

THE NEXT GENERATION IN — SUPERCOMPUTING

슈퍼컴퓨팅, 클라우드, 데이터센터를 잇는
HPE CRAY EX SUPERCOMPUTER

📅 2021. 09. 29(수), 10:30 ~ 11:40

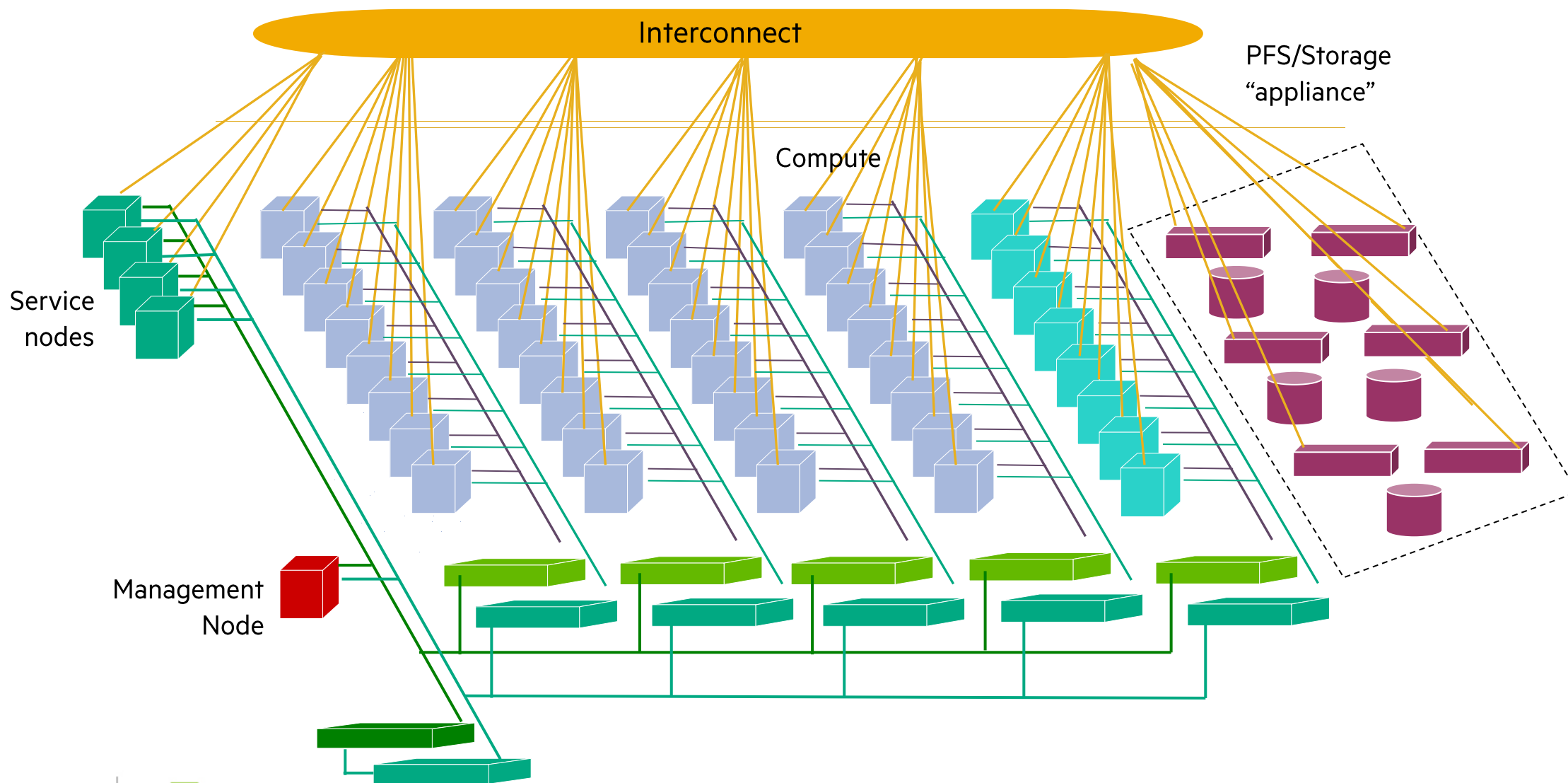
📍 ALLSHOW TV



HPE 고성능 스토리지 어플라이언스 Cray E1000 & HPFSS

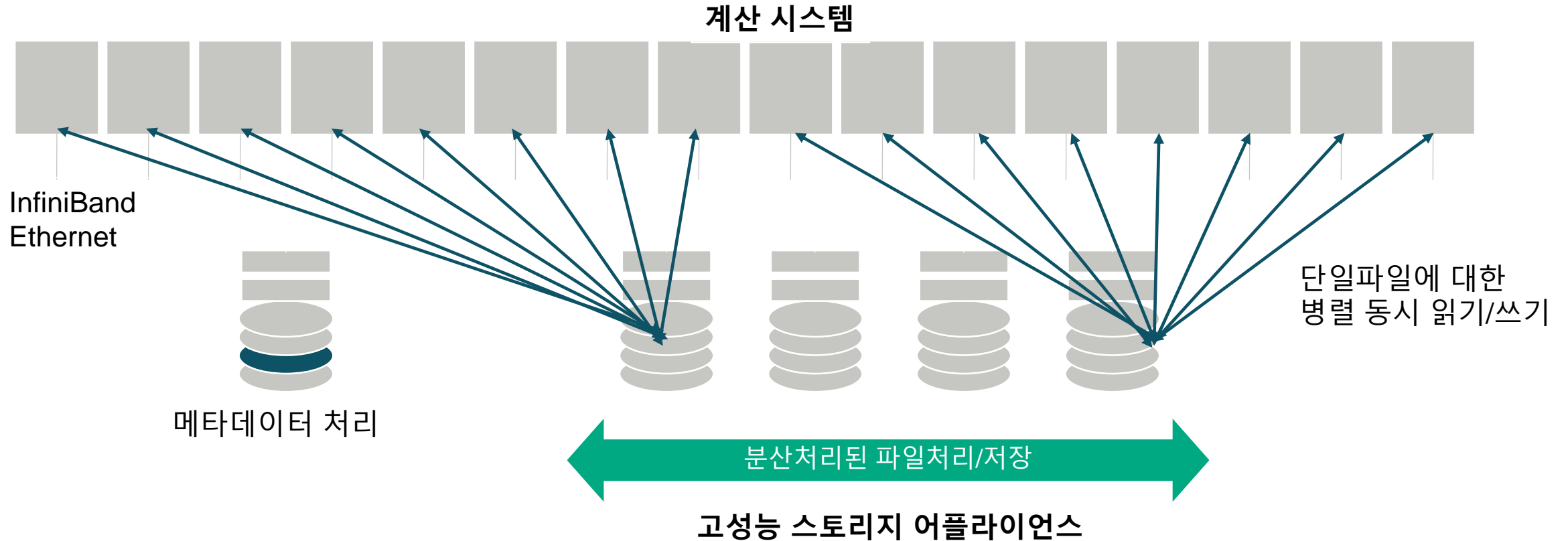
DongJin Suh | HPE

슈퍼컴퓨팅 인프라 표준 아키텍처



