

# 데이터 통합의 어려움을 해소하고, 데이터 혁신을 이끌어갈 차세대 스토리지 플랫폼 FlashBlade//S 소개 (Lablup 구축 사례 소개)

송성환 부장

퓨어스토리지 코리아



# What's happening in the world?



4x

비정형 데이터는  
2026년까지 4배 증가

(GARTNER)



60%

재구매와 재구축으로 인해  
기업은 스토리지 플랫폼 원래  
예산의 60%를 소모

(GARTNER)



54%

기업의 54%는 데이터 센터에  
소비 기반 클라우드 구독  
모델을 적용 선호

(ESG)

# What do organizations need?

이상적인 스토리지 플랫폼을 설계하는 것은 계속 진행 중.



제한 없는  
확장성



엔터프라이즈  
플랫폼



다양한 데이터  
서비스

← 모든 환경에서 배포와 사용이 간편해야 함 →

# PureStorage 철학

*“소프트웨어의 중요성을 아는 사람은  
전용 하드웨어를 만들어야 합니다.” – Alan Kay*



하드웨어와 소프트웨어의 혁신은 효율성과 확장성을 촉진



# 유니파이드 고성능 파일&오브젝트 스토리지(UFFO)

스마트 폰이 통합되는 것처럼 :

UFFO 스토리지 플랫폼 통합:



Navigation



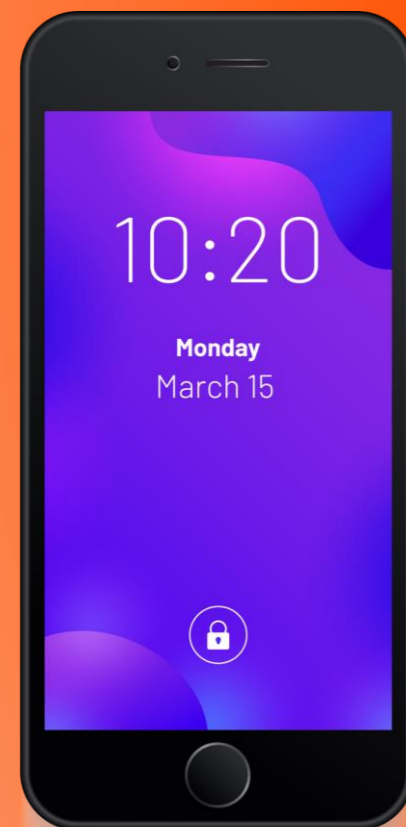
Camera



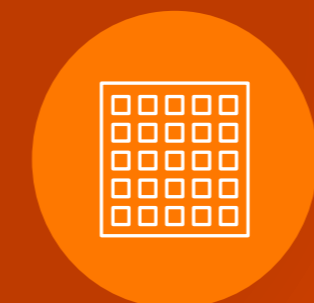
Web  
Browser



Phone



Wallet



Scale-out  
NAS



Network  
Attached  
Storage (NAS)



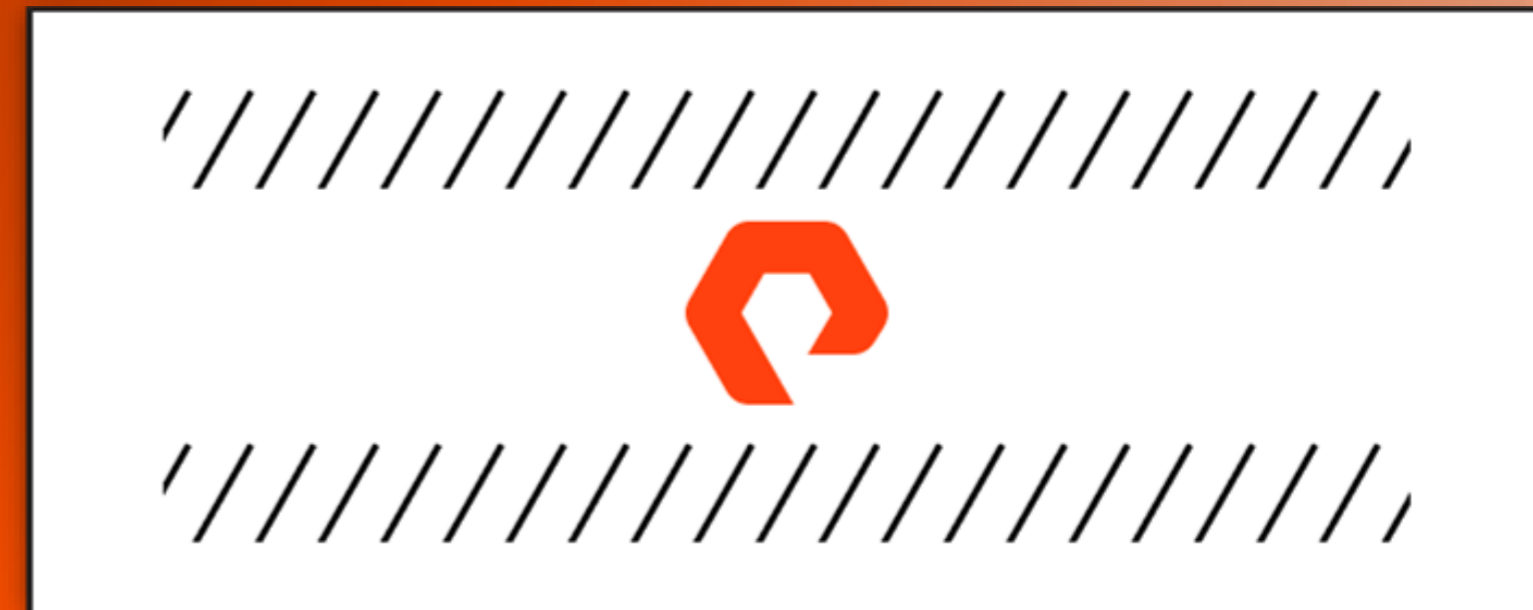
Direct Attached  
Storage (DAS)



