Cloud



인프라스트럭처

Edge



인프라스트럭처



자동화

셀프 서비스 앱 스토어

Delta Lake

Presto SQL

Jupyter NB

Livy

Hive

PyTorch

Spark

Spark

Spark

Kubeflow

Kubernetes

Kubernetes

Kubernetes

NFS

HDFS

파일/오브젝트 data fabric

데이터센터



NAS

Lake

인프라스트럭처

Cloud

인프라스트럭처

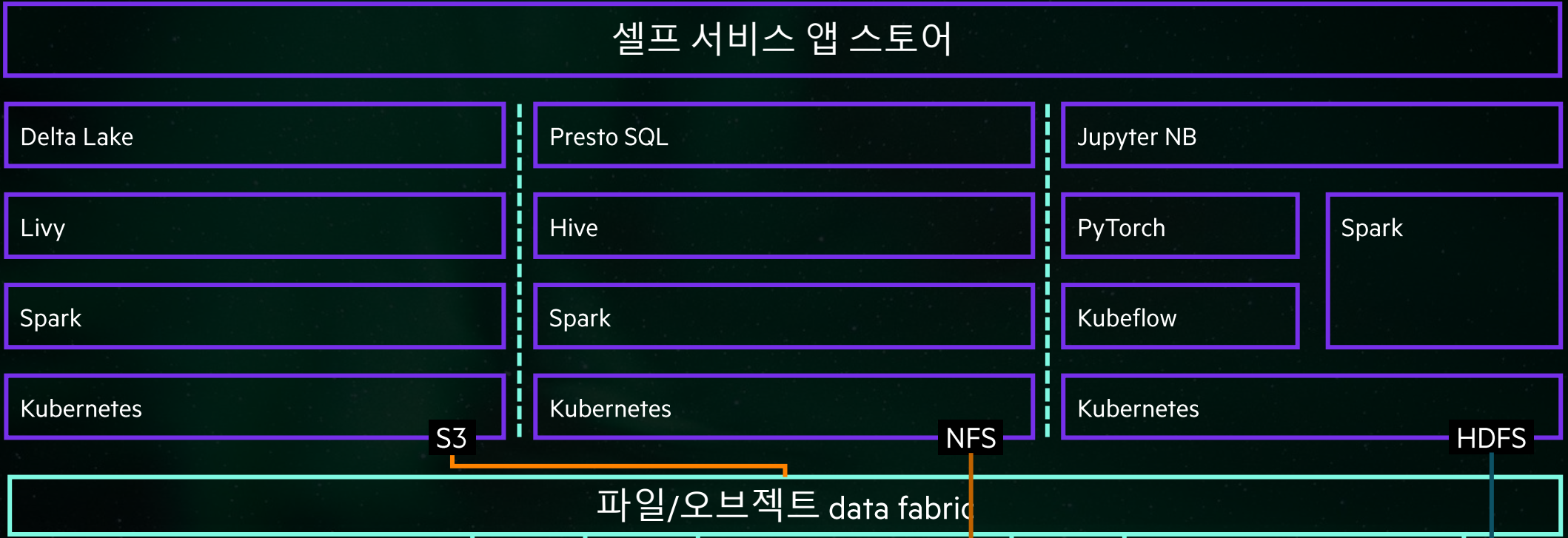
Edge

인프라스트럭처



자동화

셀프 서비스 앱 스토어



데이터센터



NAS

Lake

인프라스트럭처

Cloud

인프라스트럭처

Edge

인프라스트럭처

EZMERAL UNIFIED ANALYTICS란?