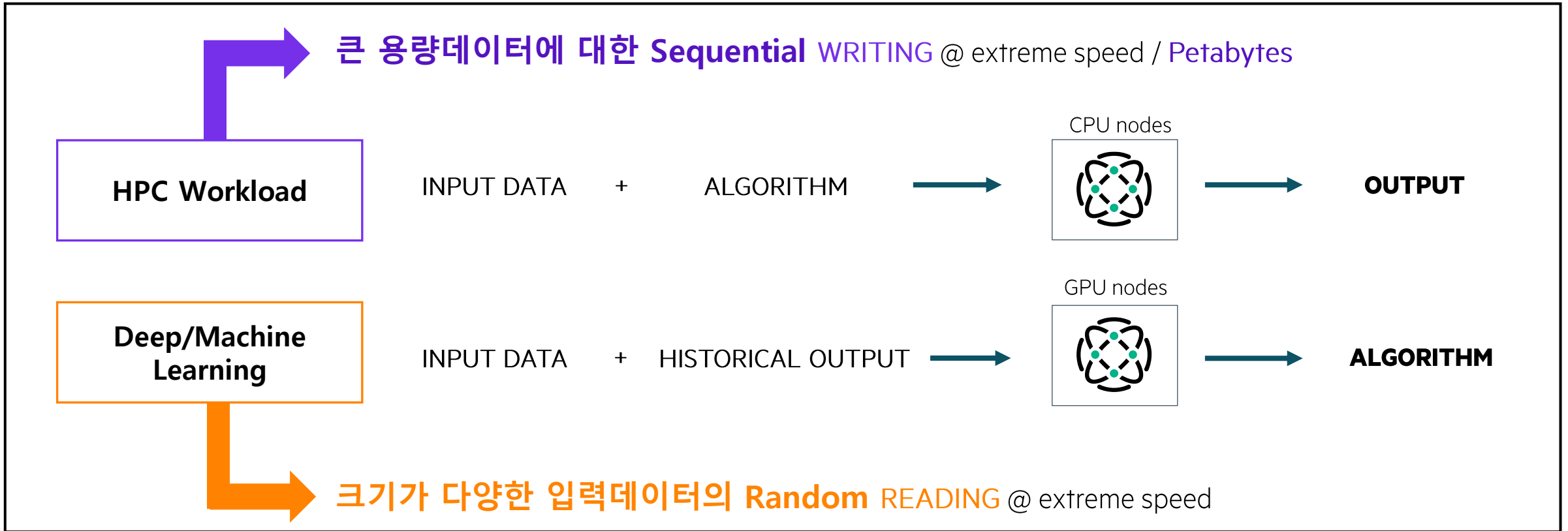
차세대 HPC/AI 스토리지 요건#1

- GB/s ~ TB/s 수준의 고성능 대역폭을 목적에 최적화하여 자유롭게 구현



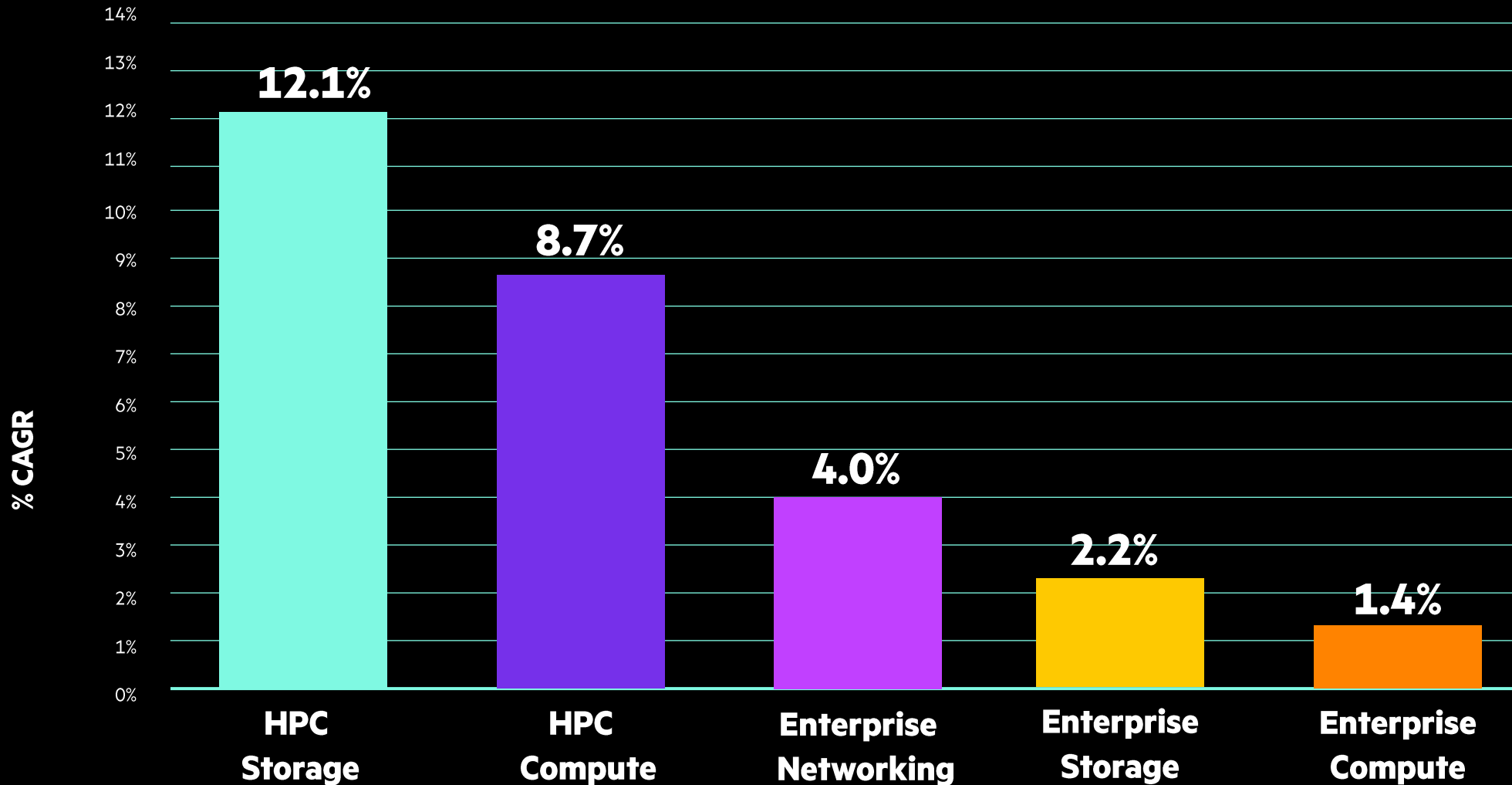
차세대 HPC/AI 스토리지 요건#2

- HPC 데이터와 새로운 AI 워크로드를 동시에 지원토록 설계된 전용 고성능 저장장치



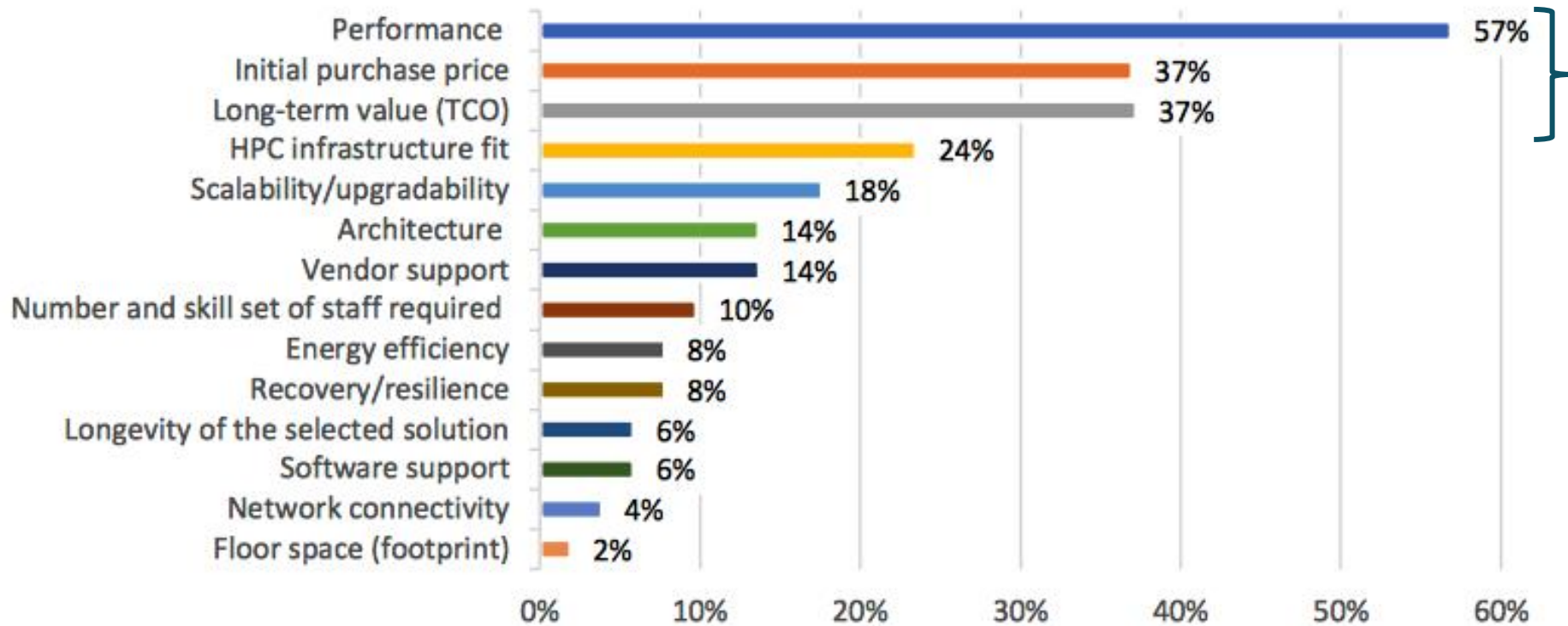
NEW FORECASTS AND HPC TRENDS