Backup  
Appliance



Fast Object





# 5+ Years of FlashBlade Innovation



Industry's  
First  
UFFO

SafeMode  
& Rapid  
Recovery

10X  
Increase  
in Scale



Uncomplicate Data Storage, Forever

# FlashBlade//S

//Sscalability  
//Sspeed  
//Ssimplicity  
//Ssustainability





## 확장성 (Scalability)

완벽한  
스케일 업&아웃

최대 2PB  
per chassis

최대  
10컨트롤러  
per chassis

## 고성능 (Speed)

>2X 성능

다이렉트 플래시

RapidFile

## 단순함 (Simplicity)

모듈형 아키텍처

Pure1

AI 레퍼런스  
아키텍처

## 지속 가능성 (Sustainability)

Future Proven

Evergreen

ESG Ready

# FlashBlade//S

Scalability, Speed, Simplicity, Sustainability

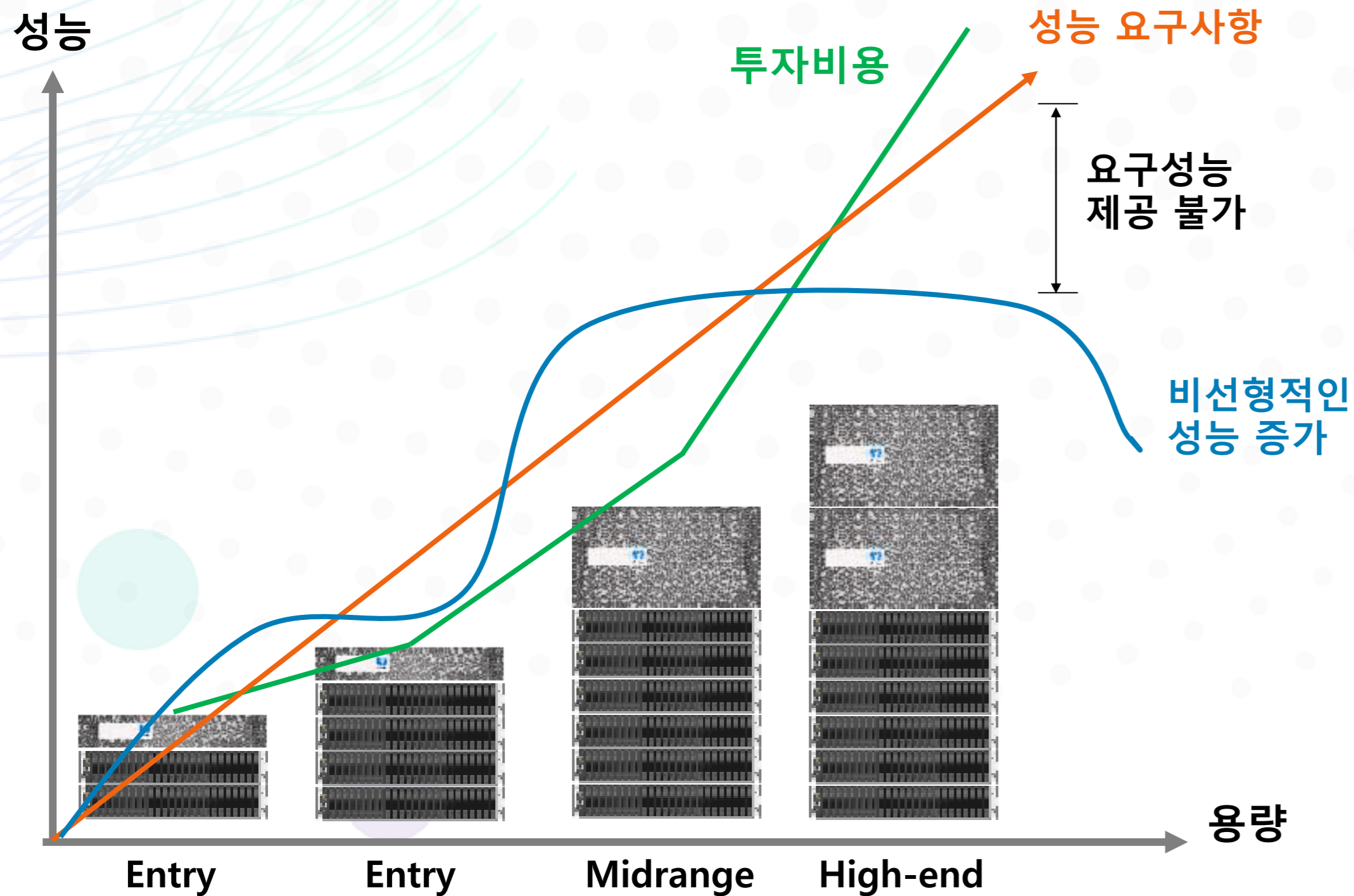


# 확장성 (Scalability)

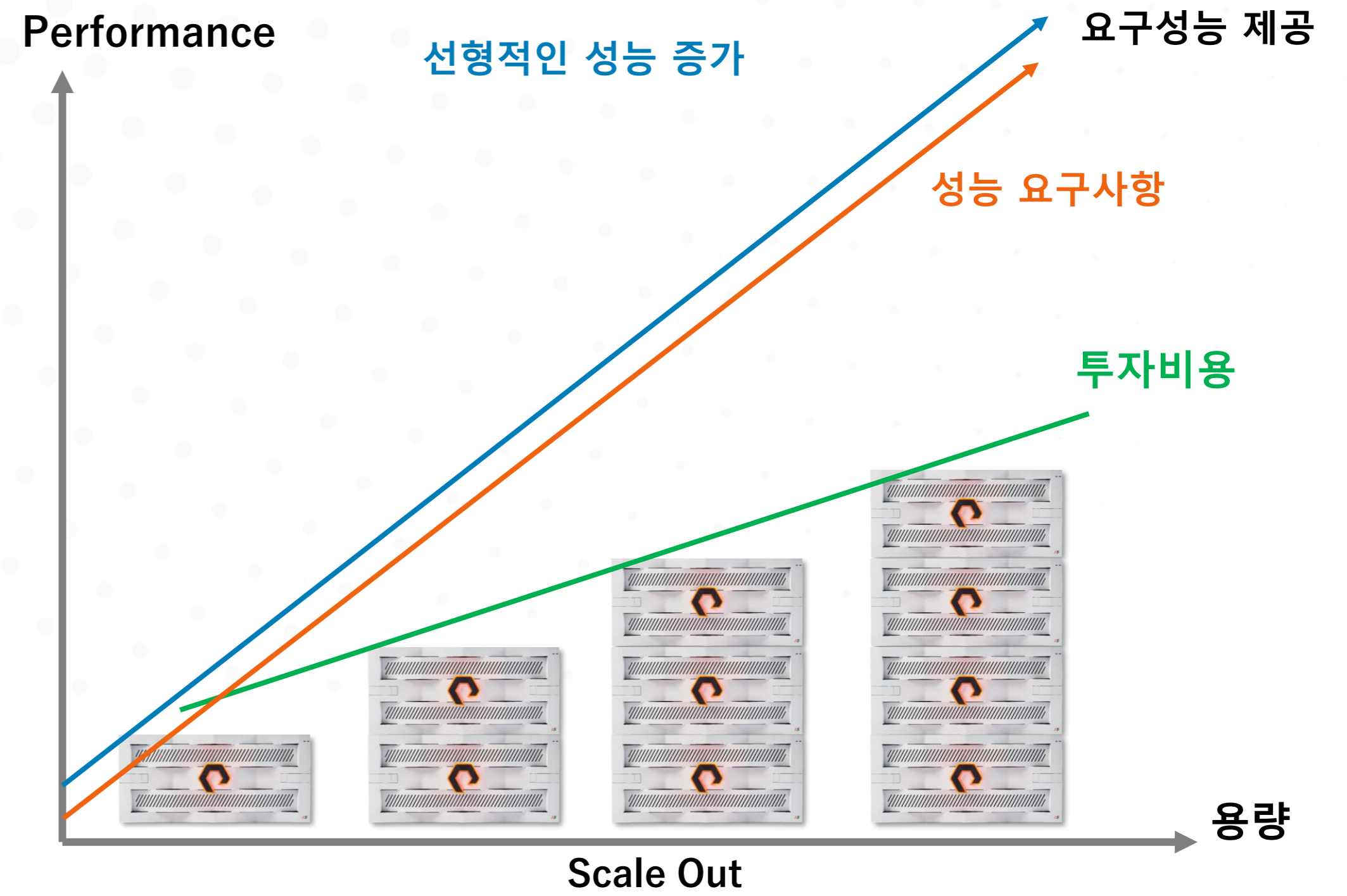
성과 용량을 자유롭게 확장할 수 있는 완벽한 스케일 업&아웃 아키텍처

타사 스토리지

Pure FlashBlade//S



- 예측이 어려운 비선형적인 확장
- 투자 예측의 어려움
- 여러 세대 모델간 컨트롤러 및 스토리지 확장 제약



- 예측이 가능한 선형적인 확장
- 효율적인 투자 가능
- 자유로운 컨트롤러 and/or 스토리지 확장



# 고 성능 (Speed)

플래시 최적화된 Pure만의 소프트웨어와 하드웨어로 업계 최고 성능 제공



## 2X

샤시당 >30GB/s 고성능

소프트웨어 최적화된  
NVMe 다이렉트 플래시 모듈

메타데이터 가속 솔루션 (RapidFile)

 DirectFlash<sup>®</sup> SOFTWARE +  DirectFlash<sup>®</sup> MODULE



**4X**  
더 높은  
성능 집적도

**3X**  
대역폭 증가

**20%**  
용량 효율성  
증가

**2X**  
용량 집적도  
증가

### Adaptive I/O Control

Flash-level에서 **일관된 고성능** 데이터 접근 보장

### Smart Endurance

플래시 수명연장, 오버프로비저닝 감소 → 도입 비용 절감

### Predictive Resiliency

Flash Level의 지속적인 모니터링으로 **높은 복원력** 제공



# 단순함 (Simplicity)

모듈형 아키텍처, AI 기반 통합 관제 시스템, 레퍼런스 아키텍처



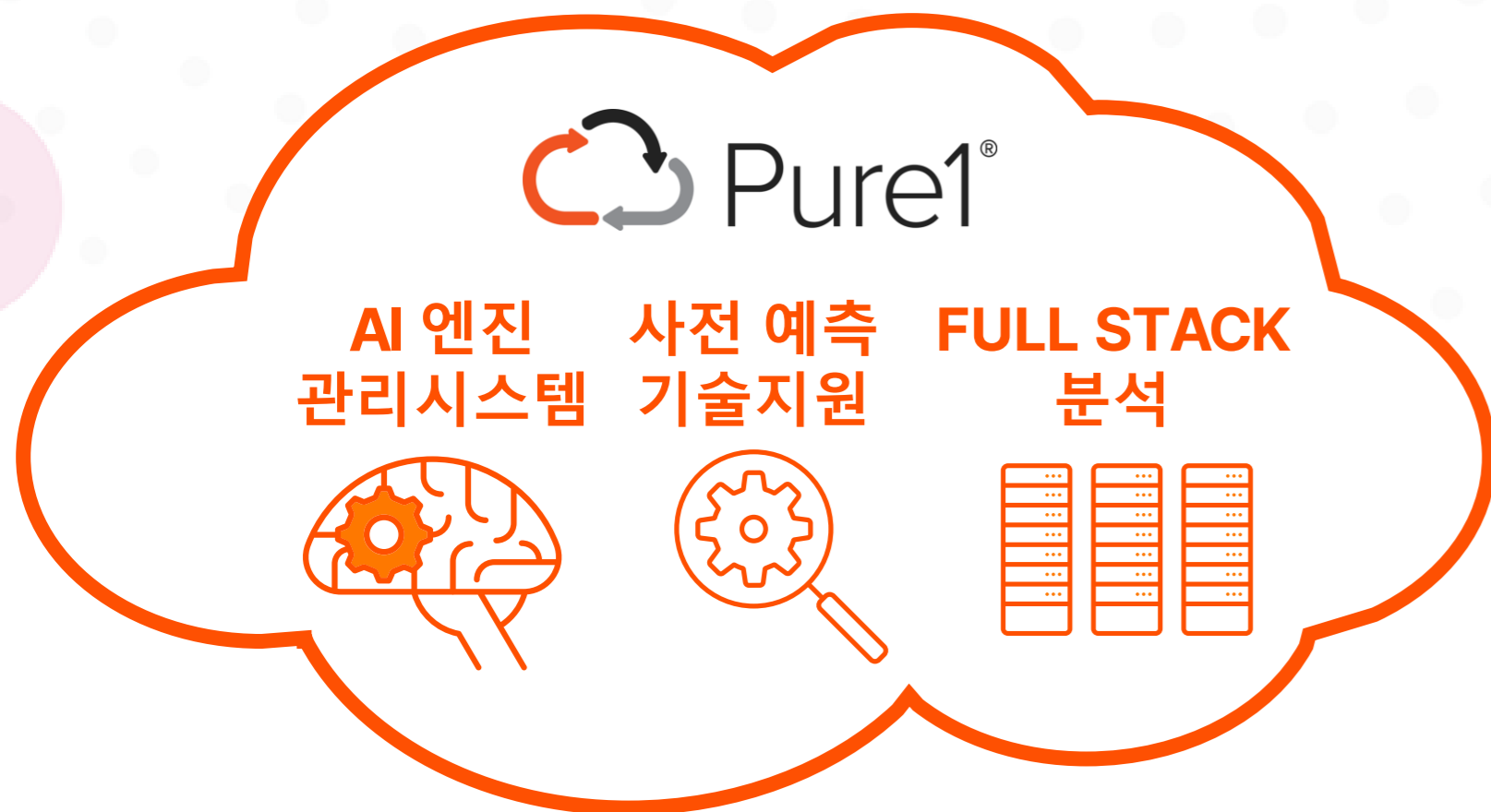
컨트롤러+다이렉트  
플래시+네트워크 통합 구성

쉽고 빠른 교체 및 확장/축소가  
자유로운 모듈형 하드웨어

직관적인 관리 대시보드 및 API

AI 기반 클라우드 통합 관제 (Pure1)