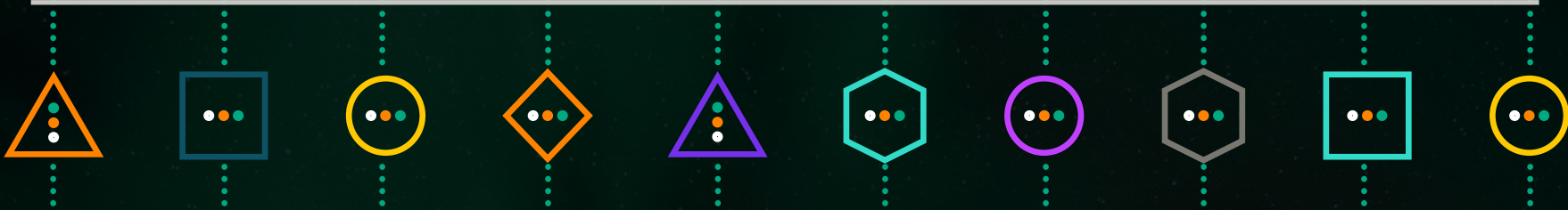
데이터 엔지니어



데이터 사용자



데이터
사이언티스트

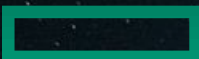
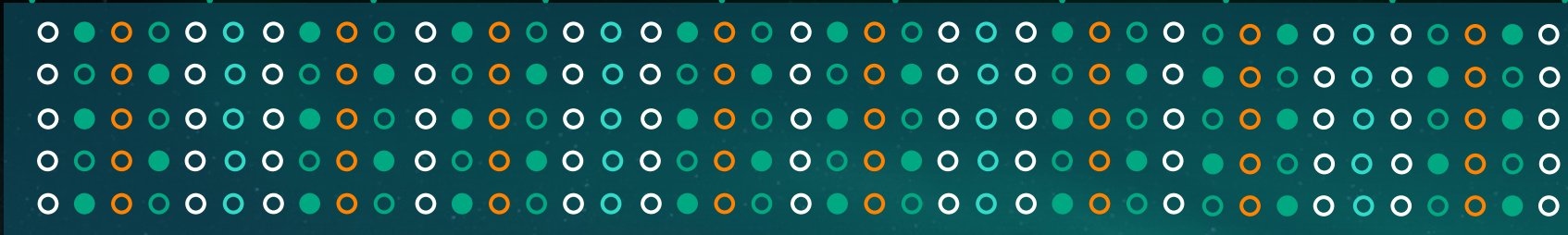


HPE EMERALD 데이터 웨어하우스+ 데이터 레이크

안전하고 통합된