Forecasted Compound Annual Growth Rate (CAGR) for customer spending 2019 – 2022



HPC/AI 스토리지 도입시 주요 고려포인트

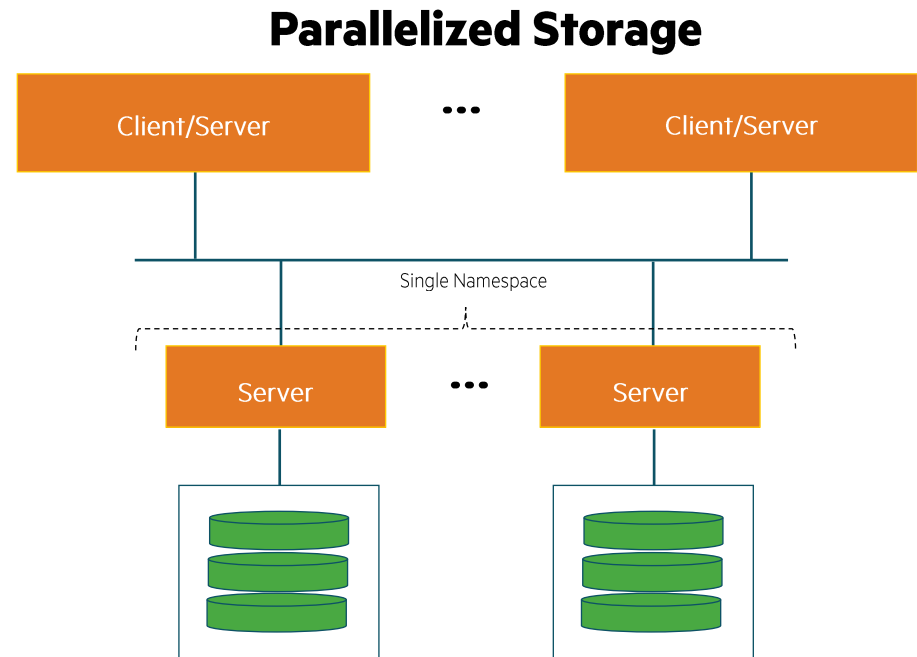
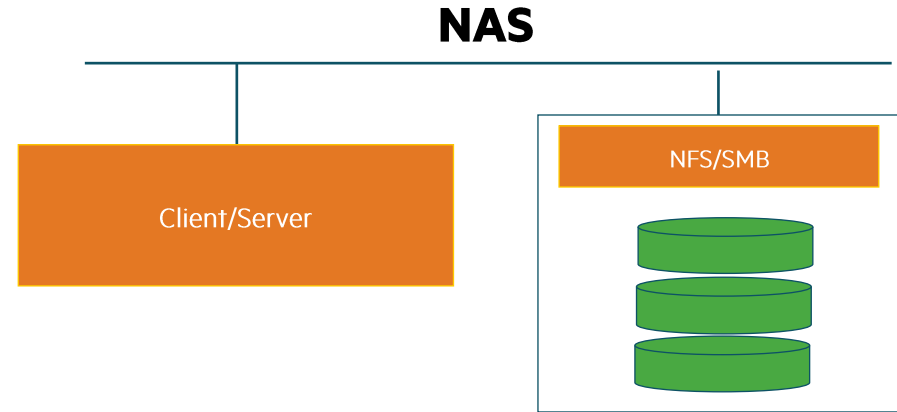
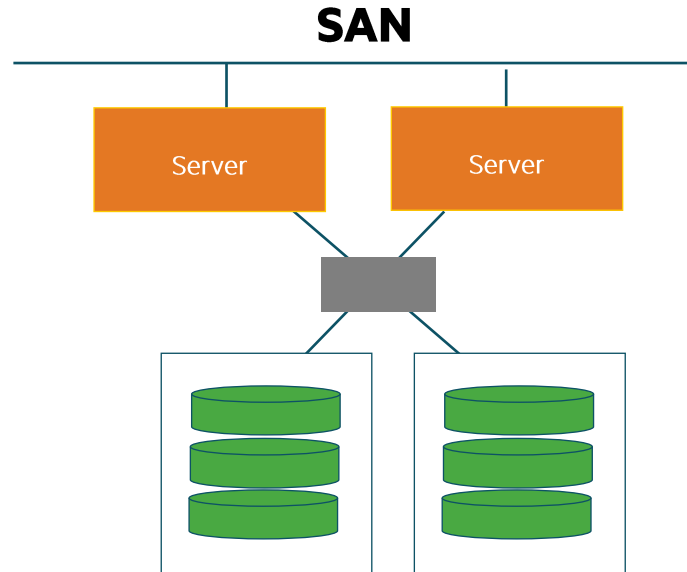
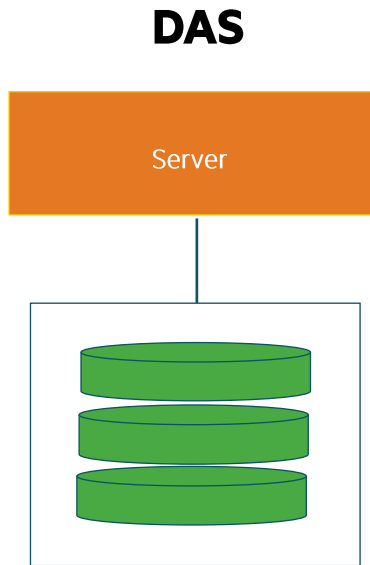
Which of these is your MOST IMPORTANT purchase criterion?