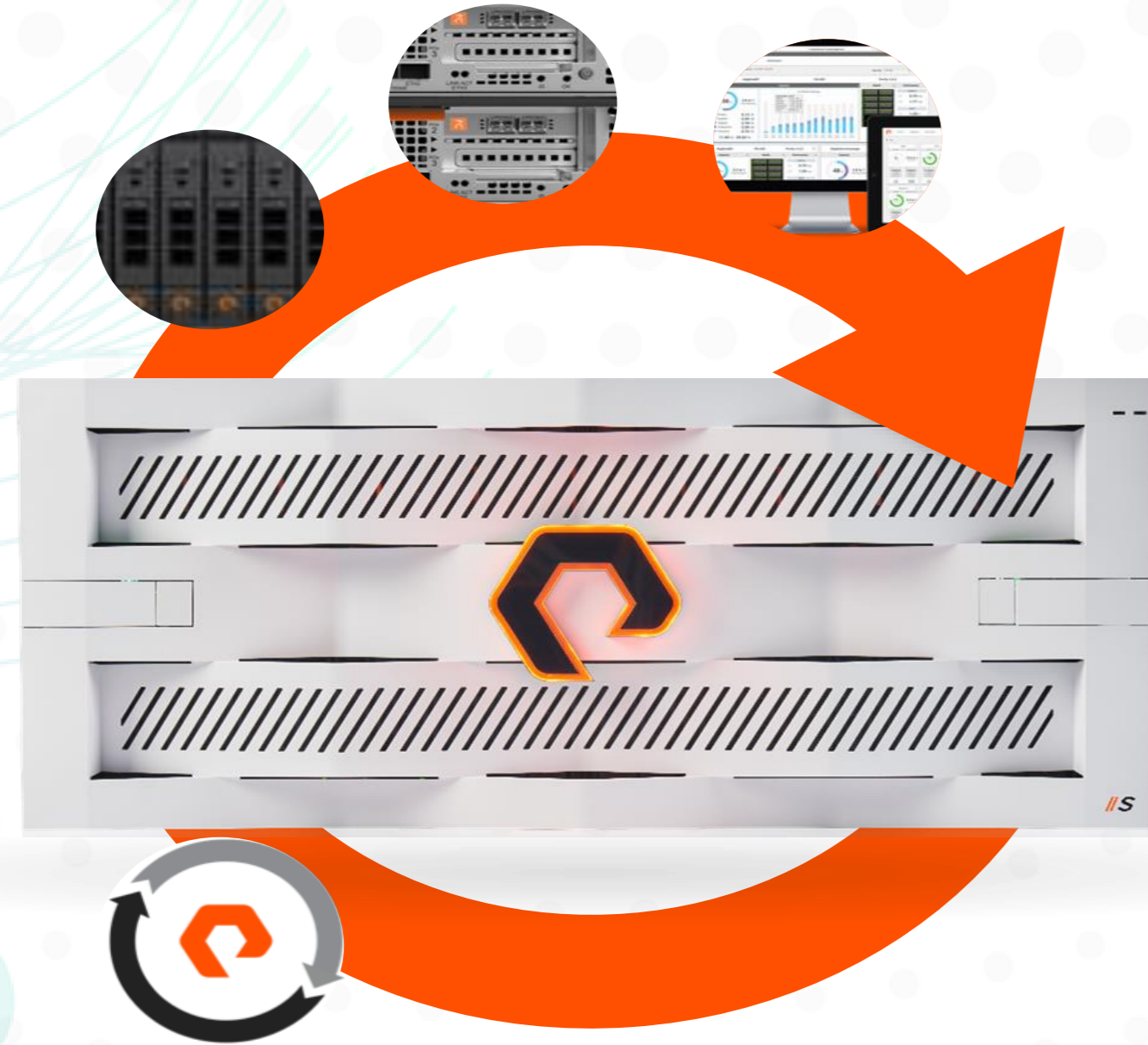
AIRI//S





# 지속 가능성 (Sustainability)

Future Proven, Evergreen, ESG Ready



28%  
Less  
Cooling

105%  
More TB  
Per RU

48%  
Less  
Power

1.3  
Watt per  
TB

미래 지향형 네트워킹

지속적인 SW/HW 현대화

No 다운타임, No 마이그레이션

비용 효율적인 구독형 서비스

지속 가능한 저탄소 기반 기술

TCO 절감







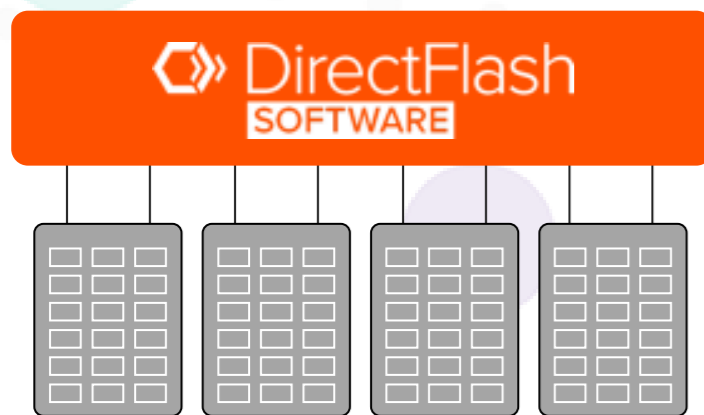
# DirectFlash MODULE

업계 최초

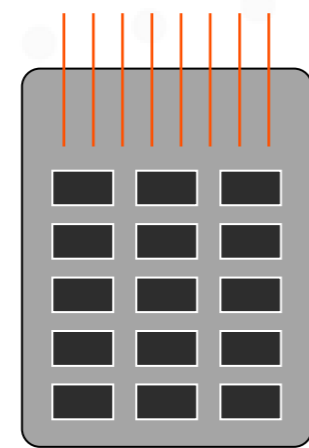
# 소프트웨어 정의 다이렉트 플래시 모듈



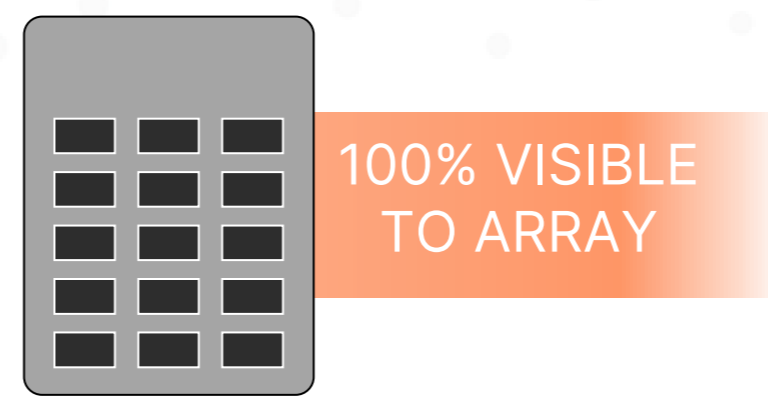
소프트웨어 정의  
스토리지 OS



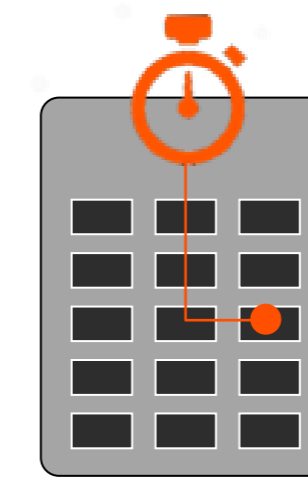
100% *nvm*  
EXPRESS



플래시 공간의 100%  
사용



NAND Level의  
응답속도 보장



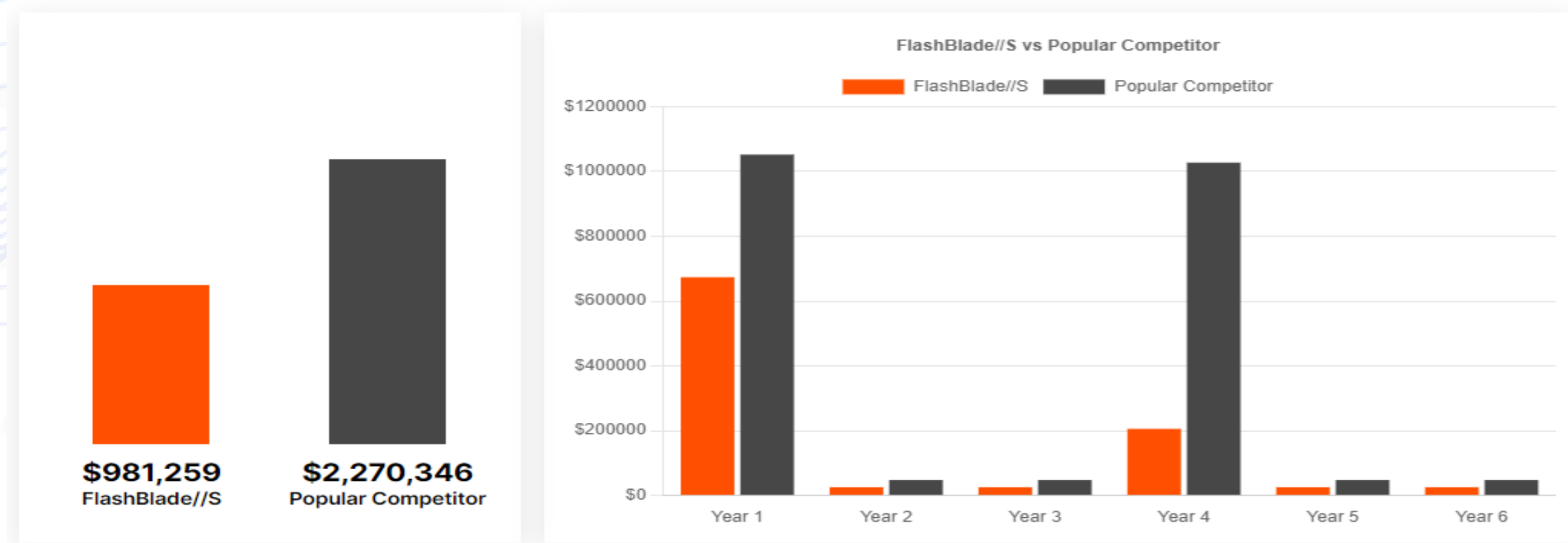
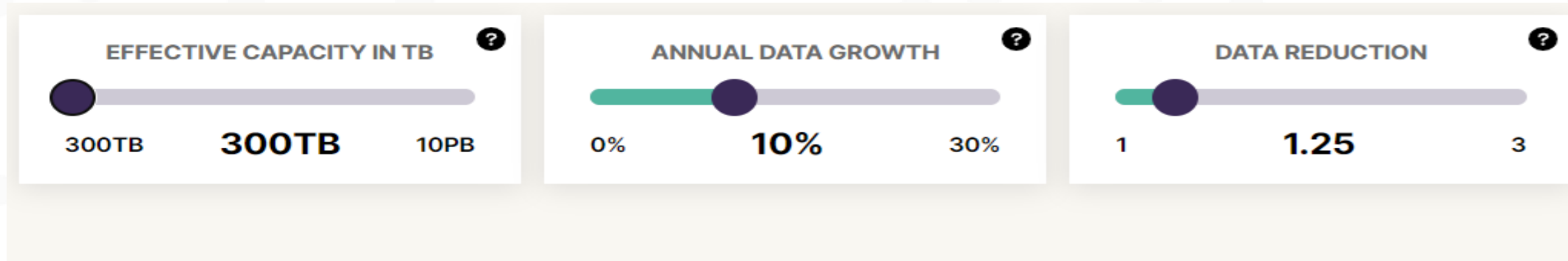
업계 최고의  
용량 집적도





# 6년 TCO 분석

고객이 직접 활용할 수 있는 직관적인 TCO 분석 툴



|                                | FlashBlade//S | Popular Competitor |   |
|--------------------------------|---------------|--------------------|---|
| Hardware/Software/Refresh      | \$468,000     | \$1,257,333        | + |
| Power/Cooling/Space            | \$40,760      | \$53,594           | + |
| Storage Management             | \$112,500     | \$233,280          | + |
| Support                        | \$360,000     | \$726,139          | + |
| <b>Total Six-year Savings:</b> |               | <b>\$1,289,087</b> |   |

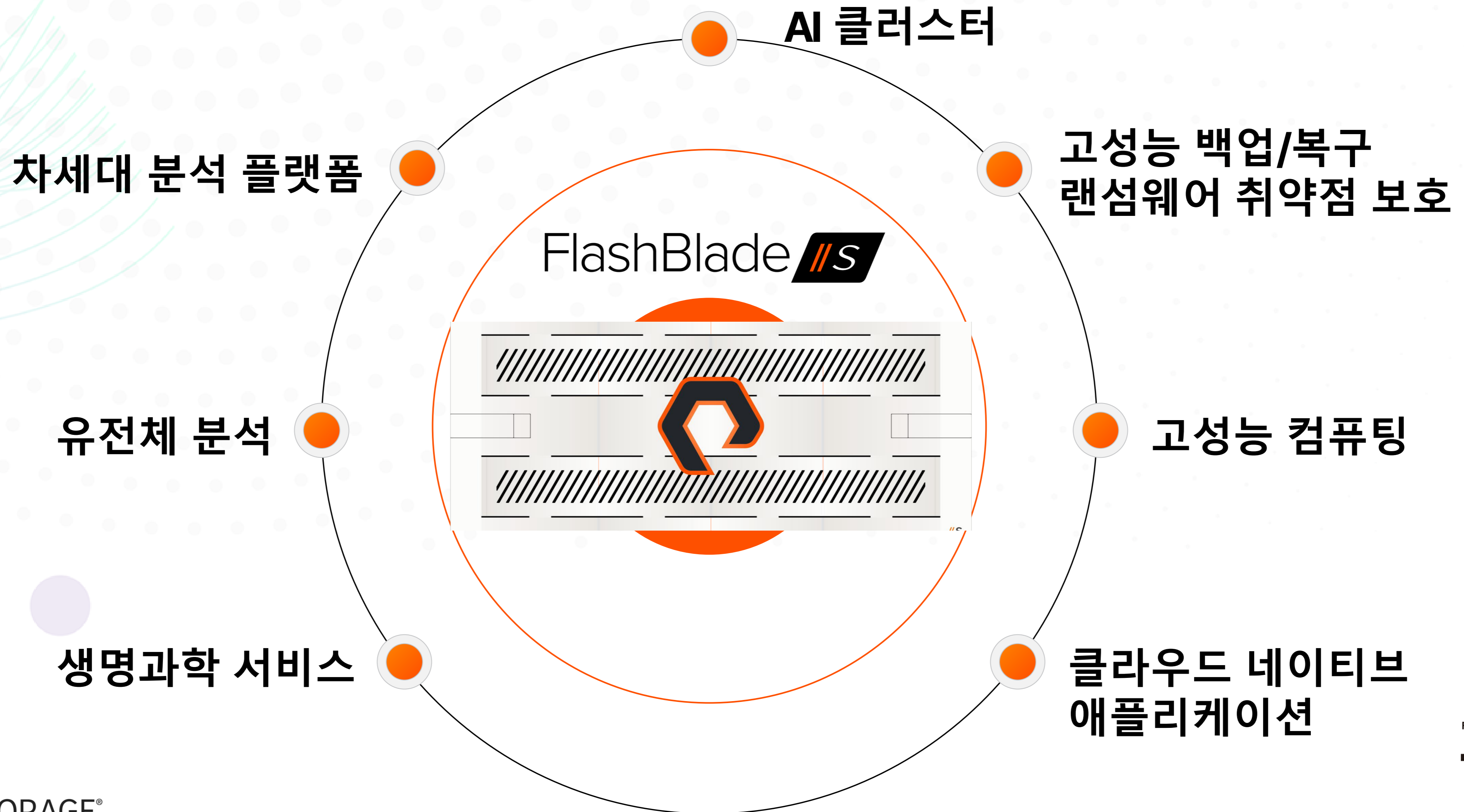
< 50%  
투자/운영 비용 절감

External TCO Calculator - <https://www.purestorage.com/products/file-and-object/flashblade-s/calculator.html>