데이터

데이터 분석팀의
생산성 향상

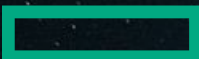
종속성은 회피하면서
혁신은 가속화



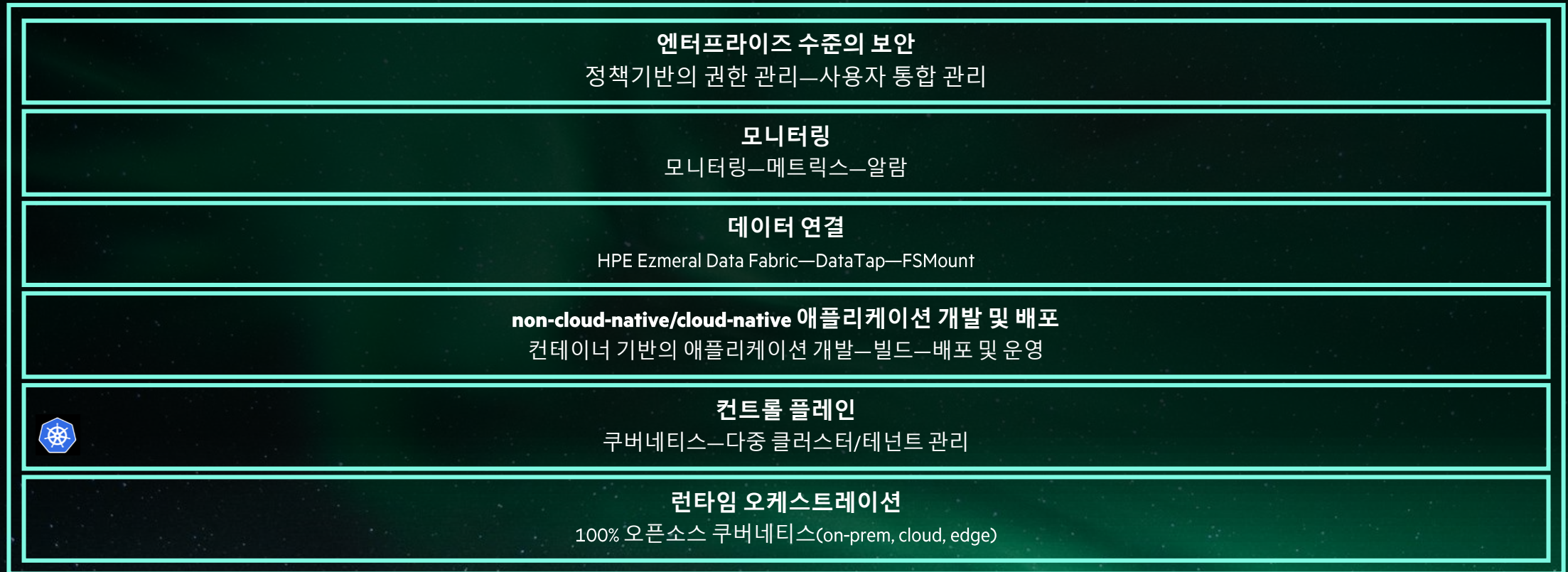
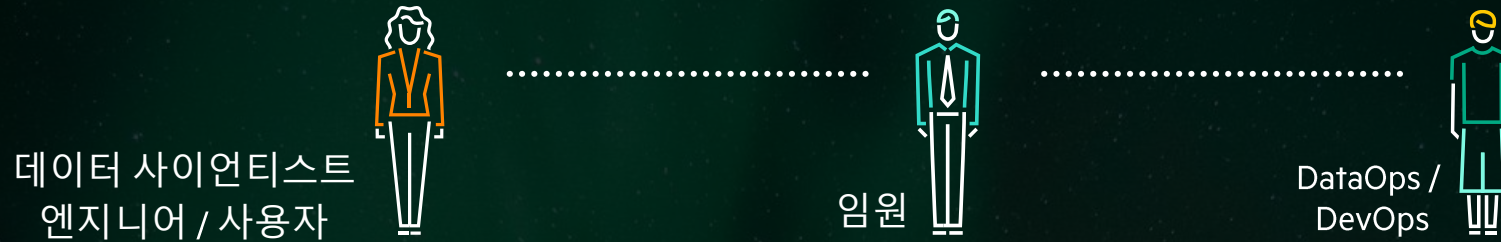
HPE EZMERAL UNIFIED ANALYTICS 아키텍처



