

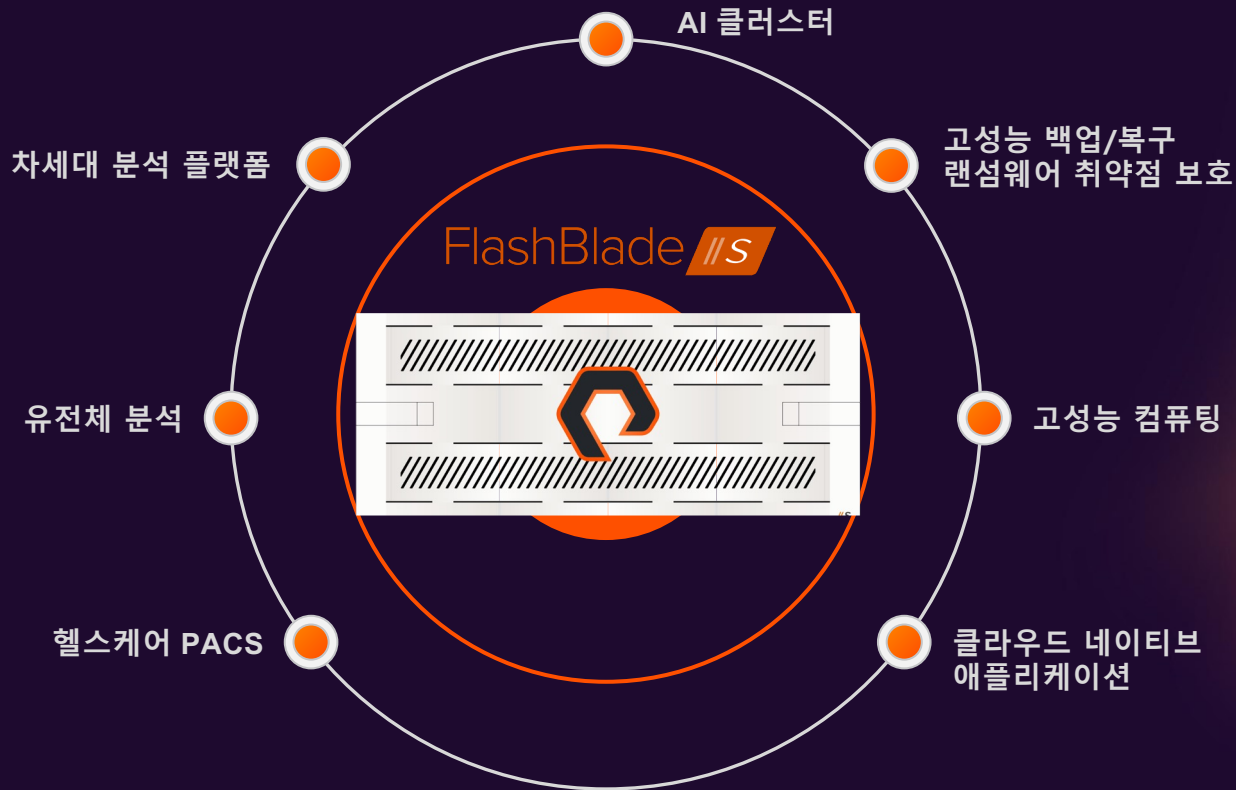
FlashBlade//S

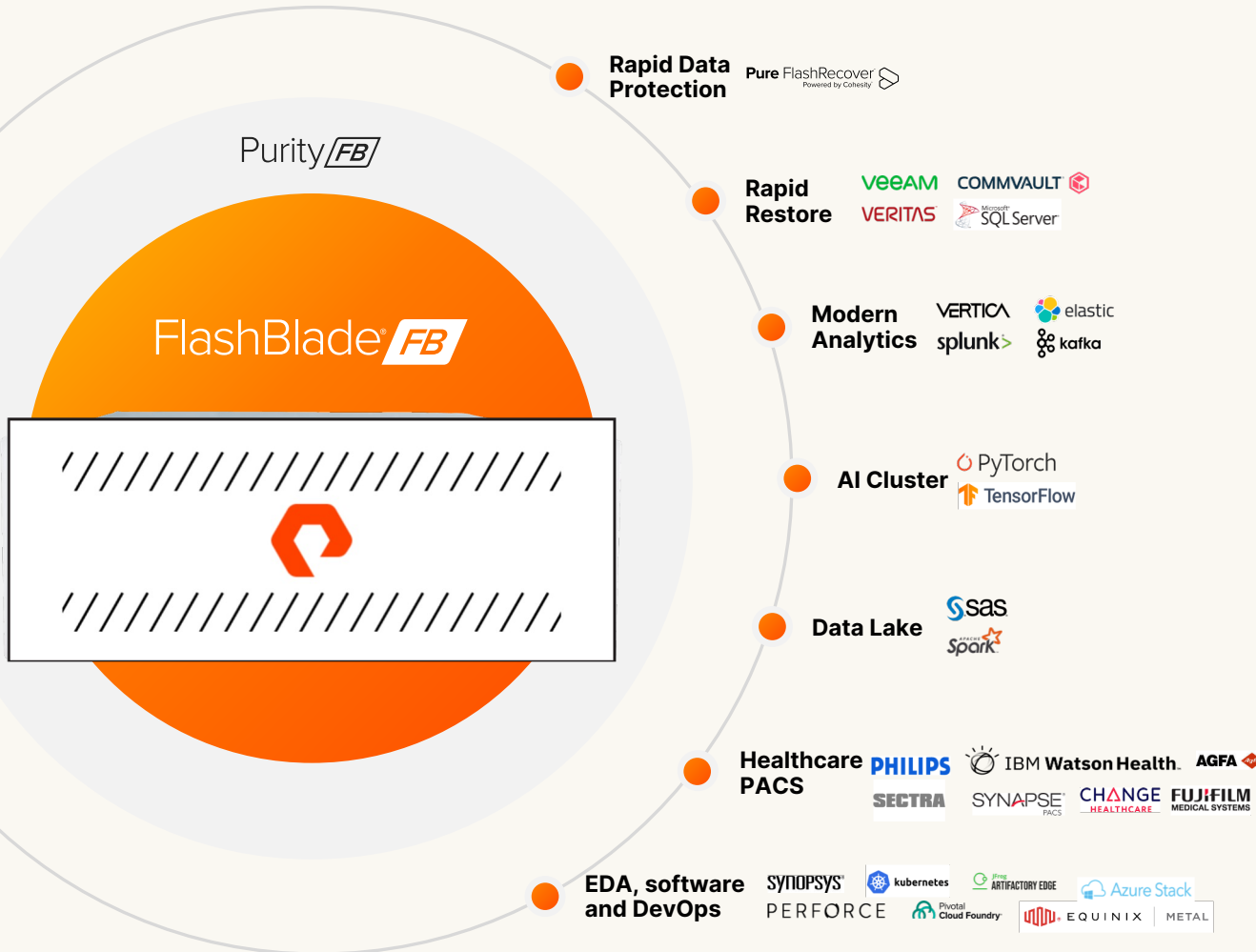
타겟 시장 및 고객 사례 소개

김민철 상무
퓨어스토리지

다양한 고객 환경 통합

개별 워크로드 사례에 대한 전용 스토리지 사일로를 제거





검증된 다양한 연계 솔루션

FlashBlade를 사용하면 최신 데이터 응용 프로그램을 확장 가능한 단일 플랫폼에 통합할 수 있습니다. 복잡하고 비효율적인 인프라 사일로를 제거합니다. 새로운 수준의 투자 보호를 달성하십시오.



Pure Storage Partners with Meta on AI Research SuperCluster (RSC)

Pure FlashArray 및 FlashBlade는 RSC를 위한 강력하고 확장 가능한 스토리지 솔루션을 제공하며, Meta는 글로벌 최대 규모의 가장 빠른 AI 슈퍼컴퓨터 중 하나입니다.