# 현대화된 솔루션 기반 다양한 고객 경험

강화된 차세대 플랫폼으로 다양한 솔루션 영역에서 PureStorage의 리더십을 확장





# Introducing AIRI//S



## 업계 리더들의 협력:

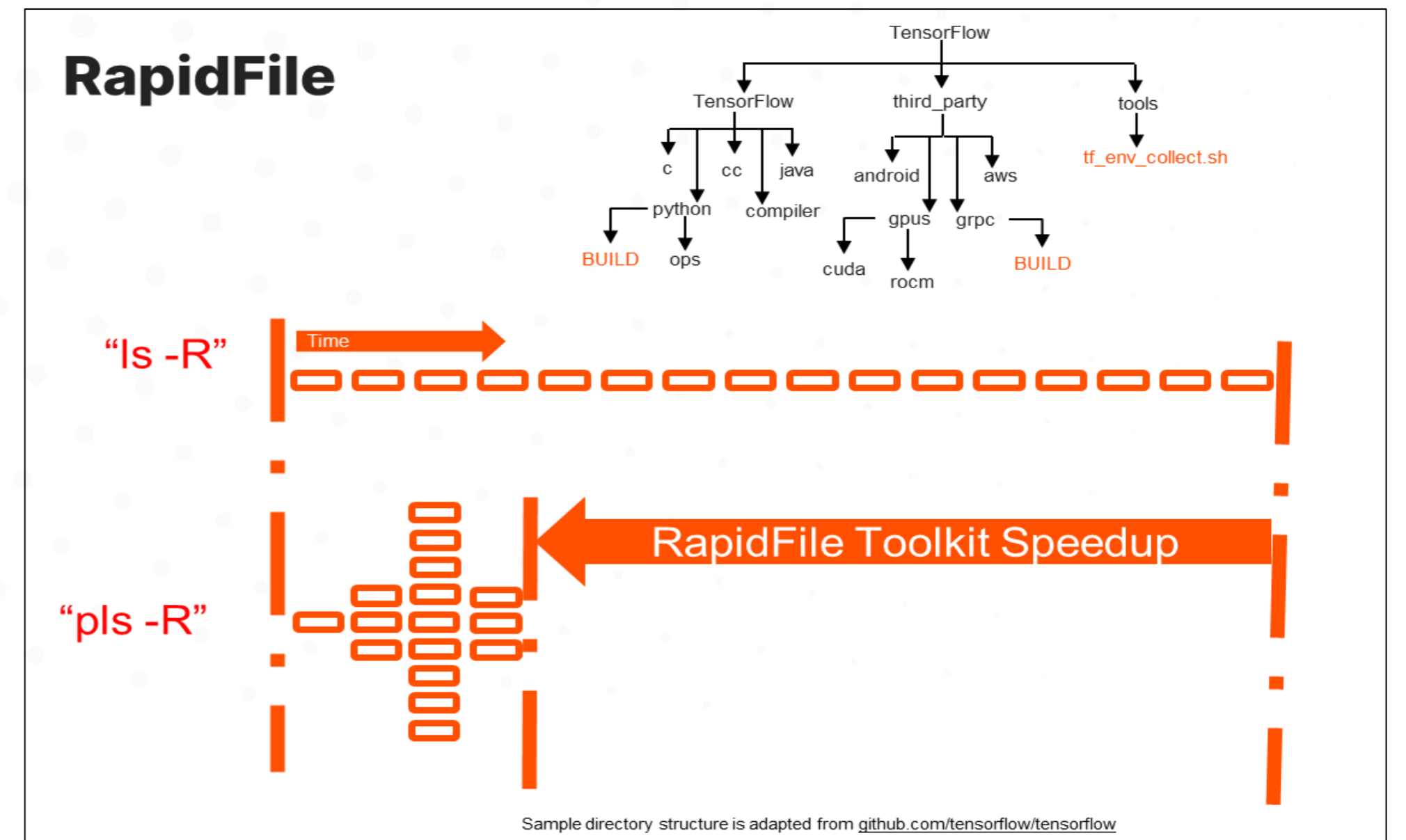
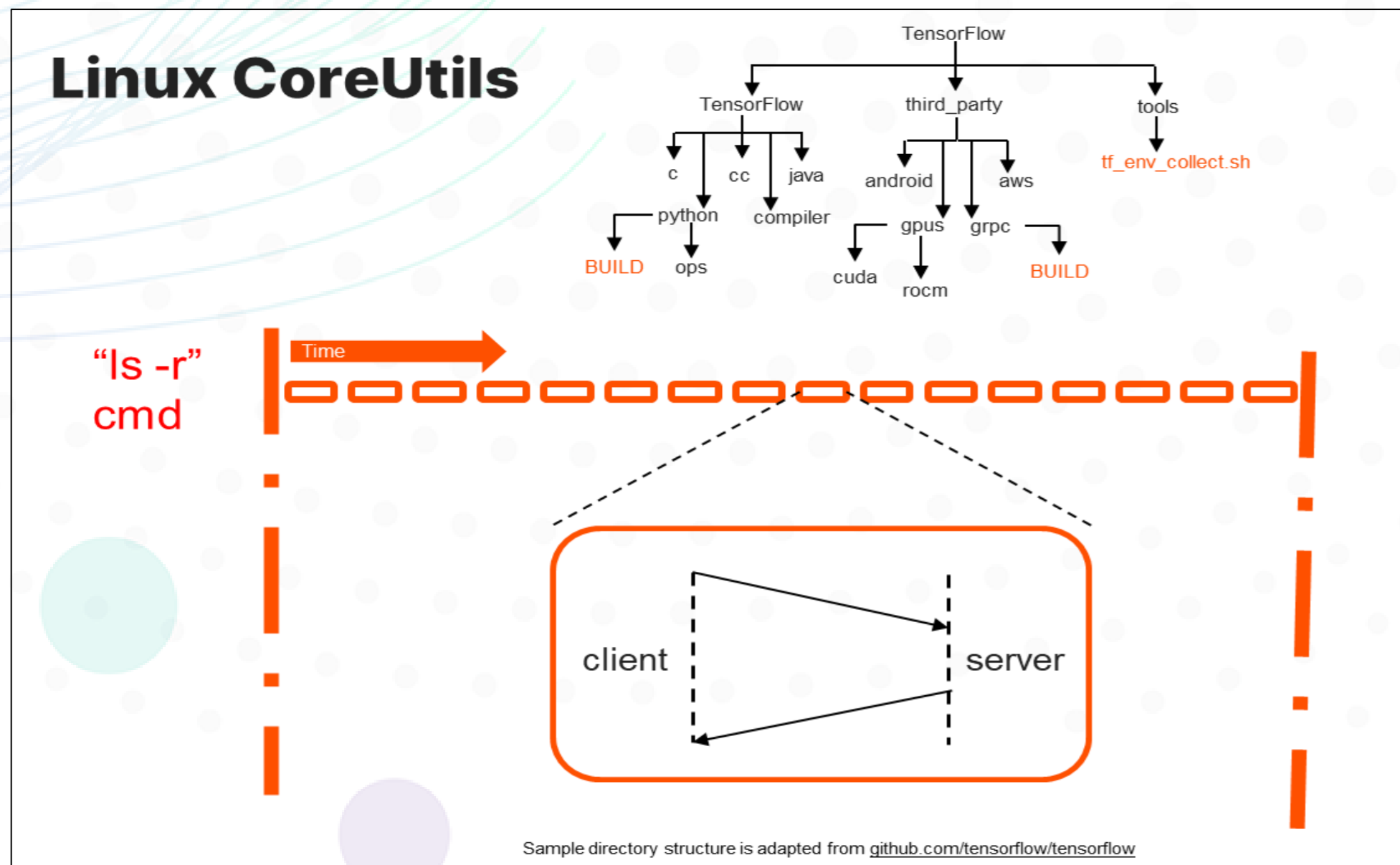
- 퓨어스토리지 FlashBlade//S
- 엔비디아 DGX Server Family
- 엔비디아 Spectrum Networking
- Lablup Backend.AI



# 업계 유일한 데이터 가속 솔루션 (RapidFile 2.0)

## Linux 코어 유틸

## Pure RapidFile



vs

### RapidFile 명령어 리스트

prm  
pchmod  
pchown

pls  
pfind  
pdu

pcp





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

Meta는 글로벌 최대 규모의 가장 빠른 AI 슈퍼컴퓨터 플랫폼을 구축했고, PureStorage는 RSC를 위한 강력하고 확장 가능한 스토리지 솔루션을 제공합니다.