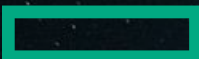
Edge Core Cloud



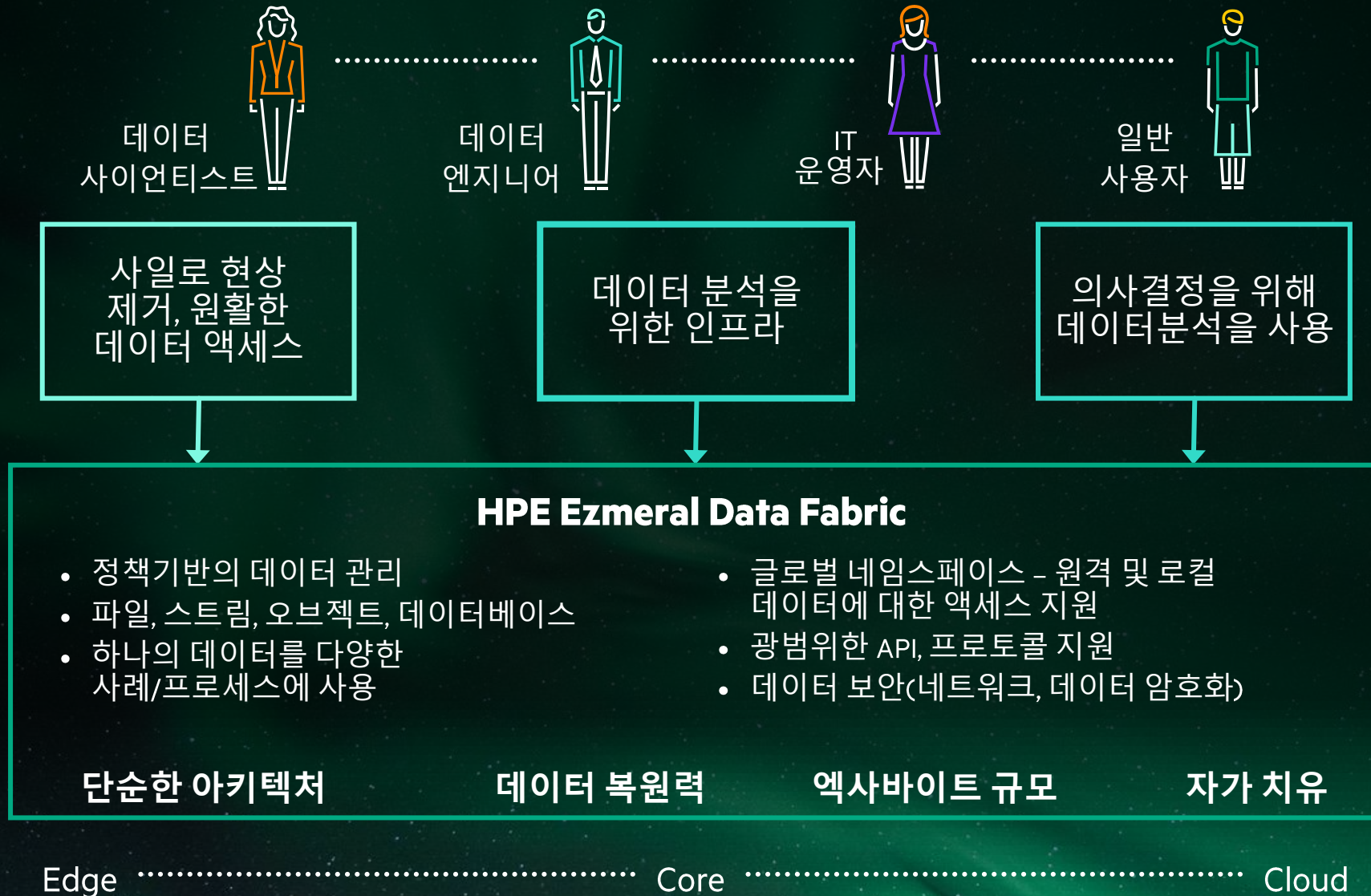
HPE EZMERAL 런타임 아키텍처



Edge Core Cloud



HPE EZMERAL DATA FABRIC

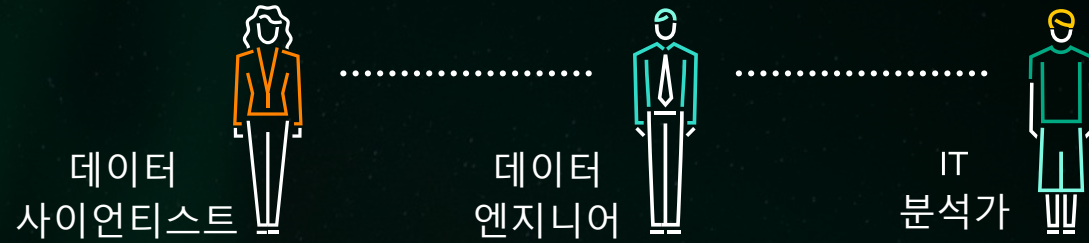


분석 엔진 및 툴

HPE Ezmeral Unified Analytics

분석 도구에 대한 통합 접근 방식:

- HPE Ezmeral ISV 마켓 플레이스
- 자체 오픈 소스 스택 가져오기
- App Workbench로 나만의 빌드 구성
- 통합적 분석도구를 지원함으로써 사일로 현상 제거.



원 클릭, 셀프 서비스 앱 스토어

Opinionated stacks HPE Ezmeral ISV Marketplace Bring/build-your-own open source stack