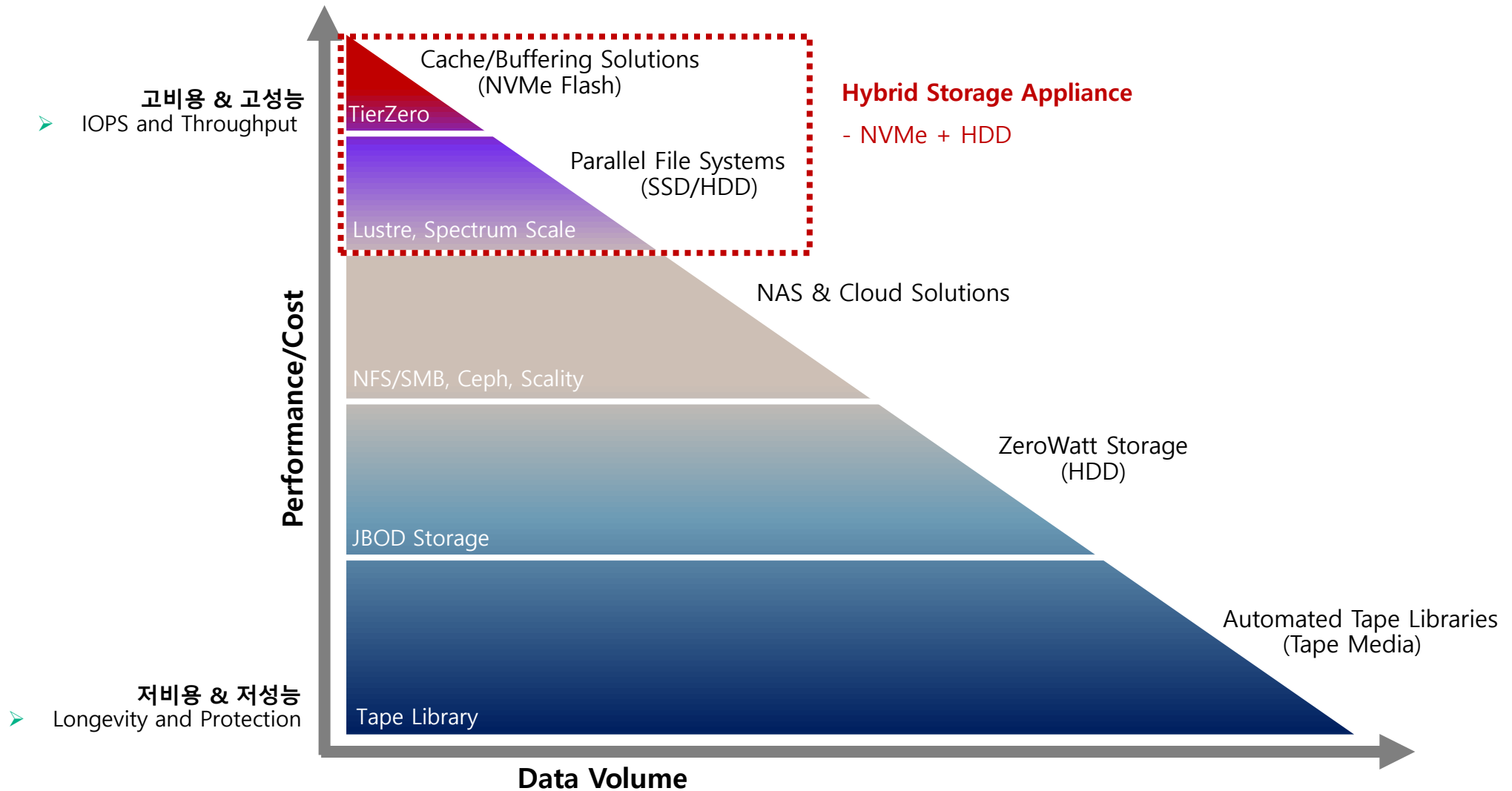
Top 3
Price/**PERFORMANCE**

Source: Hyperion Research, The hidden costs of HPC storage, May 2020

HPC/AI 영역을 위한 스토리지 아키텍처

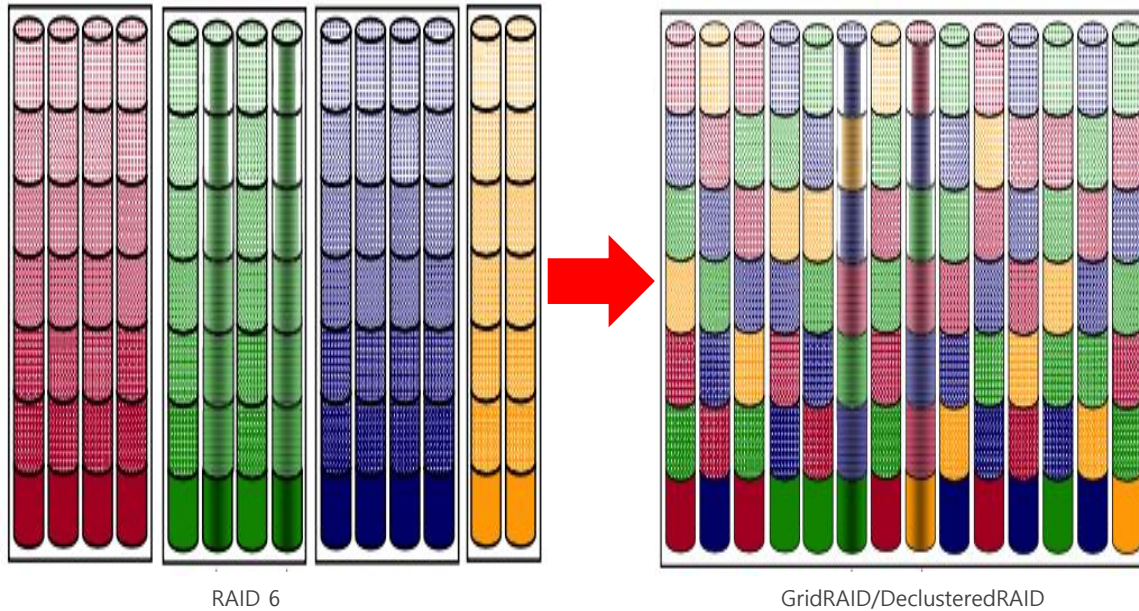


HPC/AI 스토리지 구성

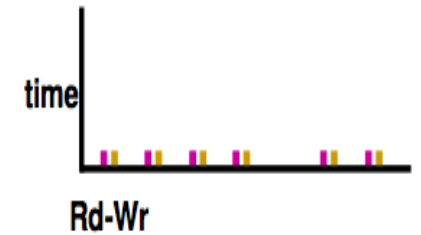
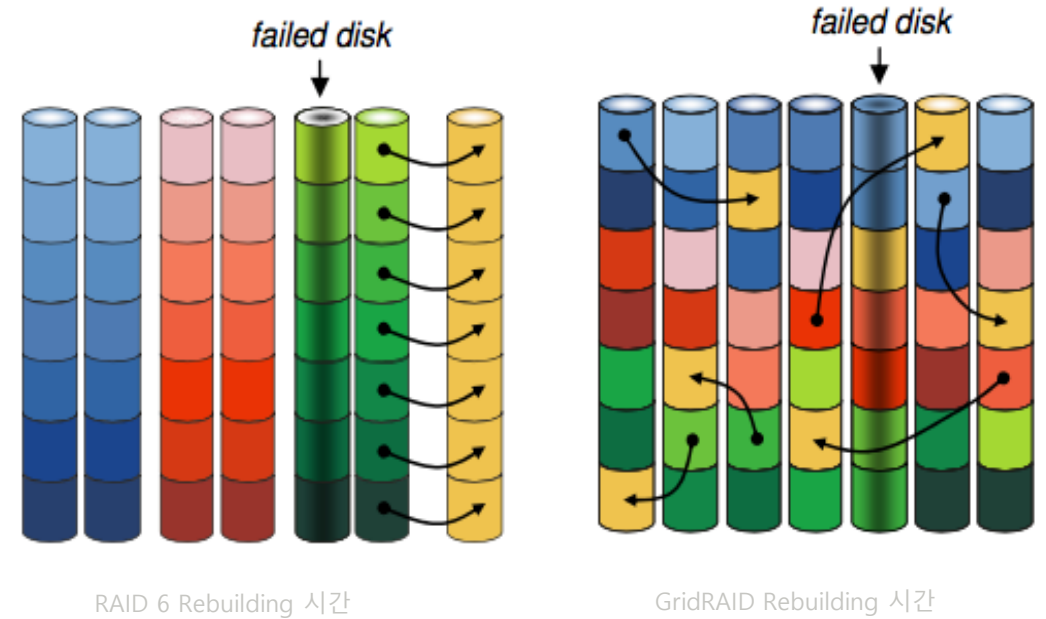