1세대(2017년) 인프라를 기반으로 구축된 RSC는 생산 훈련 속도를 20배 높이고, 완전히 구축되면 엑사바이트 규모로 확장할 계획입니다.

Meta 주요 고려 사항

- 성능 - 16TB/s of training data to server to GPU compute
- 확장성 - 수백 PB에서 EB 레벨의 하이퍼스케일로 활성 데이터 확장성
- 그린IT(ESG) - 한정된 데이터센터 규모에서 높은 상면 집적도와 전력, 냉각 효율성 향상
- 신뢰성 - 프로덕션을 오래 유지하고 임시 연구원 사용을 지원하는 중요한 인프라
- 보안 - 초기부터 보안 및 개인 정보 보호를 위해 설계된 RSC

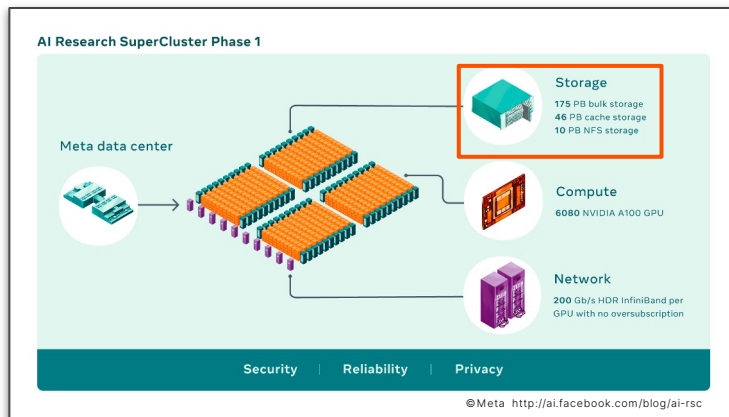
Better, Faster, Smaller, and More efficient than the alternatives



Pure Storage Partners with Meta on AI Research SuperCluster

Pure FlashArray 및 FlashBlade는 RSC를 위한 강력하고 확장 가능한 스토리지 솔루션을 제공하며, Meta는 글로벌 최대 규모의 가장 빠른 AI 슈퍼컴퓨터 중 하나입니다.

- 구성 환경
 - 760 NVIDIA DGX A100 시스템, 6080 GPUs
 - 10PB FlashBlade(File/Object) + 175PB FlashArray//C(Block/file)
 - 5 FlashBlade Chassis(storage pod) + 15 A100 서버(Compute pod)
 - **A100 서버 ↔ FlashBlade간 이더넷 구성**
- 도전과제
 - 대량의 Write 발생 시, 기존 **Cache** 구조의 스토리지 성능 저하 발생
 - DGX 노드가 RACK 단위로 구성으로 인해 **제한적인 시스템 확장**
 - AI 모델링 시간 단축을 위한 **메타데이터 가속 전용 솔루션** 필요
- 도입 효과
 - **Key-Value 데이터 관리 아키텍처로 성능 저하 방지**
 - **Storage Pod** 별도 구성으로 **선형성 확장성 보장**
 - **Pure RapidFile Toolkit** 을 통한 **메타데이터 가속 제공**



[Reference Link]