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

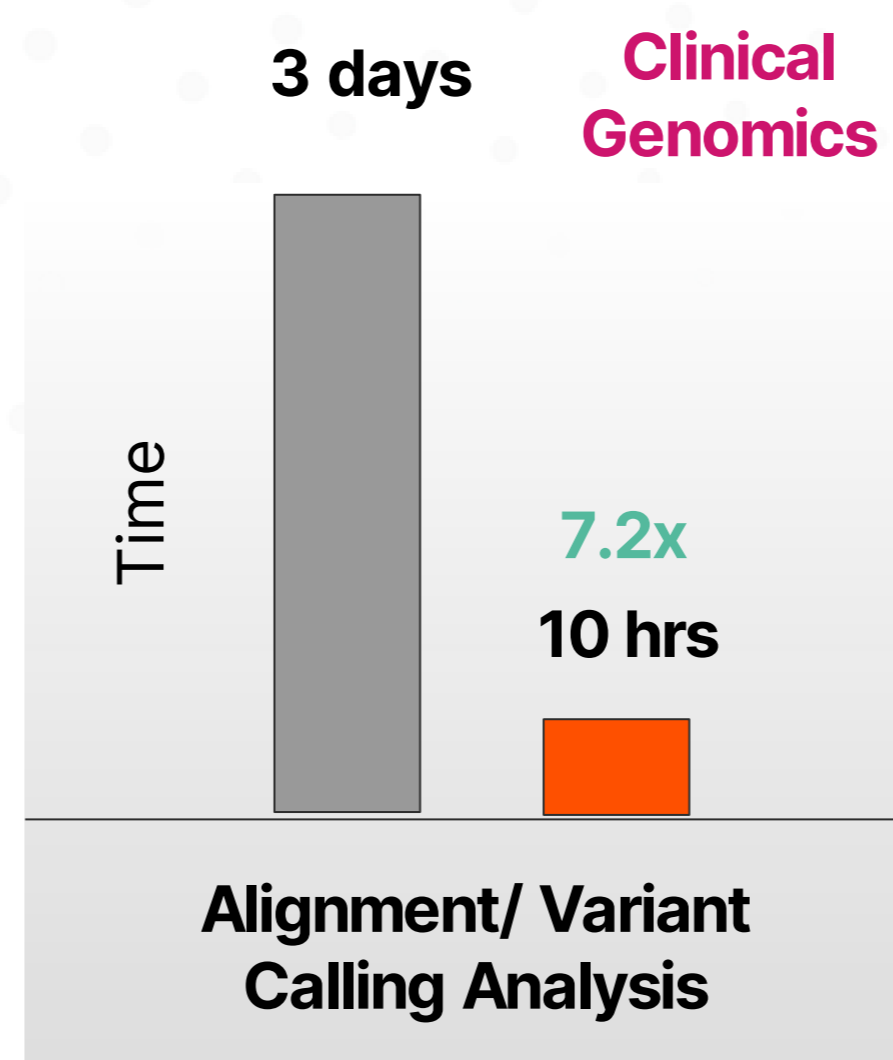
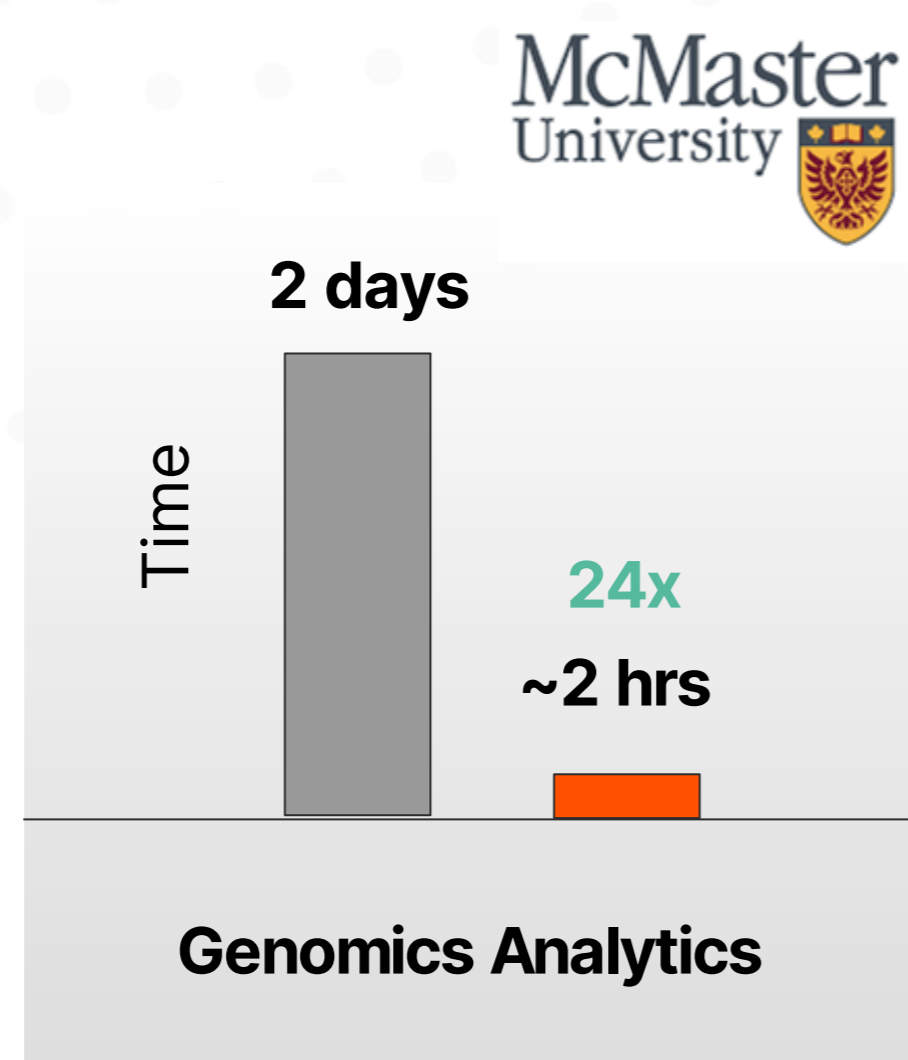
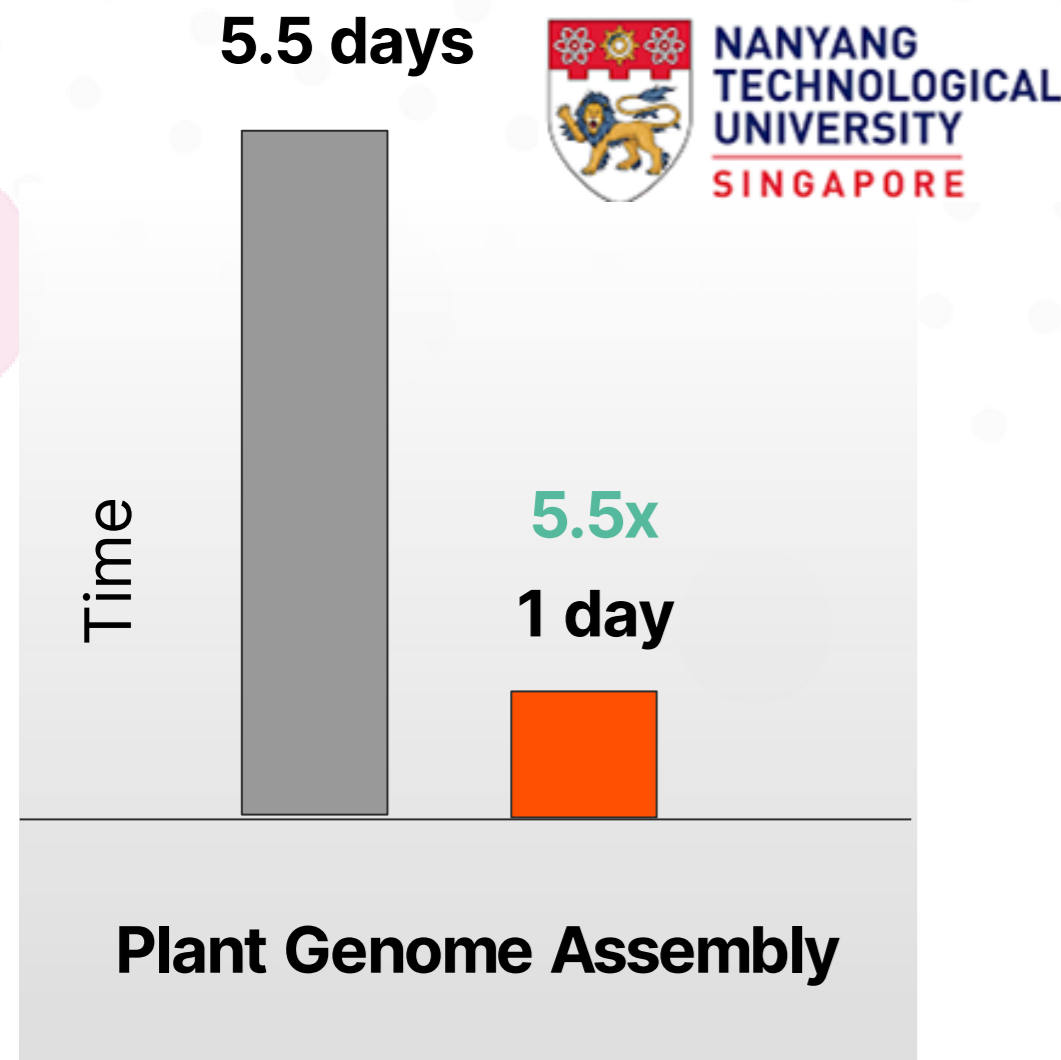
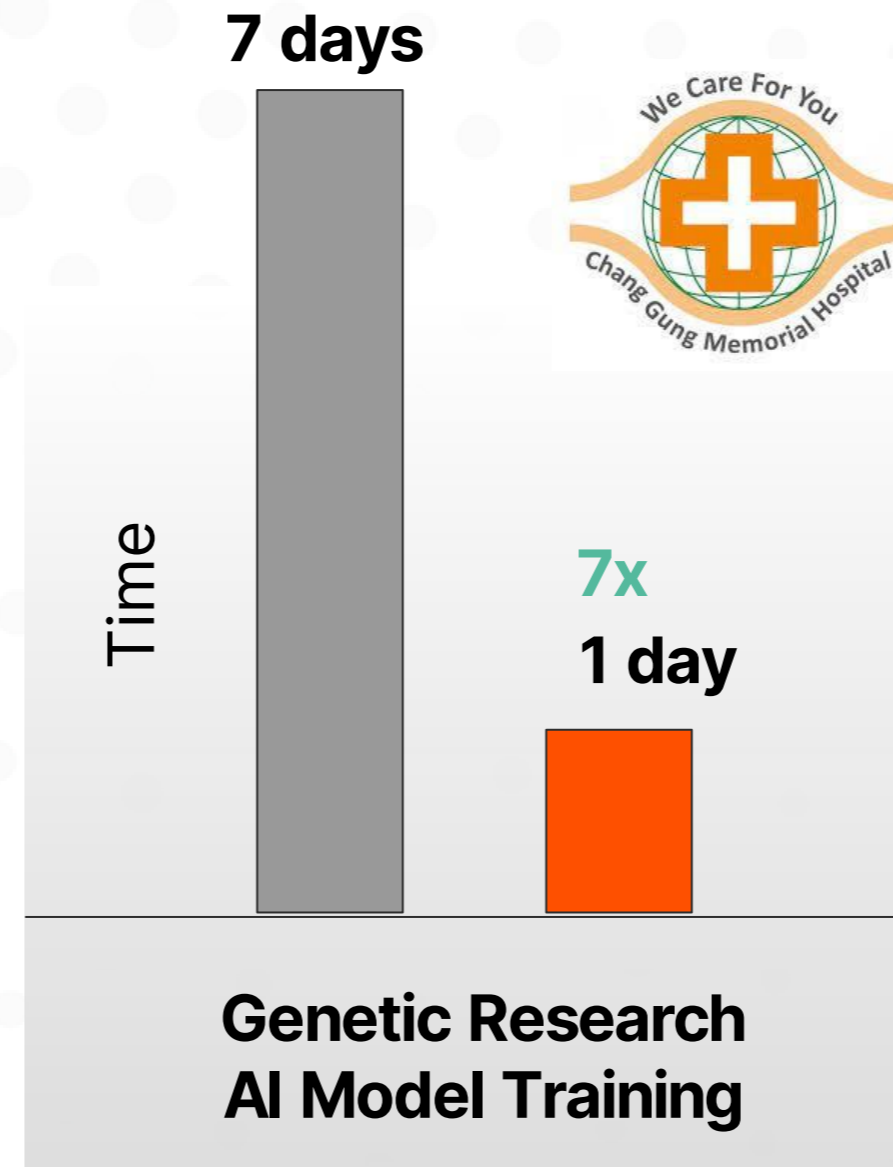
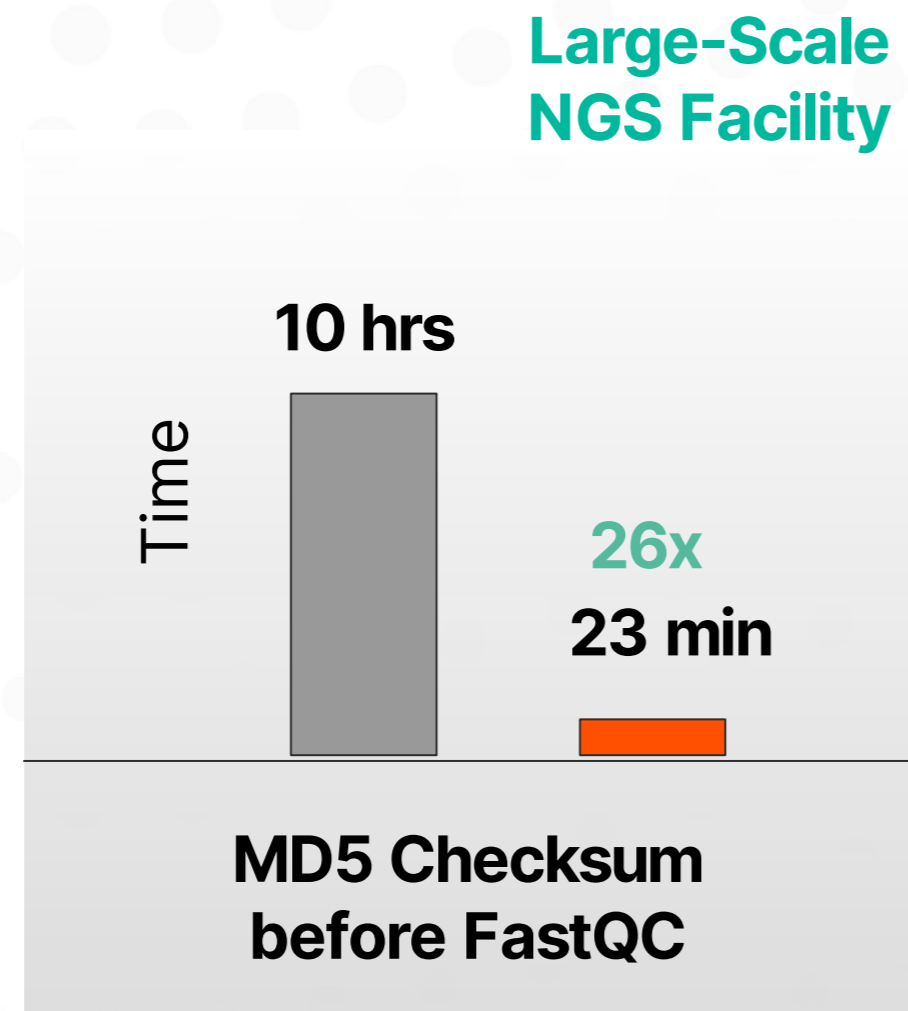
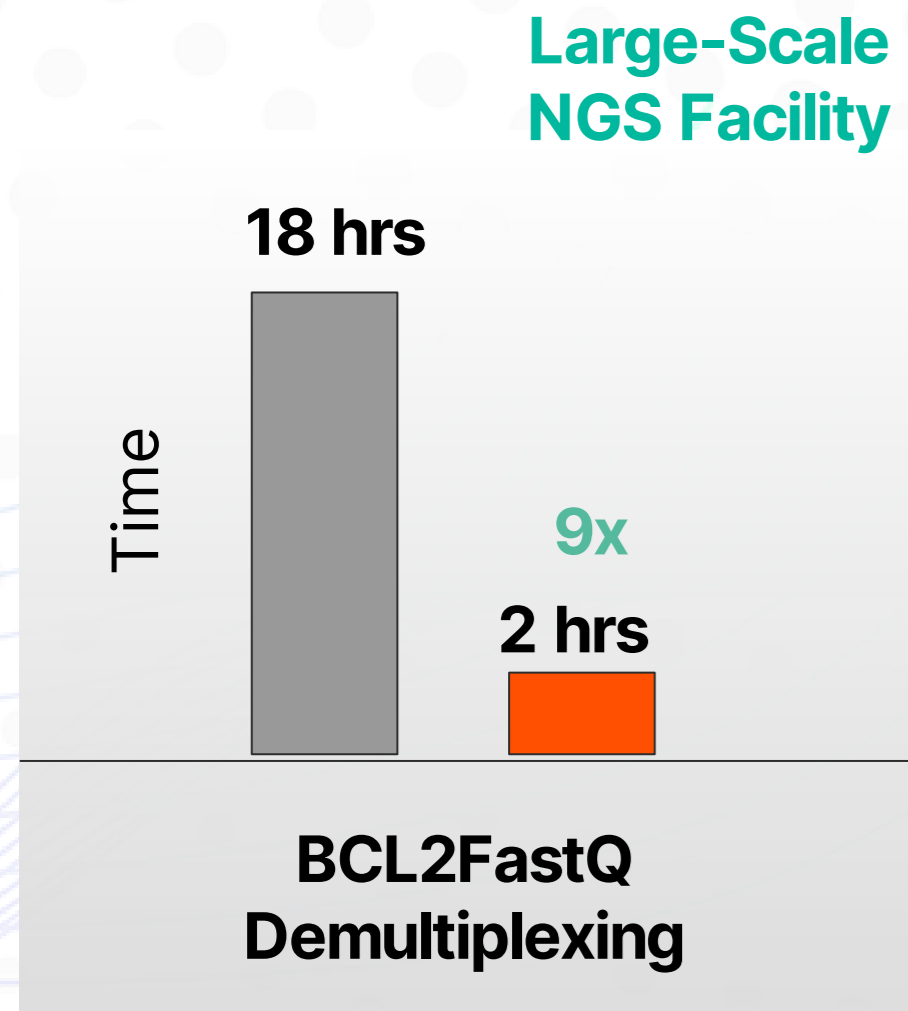
## Meta 주요 고려 사항