고성능 구현을 위한 블록기반 데이터 분산처리

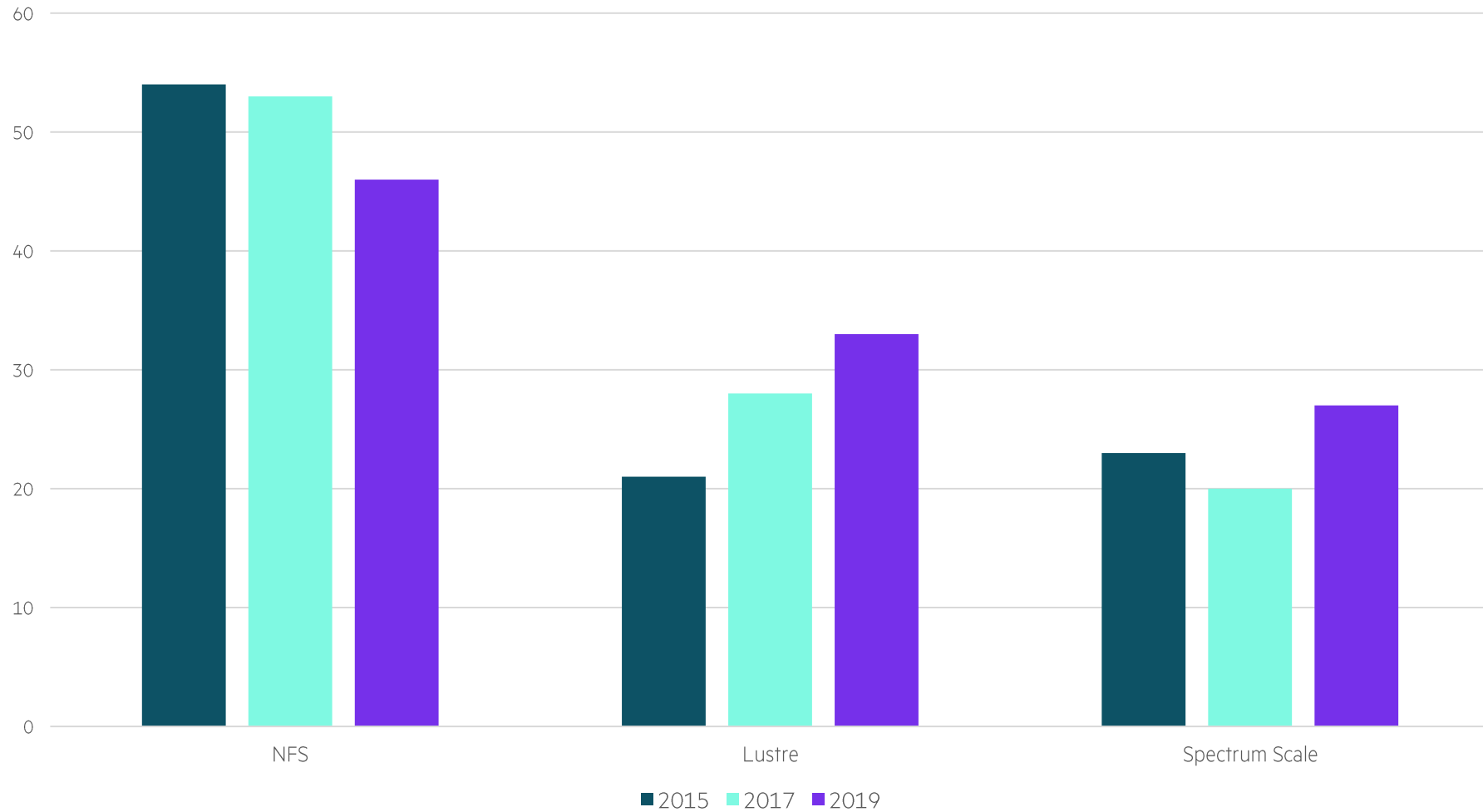
• Declustered Architecture



- 일반 RAID6에 추가기술이 적용되어 모든 디스크를 활용한 블록단위 데이터 처리
 - 데이터의 안정적인 분배 및 모든 디스크 활용으로 인한 데이터 IO 속도 증가
 - 디스크 장애 시 모든 디스크가 Rebuilding에 참여하여 Rebuilding 시간 절약

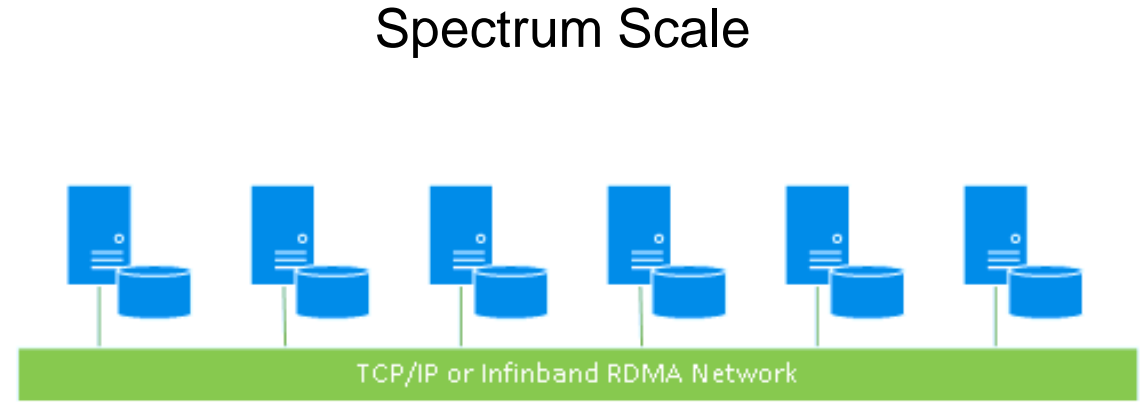
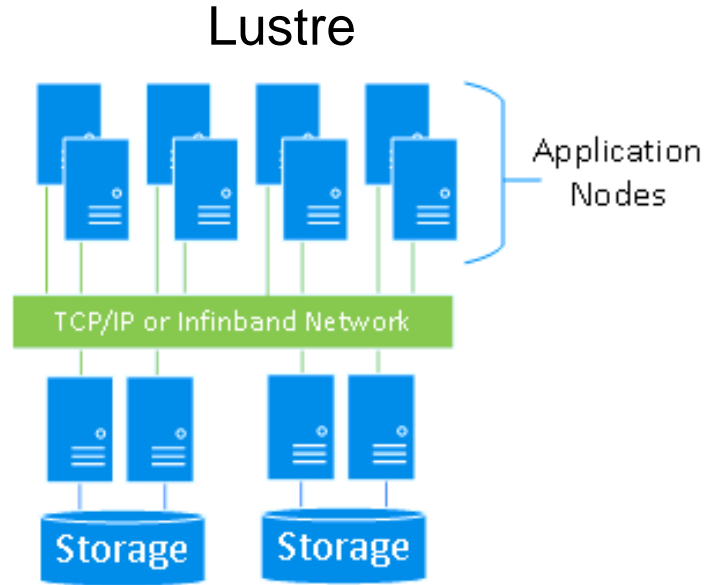


HPC/AI 영역에서의 스토리지 아키텍처 변화



Source: Hyperion Research, Special Study: Shifts Are Occurring in the File System Landscape, June 2020

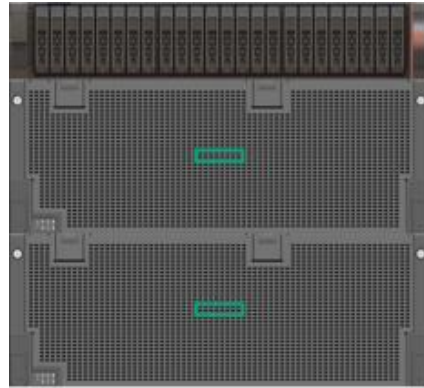
HPC/AI 영역을 위한 고성능 병렬파일시스템



| | Lustre | Spectrum Scale |
|------------|----------------------------------|------------------------------------|
| 용량 및 성능 | > 16ExaByte, Scale-Out unlimited | > 8-YotaBytes, Scale-Out unlimited |
| 개발기간 | 20년 이상 | 20년 이상 |
| 라이선스 | 오픈소스 | 상용 |
| 블럭기반 데이터처리 | 지원 | 지원 |
| 데이터 관리 자동화 | 지원 | 지원 |
| 온라인 증설/변경 | 지원 | 지원 |
| 지원 클라이언트 | 리눅스 | 리눅스/윈도우즈 |