<https://ai.facebook.com/blog/ai-rsc>

<https://www.purestorage.com/kr/company/newsroom.html>



MLops 사례

이미지 처리 기반 AI/ML 전용 클라우드 플랫폼

도입 목표

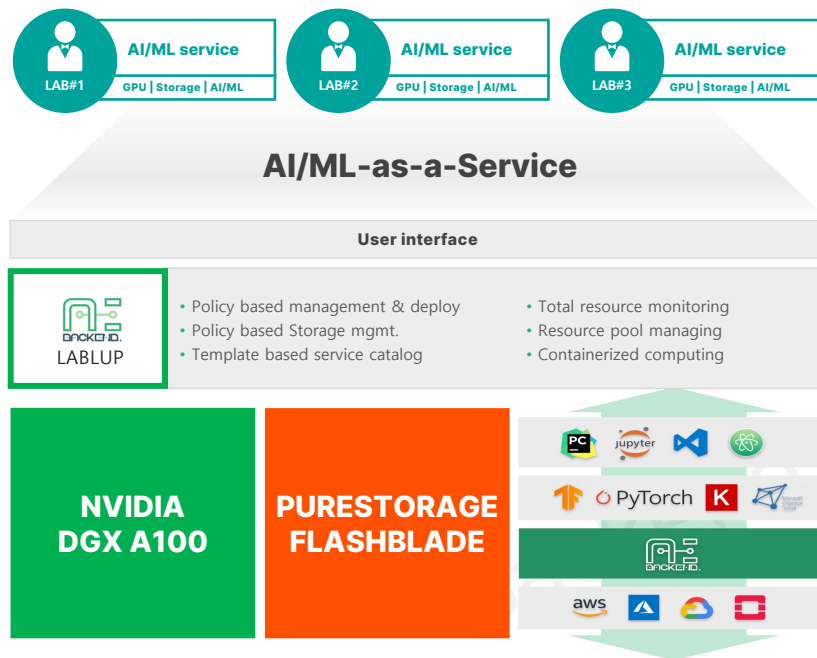
- 워크로드의 반복적인 특성을 지원하기 위한 유연하고 민첩한 AI 인프라 환경 구축
- 개발자가 AI/ML 인프라를 쉽게 사용할 수 있도록 AI-as-a-Service 형태의 AI 전용 클라우드 구축

업무 요구사항

- 기존 2주 이상 소요되는 Learning and Inference Job을 2~3일 내 완료
- AI 플랫폼과 스토리지 간의 운영 환경을 통합하여 단일화된 관리 환경 제공
- 기존 GPU 리소스 재활용

FLASHBLADE 도입 개선 효과

- A100 GPU 활용률을 개선하여 6배 이상의 성능 개선
- Silo 되어있는 그룹내 AI 플랫폼 통합
- AI 플랫폼과의 API 연동으로 관리 통합 및 운영 간소화





업계 리더들의 협력:

- 퓨어스토리지 FlashBlade//S
- 엔비디아 DGX Server Family
- 엔비디아 Spectrum Networking

차세대 AI Platform:

- AI를 위한 전체 스택 사전 통합
- AI 인프라 단순화
- 확장성 및 효율성 극대화



사용 및 관리의 용이성

AI 데이터 파이프라인의 모든 단계를 위한 단일 통합 솔루션으로, 배포, 관리 및 사용이 간편한 간단한 아키텍처로 제공

환경 효율성

ESG 친화적: 고밀도, 전력 소비 감소 및 업계 최고의 데이터 센터 효율성을 제공하는 더욱 강력한 AI 인프라

모든 규모의 성능

적게 시작하고 고객 환경 및 요구사항의 필요에 따라 원활하게 확장하고 적용

애자일 AI 플랫폼

AI 환경과 고객 요구 사항이 진화함에 따라 퓨어와 엔비디아의 미래 혁신을 위해 제작

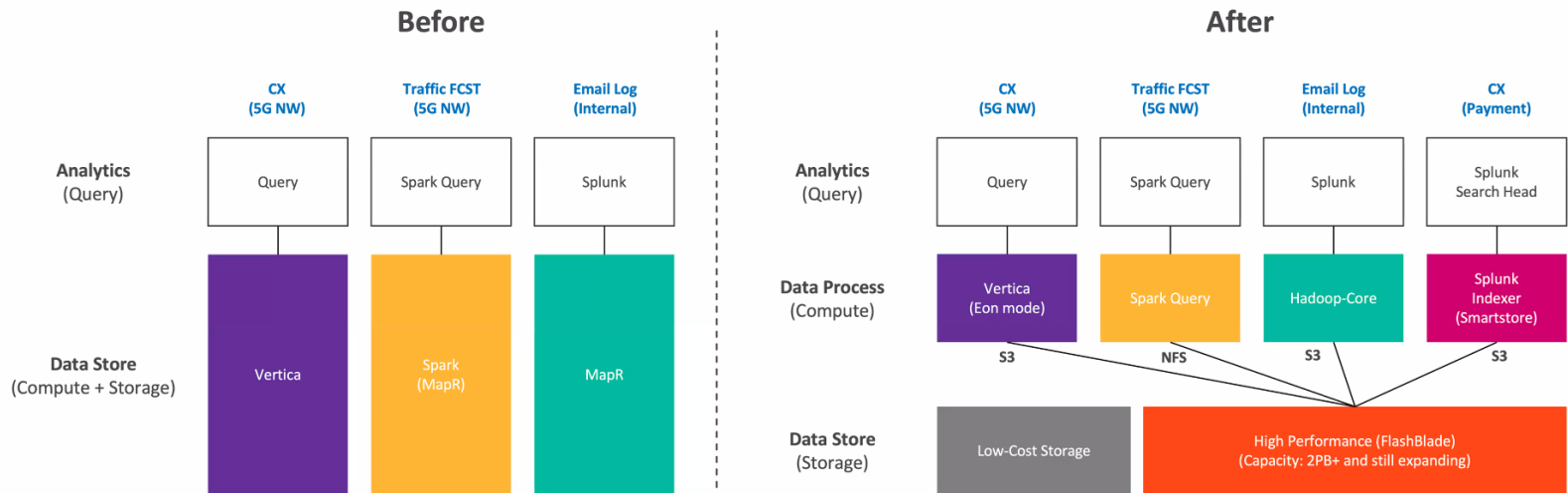
신뢰할 수 있는 업계 리더

시장의 혁신 리더들이 합류하여 2019년부터 최첨단 플래시 기반 AI 인프라를 제공



데이터 분석 통합 사례1

분리된 데이터 레이크 환경에 대한 UFFO 기반 단일 플랫폼 통합



도전 과제

- DAS 구성으로 인한 성능 및 용량 확장성 한계 도달
- 복잡한 운영 환경 및 노드 장애에 대한 서비스 문제 발생
- 지속적인 데이터 증가로 인한 고성능-확장성 스토리지 필요

개선 효과

- 컴퓨트-스토리지 분리를 통한 PB 수준의 확장성, 고성능 기반 성능 확장
- 간편한 관리 및 빠른 노드 장애 대처 가능
- 다양한 데이터 분석 애플리케이션에 대한 단일 플랫폼으로의 통합



데이터 분석 통합 사례2

도전과제

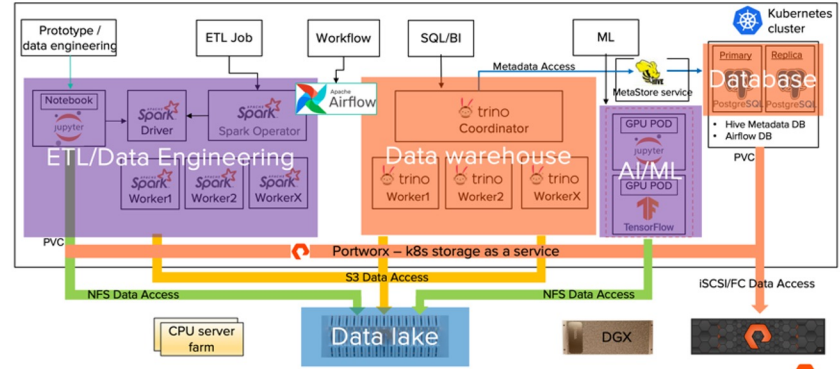
- Legacy tech debt with Cloudera CDH5
- Expensive but low-value support
- Infrastructure silos: CDH5, Elastic, and TensorFlow
- Complex operation on infra and Hadoop
- Clients also need more complex analyses to quickly model the effectiveness of potential marketing programs.