- 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**



# 데이터 분석 개선 효과



## 분석 및 진단 시간 단축

"전통적인 스토리지는 급증하는 데이터 처리에 충분한 성능을 제공할 수 없기 때문에 아무 의미가 없습니다. 퓨어 스토리지 FlashBlade를 사용하면 인간의 건강에 대한 전 세계적인 위협에 맞서 싸울 때 앞서 나갈 수 있습니다."  
 Andrew G. McArthur, Ph.D.  
 McMaster University

Competitor   
 FlashBlade 



# 기업 비즈니스 혁신

1

**FlashBlade//S**  
**A Game Changer**

2X 상면 효율  
2X 성능  
2X 전력 효율

2

**Evergreen**  
**구독 서비스 모델**  
Evergreen//Forever  
Evergreen//One  
Evergreen//Flex

3

**ESG Ready**

75% 에너지 소비 감소율  
97% 운영 지속성  
1.3 Watt/TB

# Uncomplicate Data Storage, Forever



# How to Modernize AI



**PURE**STORAGE

+



**Lablup Inc.**

PURE//ACCELERATE®

tech  
fest<sub>2</sub>  
seoul<sub>2</sub>





# FlashBlade를 활용한 Backend.AI Cloud SaaS Architecture

조만석 Solution Architect (AI/ML Specialist)

Lablup Inc.





# 래블업 주식회사



**lablup**

## 미션

클라우드 컴퓨팅과 AI 기술을 통해  
계산 기반 연구의 새로운 패러다임을 제시함으로써,  
현대의 과학 연구와 그 응용 분야 간에 존재하는  
속도 차이를 해결한다.

## 개요

회사 설립일: 2015년 4월 20일

직원 수: 21명, R&D 인력 비중 65%



**BACKEND.AI**

Backend.AI (<https://www.backend.ai>)

사용자들이 클라우드 혹은 온프레미스 환경에서  
쉽게 ML 모델을 훈련하고 서비스 하는 과정을  
통합하고 가속화 해 주는 정교한 GPU PaaS

## 독자 기술

- 자체 컨테이너 오케스트레이터
- 리눅스 커널 시스템콜 호출 가상화 기술
- 드라이버 레벨의 GPU 분할 가상화 기술
- 하이브리드 클라우드 플랫폼 기술
- 연산 노드간 직결 보안 계층 기술
- 컨테이너 내 실시간 애플리케이션 레이어 조합 기술 RE//ACCELERATE®



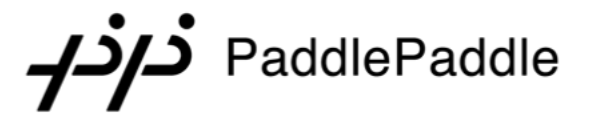
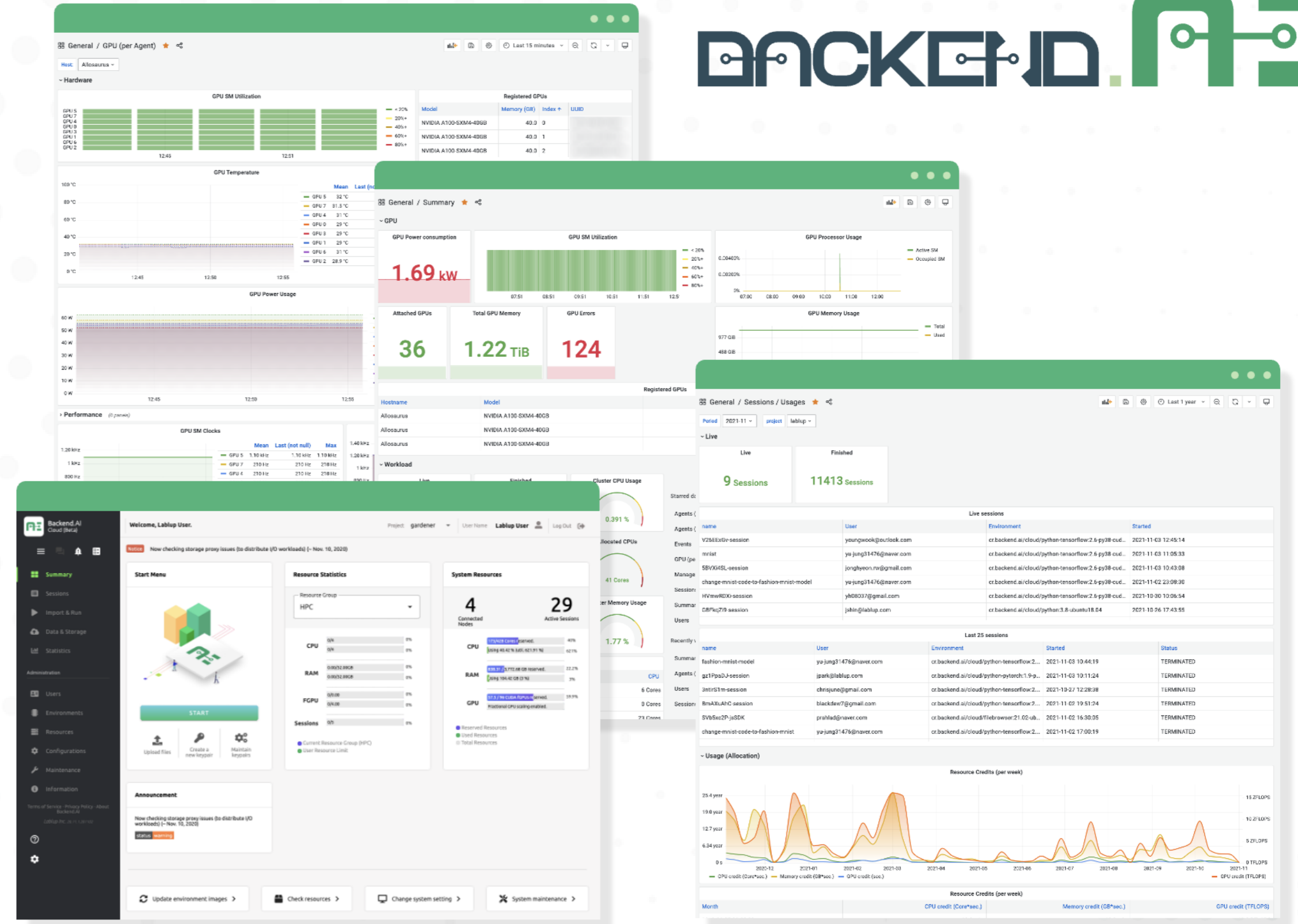
# Backend.AI

AI 개발 및 서비스를 위한  
올인원 엔터프라이즈 운영 플랫폼

머신 러닝 모델 개발의 본질에만 집중하도록  
AI 훈련 및 서비스 플랫폼에 필요한 관련 기술을  
하나의 프레임워크로 제공하는  
온프레미스 / 클라우드 통합 솔루션 및 플랫폼