HPC/AI 영역을 위한 고성능 스토리지 어플라이언스

ClusterStor e1000



HPFSS



HPC/AI 전용 어플라이언스

서버 기반 HPC/AI 어플라이언스

병렬파일시스템

Lustre

Spectrum Scale

공급형태

어플라이언스 완제품

어플라이언스 완제품

용량 및 성능

Scale-Out 형태 선형적 향상
용량: 0.5PB ~ EB
성능: > 10GB/s

Scale-Out 형태 선형적 향상
용량: ~ 2PB ~
성능: < xxx GB/s

GPU DirectIO

지원

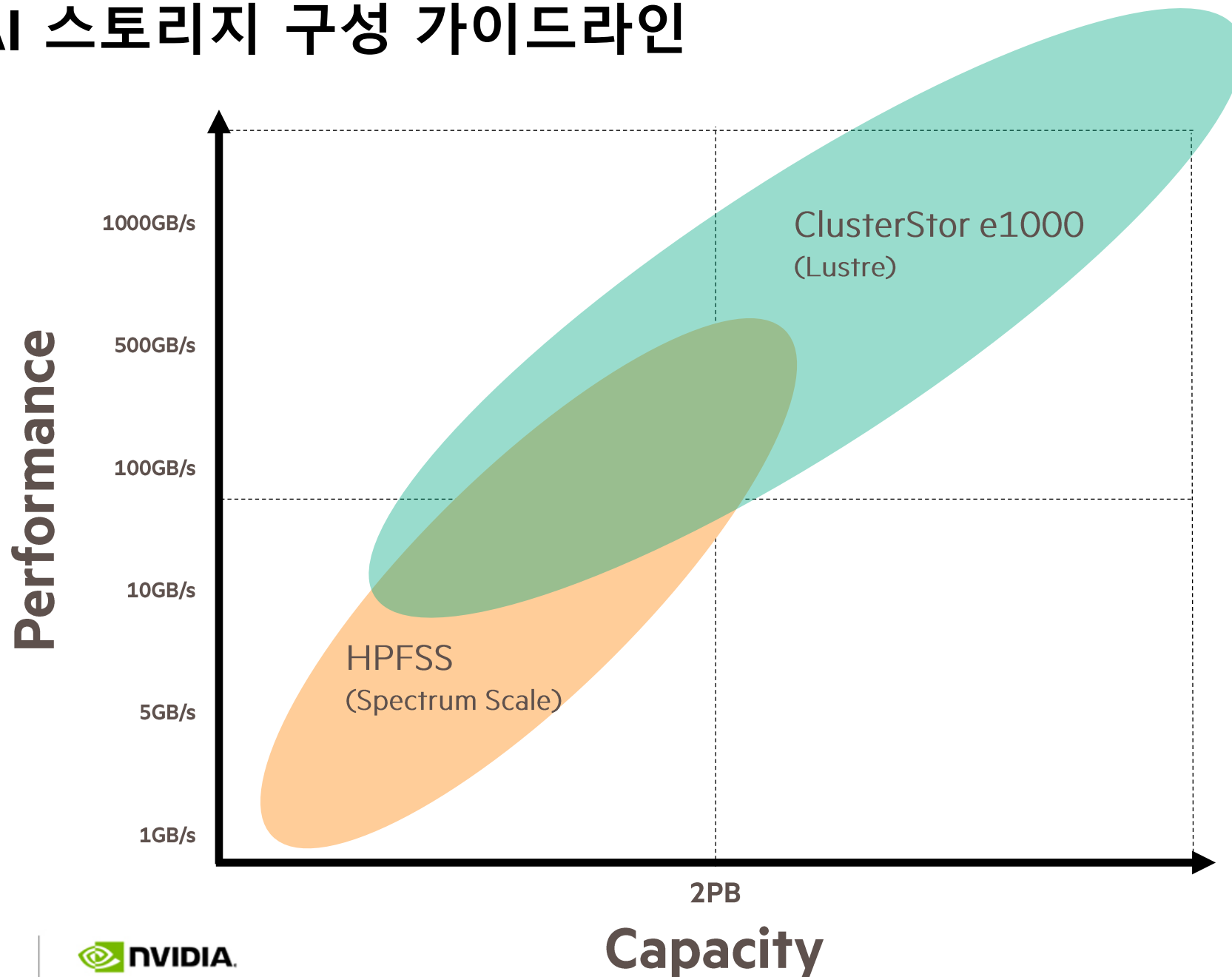
지원

기술지원

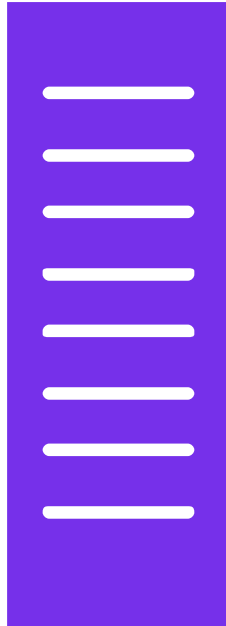
HPE 직접구축 및 유지보수

HPE 직접구축 및 유지보수

HPC/AI 스토리지 구성 가이드라인



CLUSTERSTOR E1000



Base rack

- 네트워크 장치 일체 포함 (Ethernet, Infiniband 등)
- 이중화된 메타데이터/관리엔진

Scale-Out 기반의 선형적 성능향상 및 용량확장

Flash Unit (All Flash)

데이터유닛 1-세트당

- 최대 80 GB/s read
- 최대 60 GB/s write



Disk Unit (Balanced)

데이터유닛 1-세트당

- 최대 30 GB/s read
- 최대 30 GB/s write



Disk Unit (\$/GB optimized)

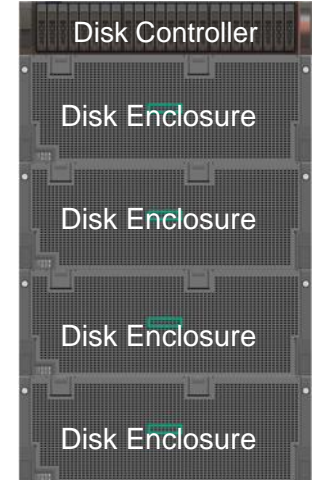
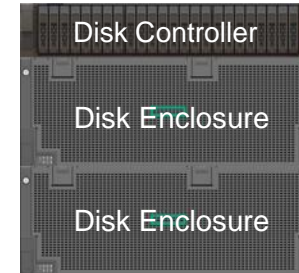
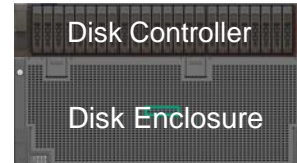
데이터유닛 1-세트당

- 최대 40 GB/s read
- 최대 40 GB/s write



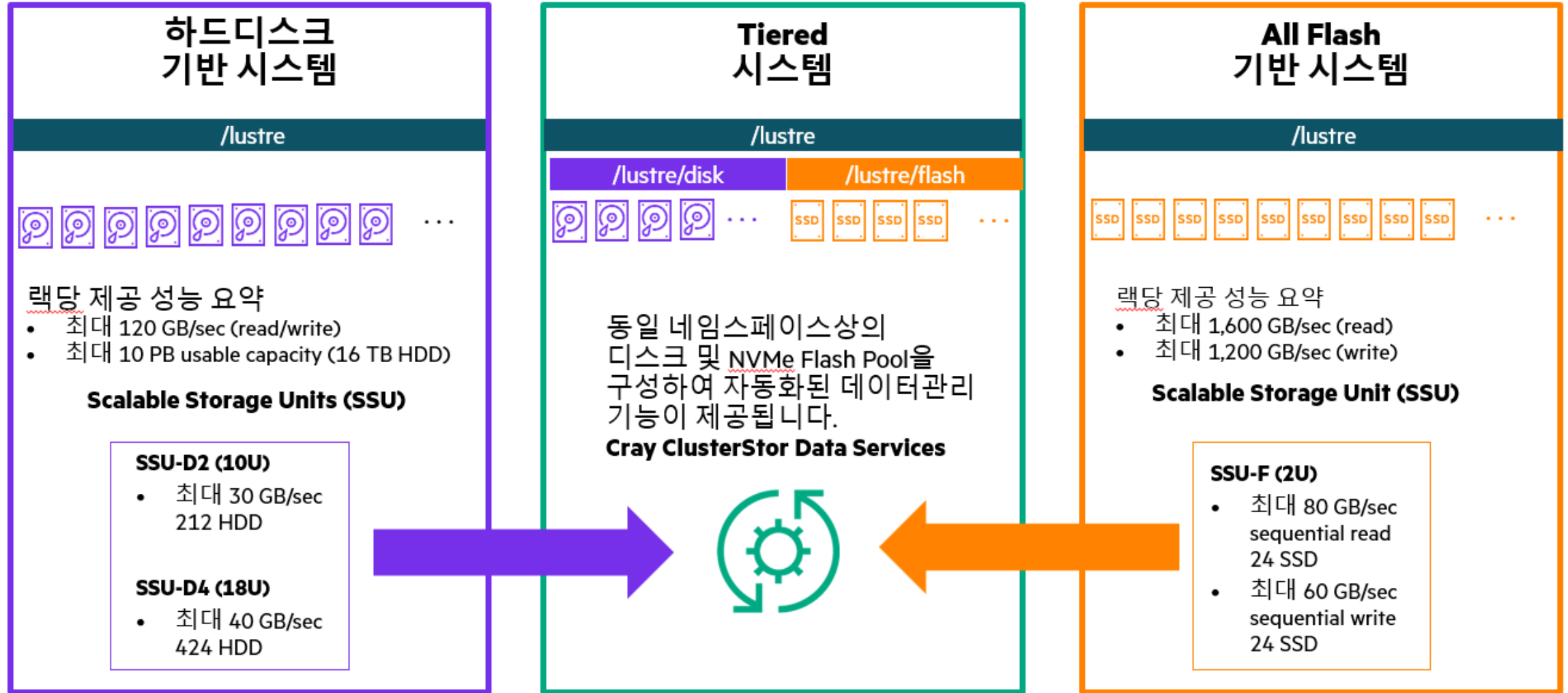
CLUSTERSTOR E1000 > 성능 및 구성

| | Extreme Performance SSU-F | Entry level SSU-D1 | HDD Performance SSU-D2 | HDD Capacity SSU-D4 |
|------------------------------|------------------------------|-----------------------|---------------------------|------------------------|
| SSD Performance (write/read) | 60 / 80 GB/s | | | |
| SSD Usable Capacity (3.2TB) | 55.3 TB | | | |
| HDD Performance | | 15 GB/s | 30 GB/s | 40 GB/s |
| HDD Usable Capacity (16TB) | | 1.25 PB | 2.51 PB | 5.02 PB |
| Network ports | 4 x 200 Gbps | 2 x 200 Gbps | 2 x 200 Gbps | 2 x 200 Gbps |



CLUSTERSTOR E1000 > 데이터 관리 자동화

- 고성능 NVMe/HDD로 이루어진 하이브리드 구성에서 자동화된 데이터티어링 기능이 제공됩니다.