퓨어스토리지 솔루션 - ANALYTICS & AI/ML AS-A-SERVICE

- **All open source:** Apache Spark, Trino, Elastic, TensorFlow, Kafka and PostgreSQL on **FB, FA & PSO**
- **Spark, Trino, Jupyter, TensorFlow on Kubernetes and Portworx**

개선효과

- 최신 오픈소스 스택 적용가능, as-a-service 구성으로 구성 관리가 용이해짐
- TCO개선 효과
- 복잡한 하둡관리



Simple, scalable and open-source Spark, data warehouse and AI with Pure



데이터웨어하우스 사례 (1/2)

Vertica ENT(DAS구성) 에서 Flashblade 고성능 S3 를 통한 Vertica Eon 모드로 구성 변경

📁 데이터 웨어하우스 고도화

도입 목표

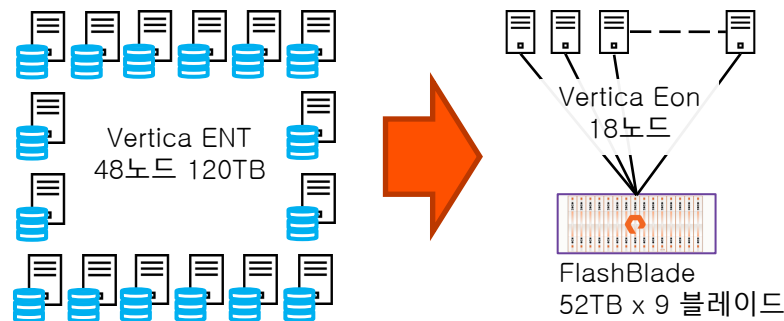
- 기존 버티카 환경의 운영 안정성 개선
- 운영 비용 절감 및 컴퓨팅과 스토리지 분리
- 분리되어 있는 DW 업무 통합

업무 요구사항

- 하드웨어 장애에 대한 일관된 성능 및 빠른 복구
- 유연하고 민첩한 노드 확장

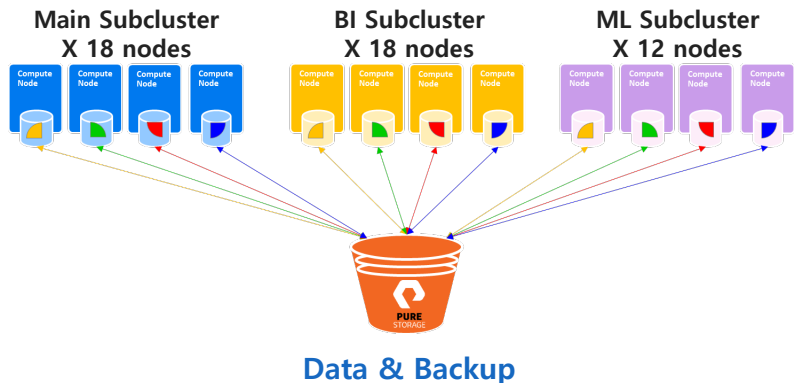
FLASHBLADE 도입 개선 효과

- 컴퓨팅 리소스 활용률 65% 개선(기존 48노드 → 18노드로 절감)
- 상면, 전력, 냉각 비용 절감 2배 절감
- 유연한 스케일아웃 데이터허브 플랫폼 구축



데이터웨어하우스 사례 (2/2)

컴퓨팅과 스토리지 분리 아키텍처를 통해 유연한 컴퓨팅 자원 활용 기반을 구축하였고, 마케팅, 빅데이터, 마이데이터, 머신러닝 등 분석 업무 영역으로 확장 진행중.



- Data 분석에 최적화된 Analytic DW 구축
- Cluster 분리를 통한 다양한 workload 대응 가능
- 하이브리드 클라우드 연계 플랫폼 확보
- 비즈니스 독립성 확보 (SLA 지연 리스크 제거)
- Sub Cluster #1: Data 분석가 Play-ground 제공
- Sub Cluster #2: Vertica Data 활용한 AI학습 지원
- One DB 구축을 통한 비용 절감 및 관리 효율성 개선

Thank you !

✓ 공식 웹사이트

www.purestorage.com/kr

✓ 공식 유튜브

www.youtube.com/c/PureStoragekr

✓ 공식 페이스북

www.facebook.com/purestoragekorea

✓ 네이버 블로그

blog.naver.com/purestorage_korea