## 사용자 강점

독자 구현 GPU 분할 가상화를 통한 자원 확장 및 공유  
쉬운 GUI 와 강력한 CLI / SDK  
관리 용이성 및 높은 활용도



PURE//ACCELERATE®

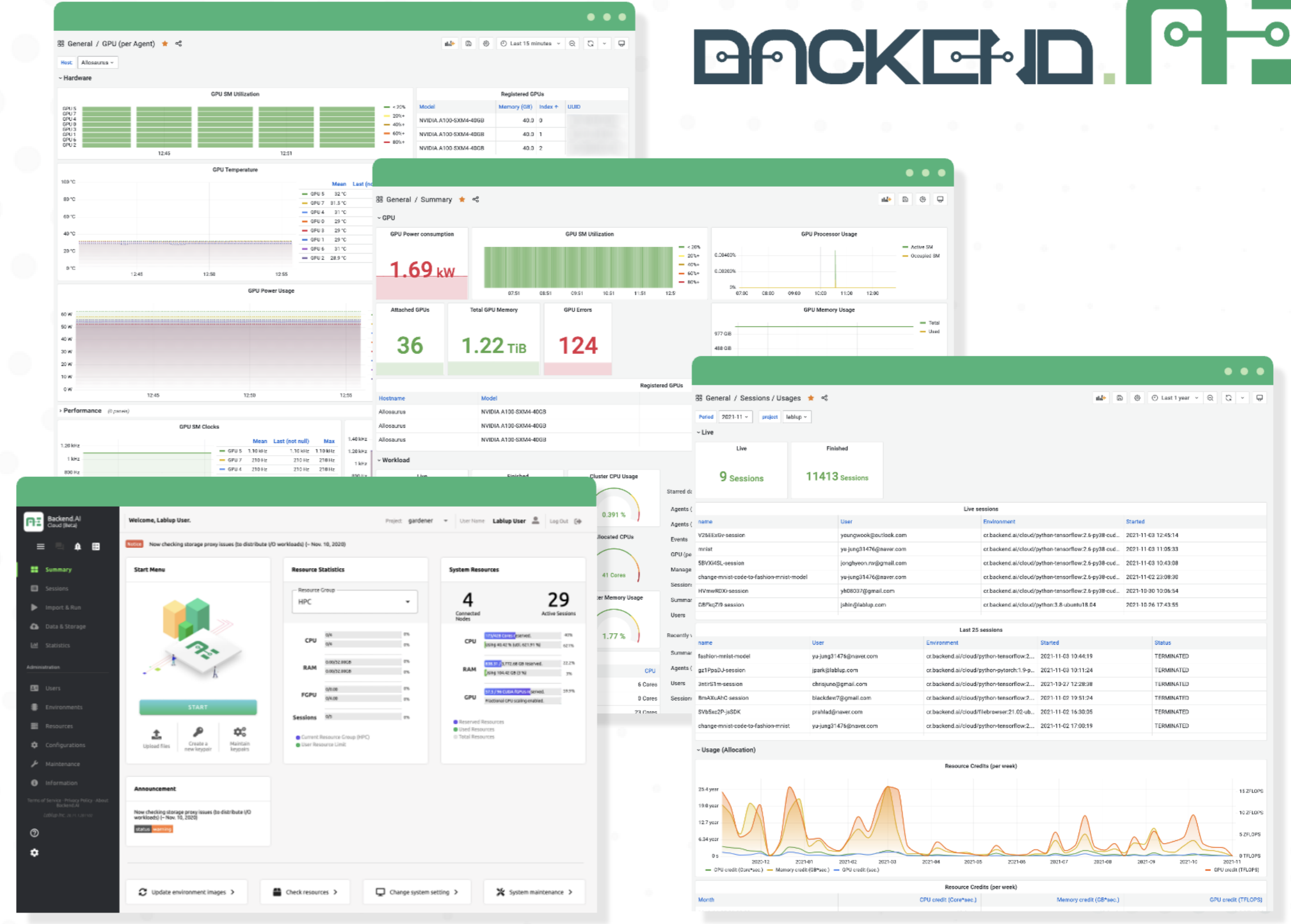




# Backend.AI

AI 개발 및 서비스를 위한  
 올인원 엔터프라이즈 운영 플랫폼

자체 오픈소스 코어 / 하이퍼스케일러  
 GPU 자원 효율화+AI 클러스터 관리 및 최적화  
 +개발자 및 사용자 서비스+MLOps





# 당면한 문제

## 프라이빗 AI 시스템 고객

- **고성능 고효율의 AI 모델 개발**
  - AI 모델 훈련의 병목: 멀티 노드 훈련
    - 데이터 피딩, GPU간 통신
  - 데이터를 분산 노드로 균일하고 빠르게 스트리밍할 필요성
- **딥 러닝과 스토리지의 어려운 통합**
  - 멀티 노드를 아우르는 데이터 레이크
  - 같은 데이터를 여러 노드에 걸쳐 전달
  - 대역폭, 안정성 및 확장성
  - 초고속 스토리지 솔루션들의 높은 사용 난도

## 래블업 클라우드

- **대규모 퍼블릭 클라우드 구축**
  - 다수의 사용자들을 담을 수 있는 스토리지
  - 딥 러닝 훈련 데이터를 빠르게 전달
  - 쉽게 스케일 조정이 가능한 스토리지
- **적정 비용으로 사용자에게 자원 공급**
  - 보관하는 데이터의 크기가 아주 큼
  - 대부분의 데이터는 즉각 사용되지 않음



# 래블업의 해결

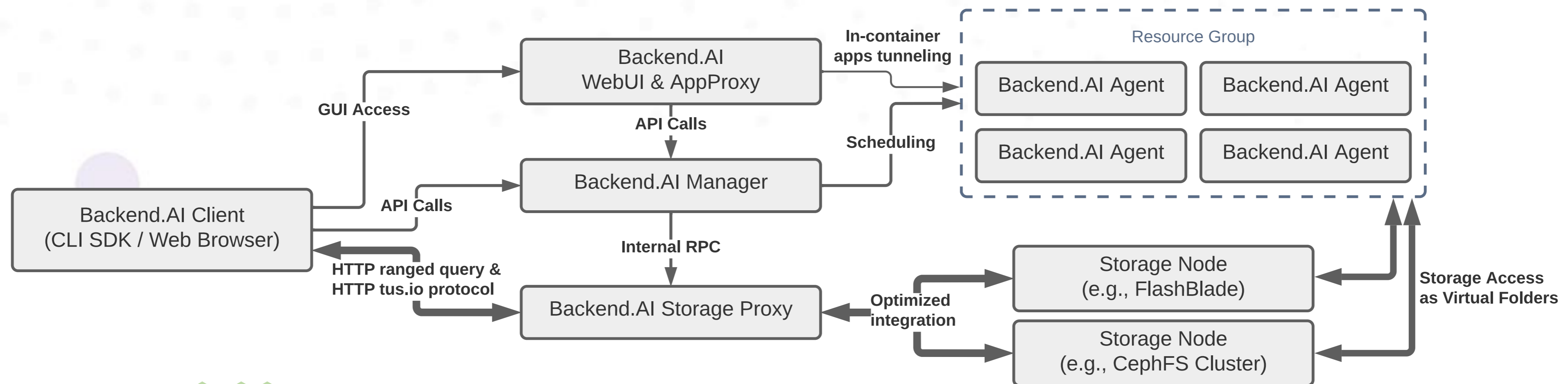
## Lablup + PureStorage -> Supercharged Backend.AI Enterprise

- **스토리지와 함께 최적화된 AI 개발 및 서비스 솔루션**
  - 네트워크 파일 시스템 및 오브젝트 파일 시스템 **자동 할당**
  - 멀티 노드를 아우르는 데이터 레이크에서의 편의성 극대화
  - **쉬운 네트워크 데이터 공유**: 간단한 노드간 데이터 공유 과정
  - 대역폭, 안정성 및 확장성
  - **스토리지 네트워크**를 유저 네트워크 및 **딥 러닝 워크로드와 분리** 설정 및 **사용 자동화**
  - 스토리지 전용 프록시 기반의 FlashBlade + Backend.AI 통합
    - 초고속 스토리지 솔루션의 **가속 기능을 학습 없이 바로 사용**



# Backend.AI: 스토리지 입출력 분산 최적화 및 통합

- 대용량 파일 업로드·다운로드를 API 서버를 거치지 않고 별도 proxy로 분리
- tus.io 업로드 프로토콜 및 HTTP ranged request 다운로드 프로토콜 적용 (끊겨도 이어서 전송)
- 파일시스템·NAS 장비 별 최적화된 기능 제공
- FlashBlade 통합
  - FlashBlade의 스토리지 가속 기능에 대한 지식 없이도 자동적으로 통합 및 가속 기능 적용
  - RapidFile Tools : 고속 스캔 및 사용 통계
  - Purity REST API : 파일 시스템 당 성능 지표 측정 (IOPS, I/O latency, I/O throughput)

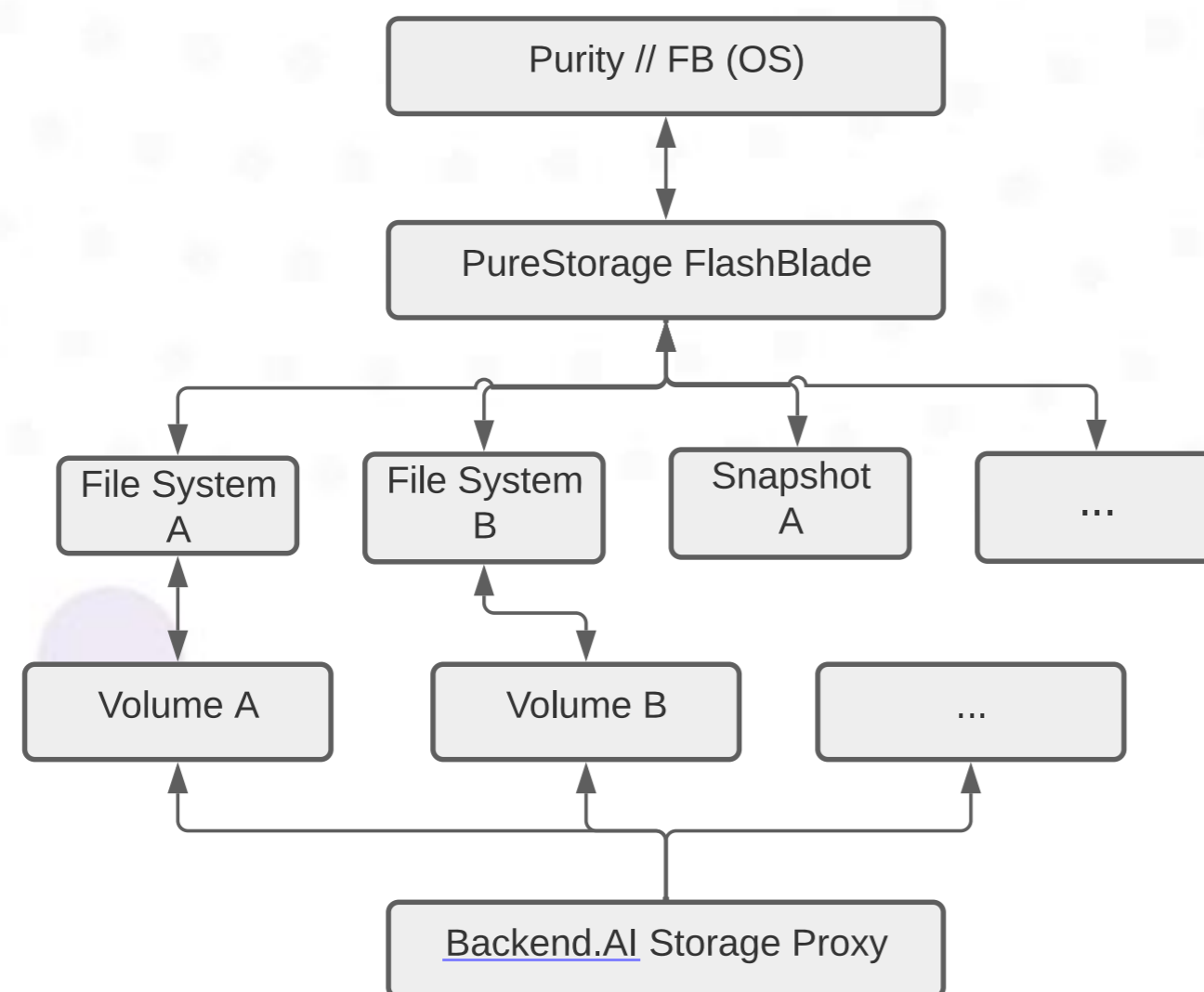




# Backend.AI: 스토리지 입출력 분산 최적화 및 통합

- FlashBlade 통합

- FlashBlade의 스토리지 가속 기능에 대한 지식 없이도 자동적으로 통합 및 가속 기능 적용
- RapidFile Tools : 고속 스캔 및 사용 통계
- Purity REST API : 파일 시스템 당 성능 지표 측정 (IOPS, I/O latency, I/O throughput)



The screenshot displays the Backend.AI web interface for a Flashblade node. The main content area shows a table of Flashblade Nodes with the following data:

| # | Endpoint                                   | Backend Type        | Resources | Capabilities                   | Control        |
|---|--|---------------------|-----------|--------------------------------|----------------|
| 1 | local:mypure<br>/mnt/pure_nfs3/demo-volume | Backend purestorage |           | vfolder<br>fast-scan<br>metric | [Control Icon] |
| 2 | local:mypure2<br>/mnt/pure1                | Backend purestorage |           | vfolder<br>fast-scan<br>metric | [Control Icon] |

Below the table, there are several panels:

- Array:** Shows details for 'New-Pure' including Model/Name, Serial Number (78deaf7-6c35-40ad-8d0c-12453532cbfa), and 1 Blade.
- Software:** Shows Purity//FB OS version 2.4.7 and revision 2019.12.06\_c1a19b1d.
- Configuration:** Shows local:mypure Storage Proxy ID and /mnt/pure\_nfs3/demo-volume Mount point.
- Statistics:** Shows Usage of 0.03 TB.
- Performance (Current):** Shows IOPS and Traffic metrics for Read and Write operations.
- Operations:** Includes an 'OPEN CONSOLE' button.