SPARK 통합

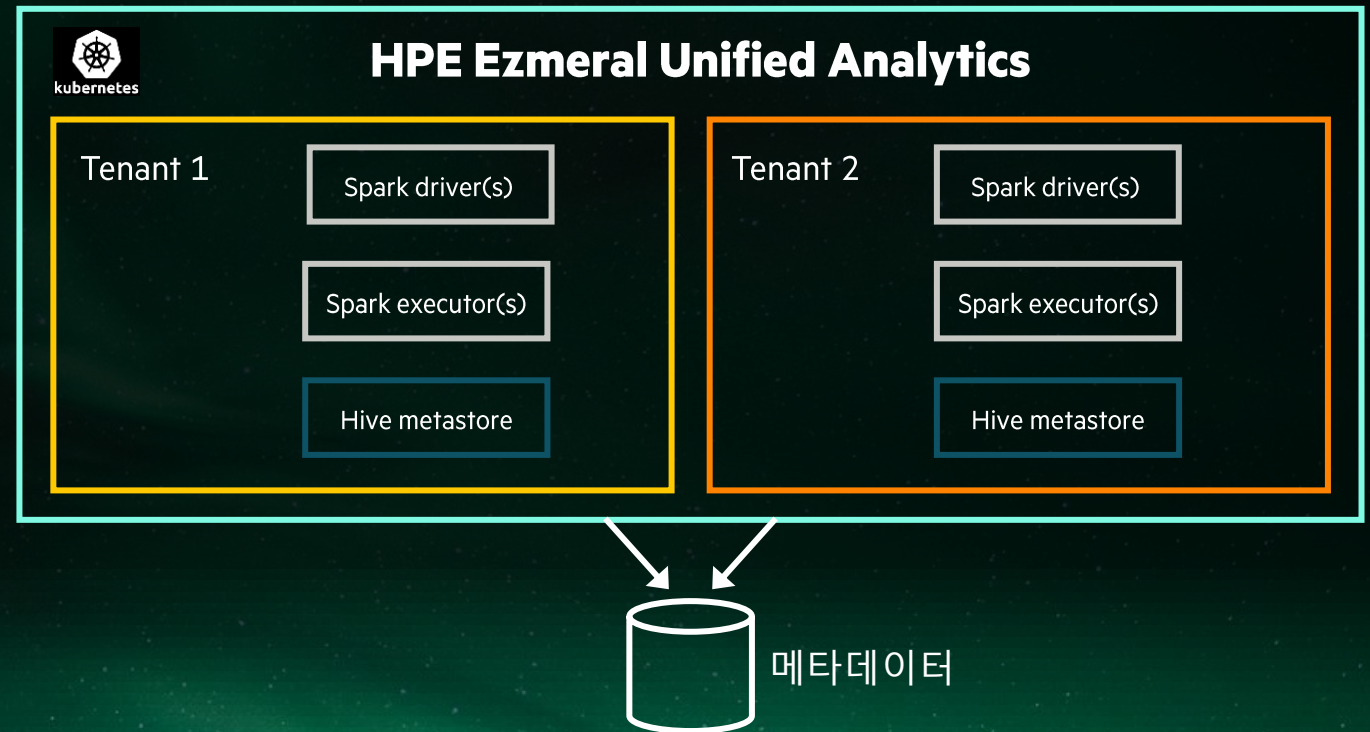
데이터 분석가, 사이언티스트 및 엔지니어는 데이터 분석을 위해 Apache Spark를 선호

- 다양한 Spark 버전을 실행
- 24x7 기술지원
- NVIDIA® GPU로 작업 가속화
- Delta Lake와의 통합 인터페이스 제공
- MLOps, AirFlow, ML Flow 및 Kubeflow와 통합
- 쿠버네티스에서 Spark 3.x operator 지원

HIVE 메타데이터

HPE Ezmeral Unified Analytics

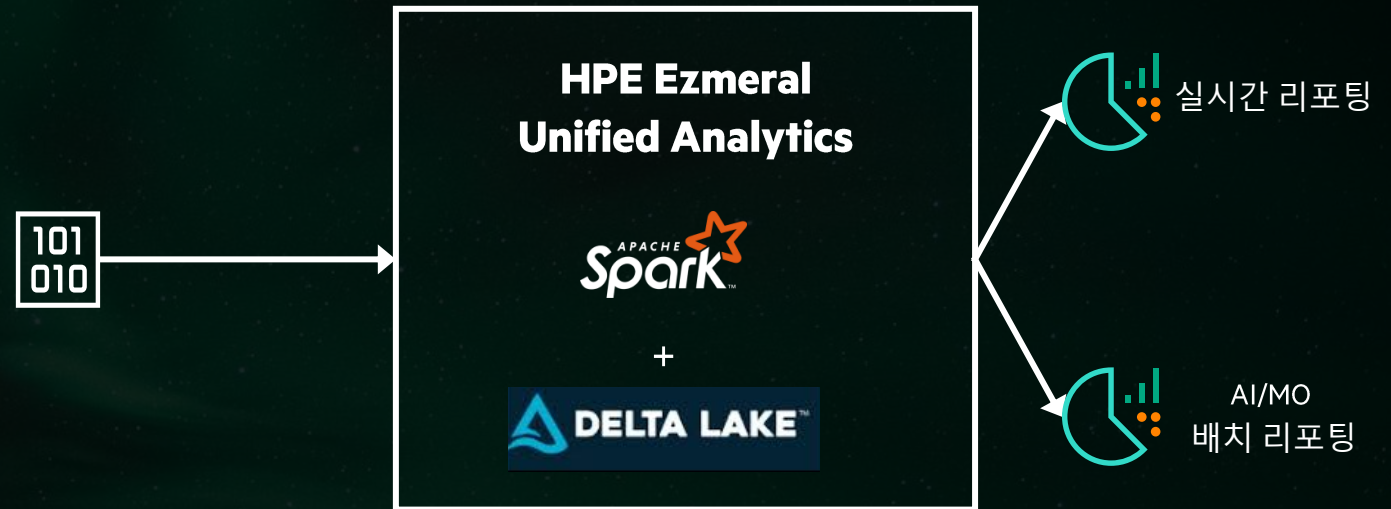
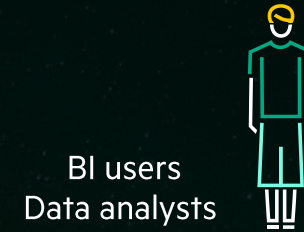
- 메타데이터 공유 저장소
- 다중 테넌트 간의 협업 향상
- 편리한 리소스 관리