# 사례: AI 슈퍼컴퓨팅 클러스터 구축

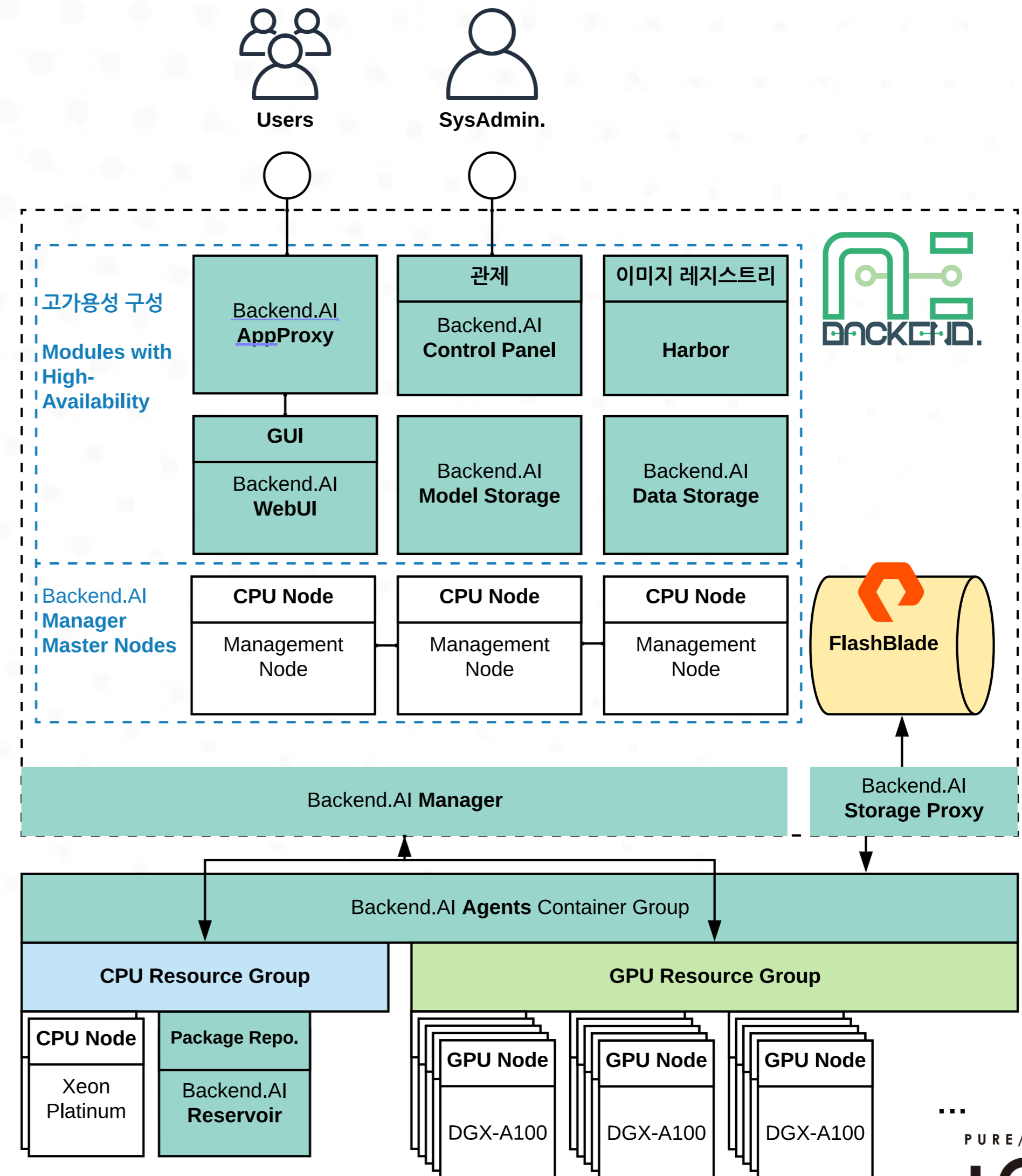
멀티 리전, 다중 조직 사용자를 위한 AI 개발 클러스터

## 구성

- (GPU 워크로드용) A100 GPU 200대  
(빅데이터 분석 및 데이터 전처리용) 고성능 CPU 노드 다수
- 주기적 업데이트 및 사용자별 패키지 설치를 포함한 완전 Air-gapped 환경 구축

## 고객 혜택

- **대단위 팜 구성 설계** 제공 및 SLA 극대화 고가용성 구성
- 멀티노드 분산 훈련 및 GPU간 직접 네트워크 기반의 **대규모 초고속 딥러닝 훈련**
- 로컬 패키지 저장소 솔루션인 Backend.AI Reservoir와 결합하여 **PyPI 및 Ubuntu 저장소를 완전 폐쇄망 내에서 자유롭게 사용** 지원
- 기관 내외부 동시 서비스 시 시스템/데이터 보안을 위한 격리 도메인 구성





# 사례: 래블업 Personal Backend.AI Cloud SaaS

## 달성 목표

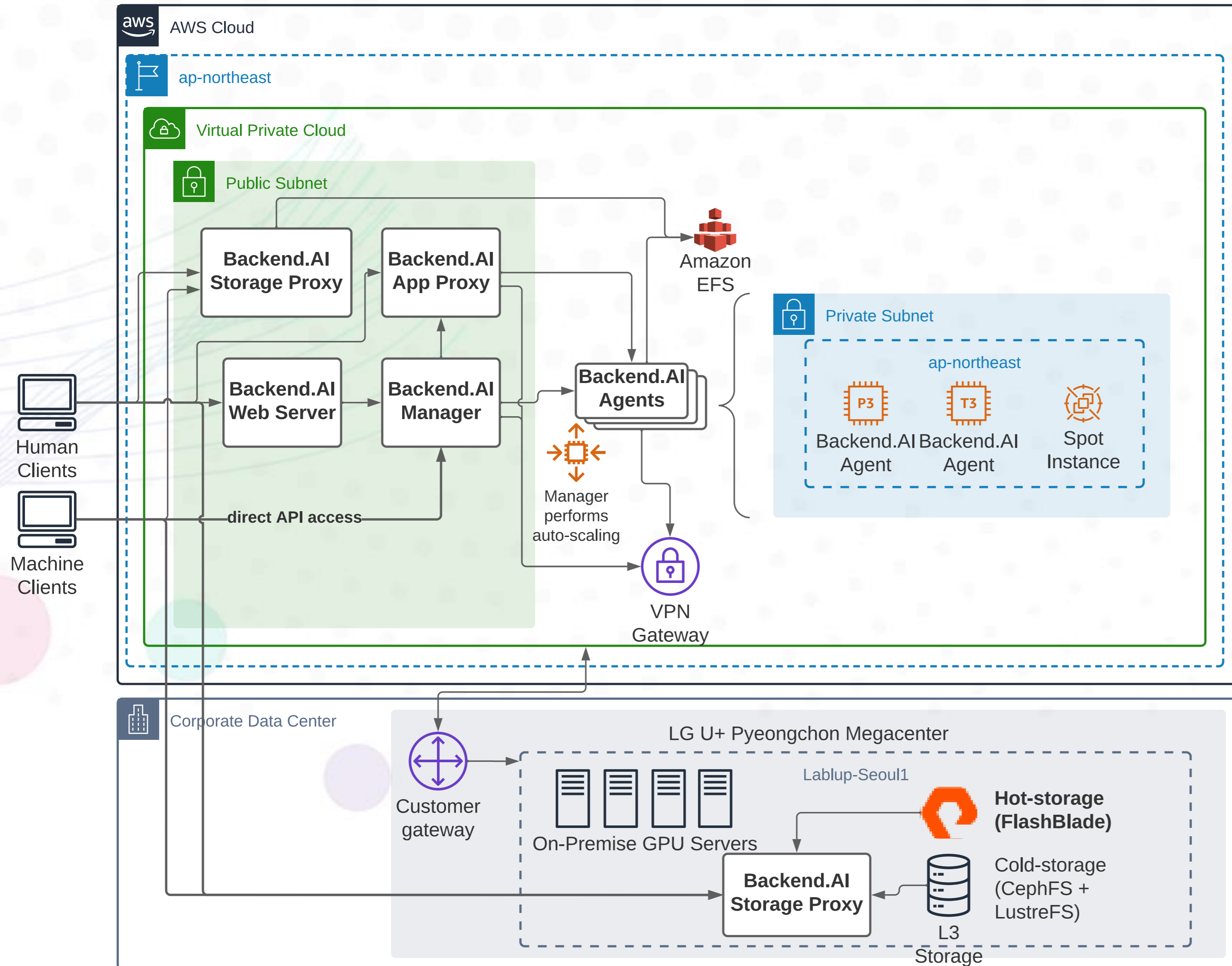
- Backend.AI 멀티 스토리지 프록시 기반의 유연한 사용자 데이터 액세스 환경 제공
- 스토리지 종류에 따른 별도의 메트릭 및 과금 모형

## 사용자 시나리오

- 자주 사용되지 않는 데이터는 HDD 기반의 분산 콜드 스토리지에 보관
- 스크래치 및 액세스가 빈번한 데이터는 PureStorage FlashBlade 상에 보관
- 사용자가 원할 때 폴더를 Hot storage – Cold storage간 전환하는 명령 및 옵션 제공



# 사례: 래블업 Personal Backend.AI Cloud SaaS



## L3 (Lablup LargeFile Lake) 구성

- Hot storage: **FlashBlade** 기반
  - 사용자 기본 파일 및 작업공간
  - 멀티노드 훈련시 데이터 입출력
  - 유료 사용자들의 기본 사용 공간
- Cold storage: CephFS+LustreFS
  - 장기간 보관 및 사용률이 낮은 파일
  - 무료 사용자들의 기본 사용 공간

비용 비교

## 달성

- 적정 비용으로 **초고속 입출력 속도**와 대용량 스토리지 제공 목표를 모두 달성
- 사용자들의 **AI 분산 훈련 경험**: 빠른 스토리지 사용의 필요성 절감



# 요약

## PureStorage



- FlashBlade 초고속 스토리지 토탈 솔루션
- 초고속 스토리지의 사용자 경험을 Backend.AI 를 통해 쉽게 전달

## Lablup Backend.AI



- AIOps 올인원 솔루션
- 엔터프라이즈 및 클라우드 온프레미스부터 클라우드까지 폭넓게 적용할 수 있는 고성능 스토리지 가속 기능 통합
- GPU 특화 기능들과 결합한

## 엔터프라이즈 고객



- 딥 러닝 훈련의 병목 최소화를 위해 고속 입출력 스토리지 솔루션이 필요
- 학습 곡선 없는 초고속 스토리지 입출력 사용
- 딥 러닝의 병목 구간 최소화

## Backend.AI Cloud



- 비용과 용량의 타협점을 찾아 서비스 제공
- AI 훈련 데이터 입출력을 FlashBlade를 이용하여 가속
- 작업후 데이터 및 장기간 미사용 데이터는 HDD 기반 환경에 보관
- 비용과 속도의 균형에 엄청난 도움



# Thank you