배치 및 스트리밍 데이터 처리

HPE Ezmeral Unified Analytics

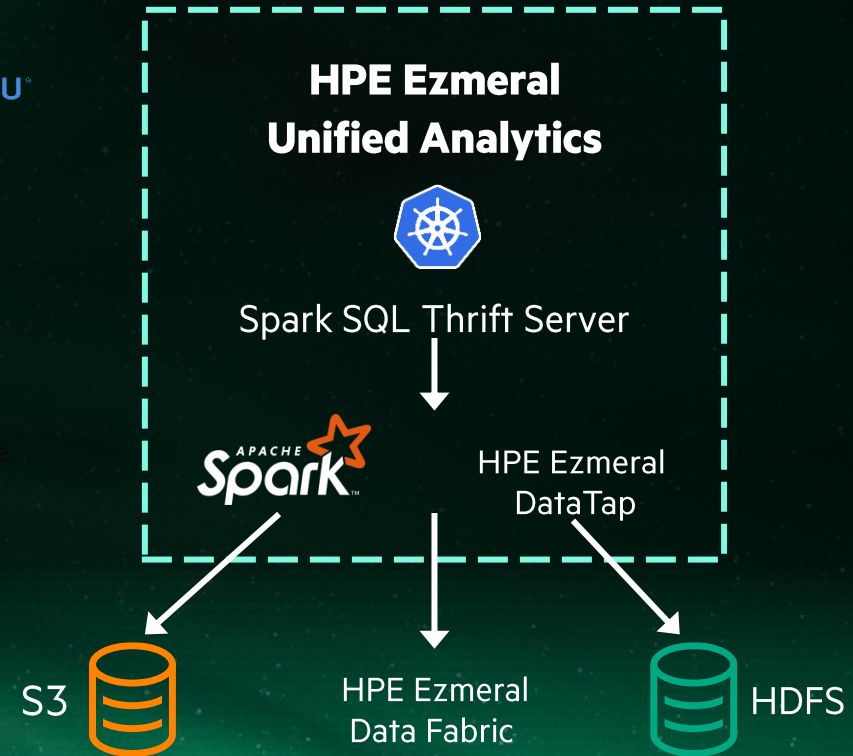
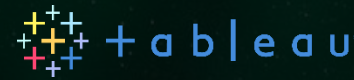
- 통합 인터페이스로 배치, 스트리밍 워크플로 지원
- 일관되고 안정적인 데이터를 위한 ACID 트랜잭션 지원



앱 및 데이터 연결

HPE Ezmeral Unified Analytics

- HPE Ezmeral을 사용하여 Spark가 액세스하는 데이터를 호출하고 활용
- HPE Ezmeral DataTap은 다양한 인터페이스 또는 스토리지(HDFS, NFS)에 대한 액세스를 제공
- Thrift Server를 사용하여 BI 앱과 Spark 작업을 연동



HPE EMERALD

데이터 레이크와 데이터
웨어하우스를 통합
플랫폼으로 연결



THANK YOU



a00118475enw

CONFIDENTIAL | AUTHORIZED HPE PARTNER USE ONLY