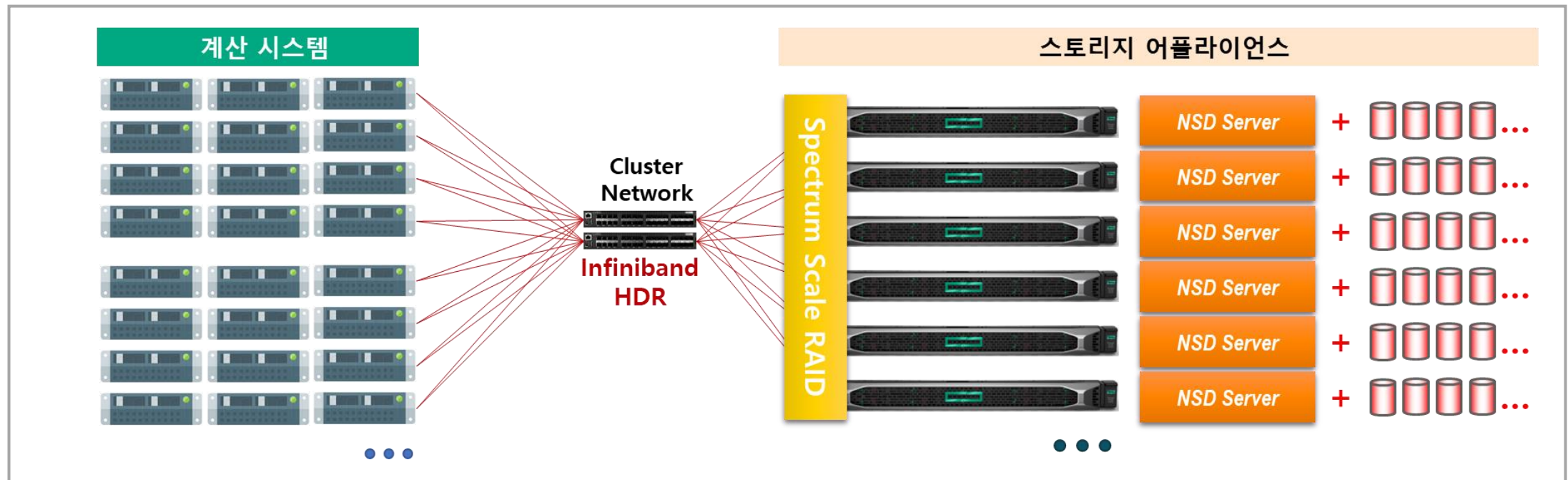


HPFSS(HPE PARALLEL FILE SYSTEM STORAGE)

Flash Storage Server



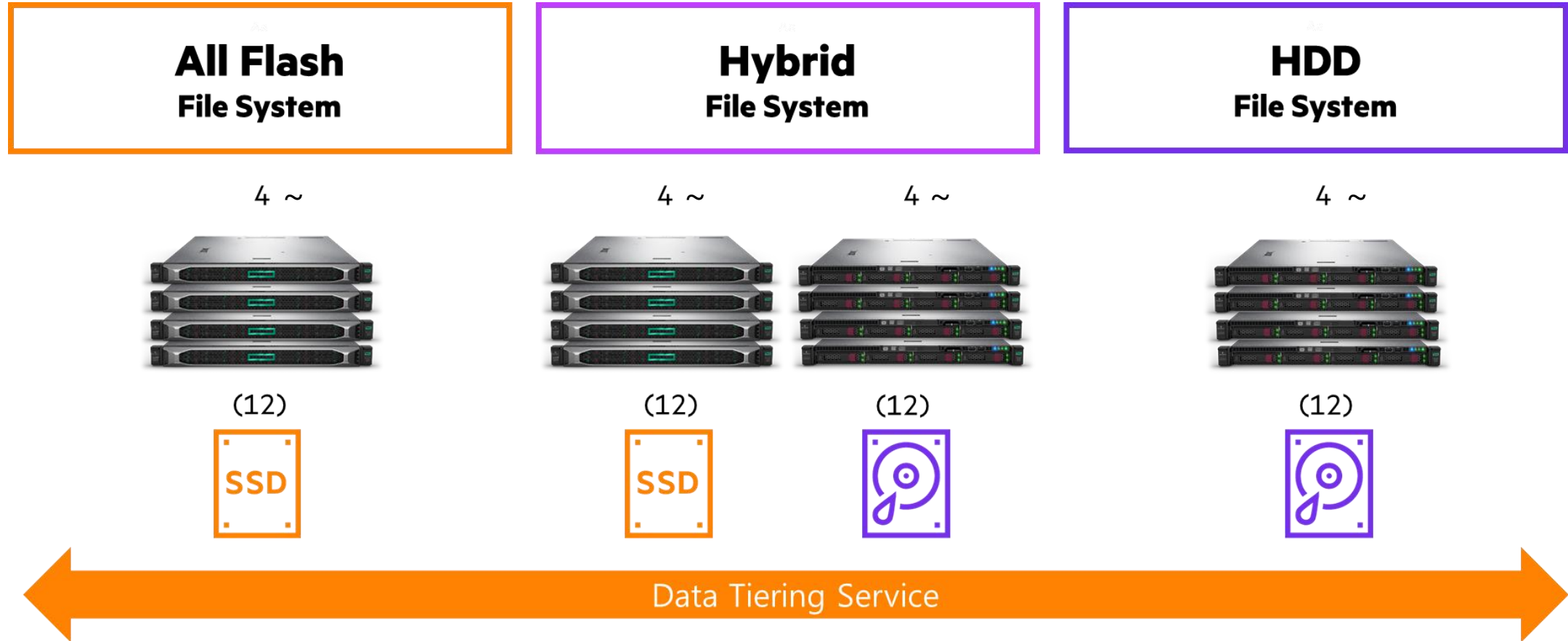
HDD Storage Server



HPFSS > 구성 및 성능

HPE Parallel File System Storage

- 클러스터당 4~32 스토리지노드
- 읽기쓰기 성능: 1~xxx GB/sec



HPC/AI 파일서비스를 위한 고성능 네트워크

HPE Apollo or HPE ProLiant Clusters

(100/200 Gbps interconnect)

Cray Supercomputers

(water-cooled / air-cooled)

Infiniband HDR



QM8790 Mellanox Quantum™

HDR Edge 1U Switch:

- 40 x HDR 200Gb/s ports
- 80 x HDR100 100Gb/s ports

200 Gigabit Ethernet



Cray Slingshot 200 Gbps

1U Switch

- 64 x 200Gb/s ports
- Ethernet compatible

Cray Slingshot

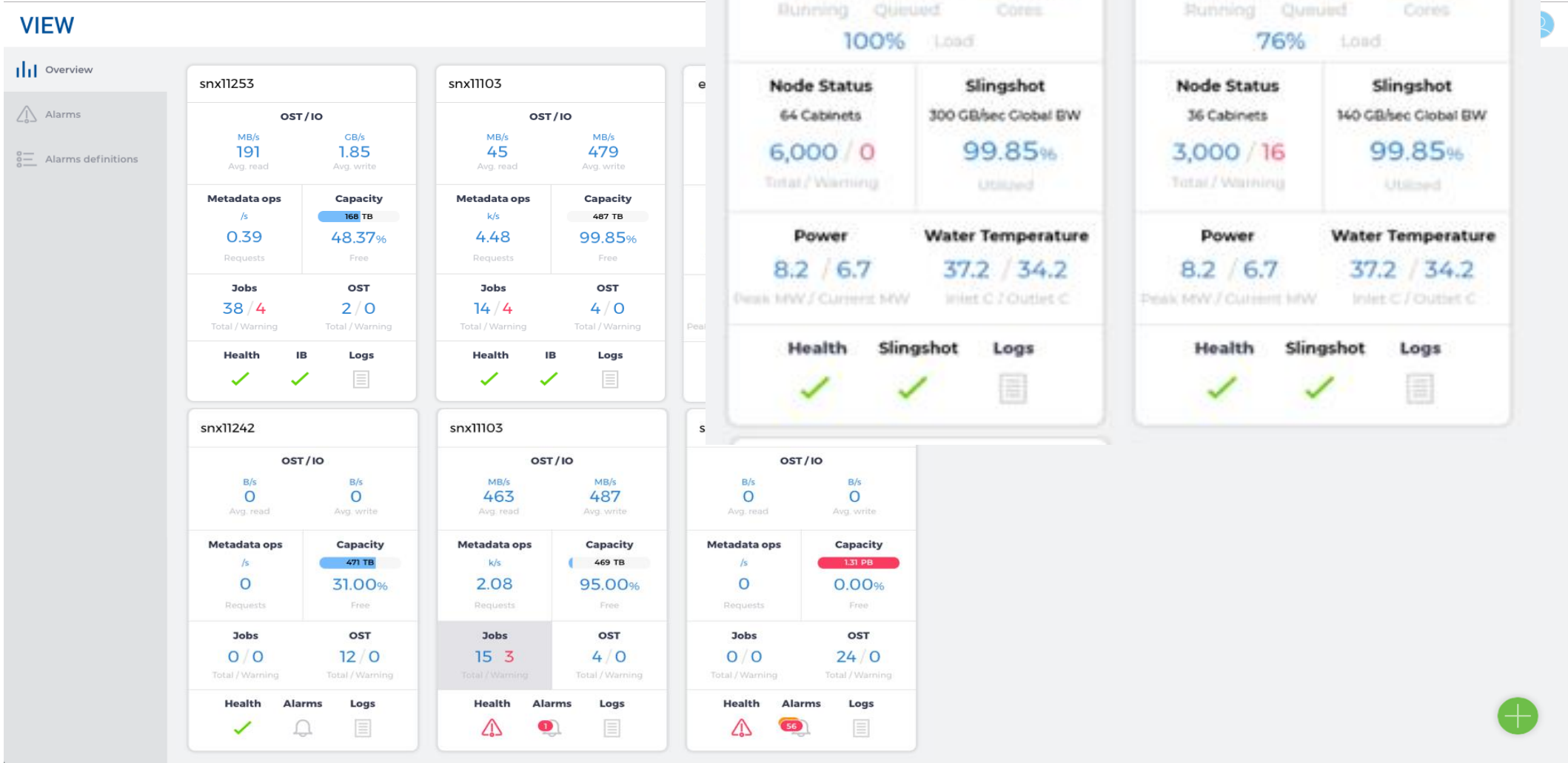


Cray Slingshot 200 Gbps

1U Switch

- 64 x 200Gb/s ports
- 32 cables

효과적 스토리지 관리



계산자원과 연계한 진보된 스토리지 분석활용

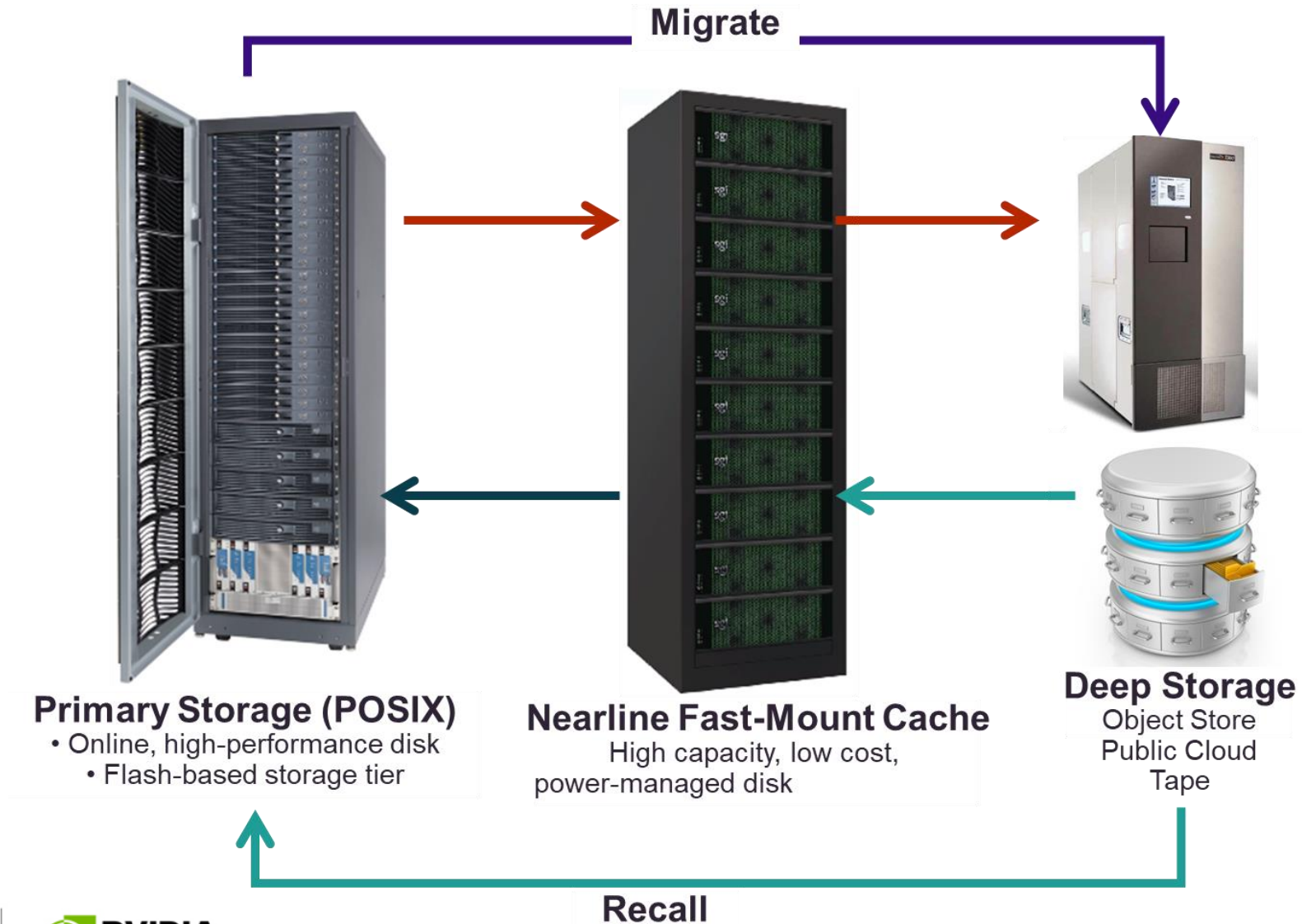
가시화

스케줄러와 연동된
Job 관련 읽기/쓰기/
메타데이터
부하 상세 정보



Job 특성비교
특정 JOB과 기타
일반 Job
동작현황에 대한
직관적 비교

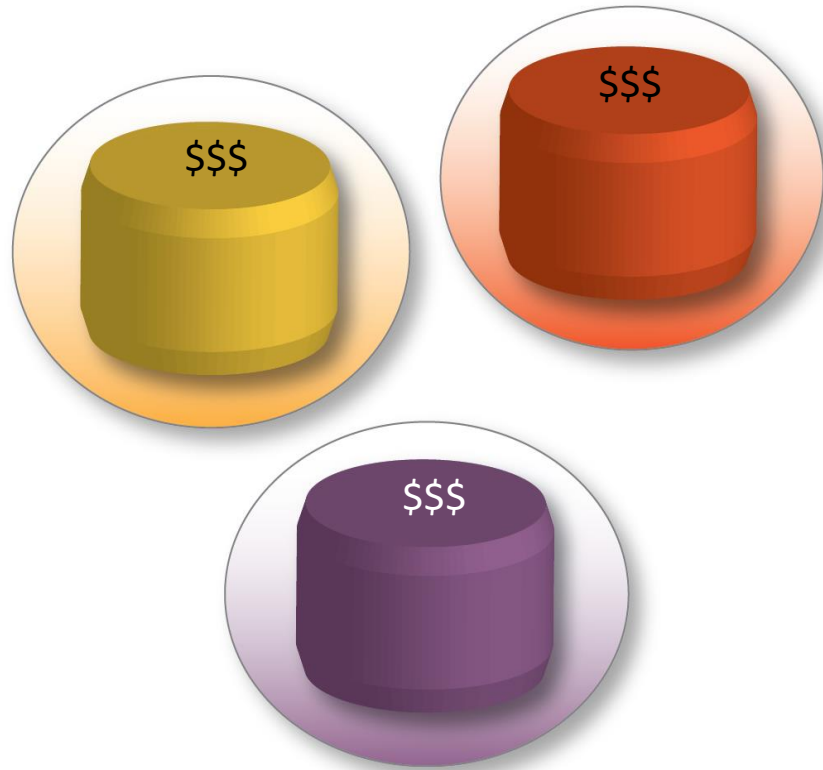
DATA MANAGEMENT > 이기종간 데이터관리 자동화



DATA MANAGEMENT

기존 환경

각각의 스토리지에 분산된 데이터

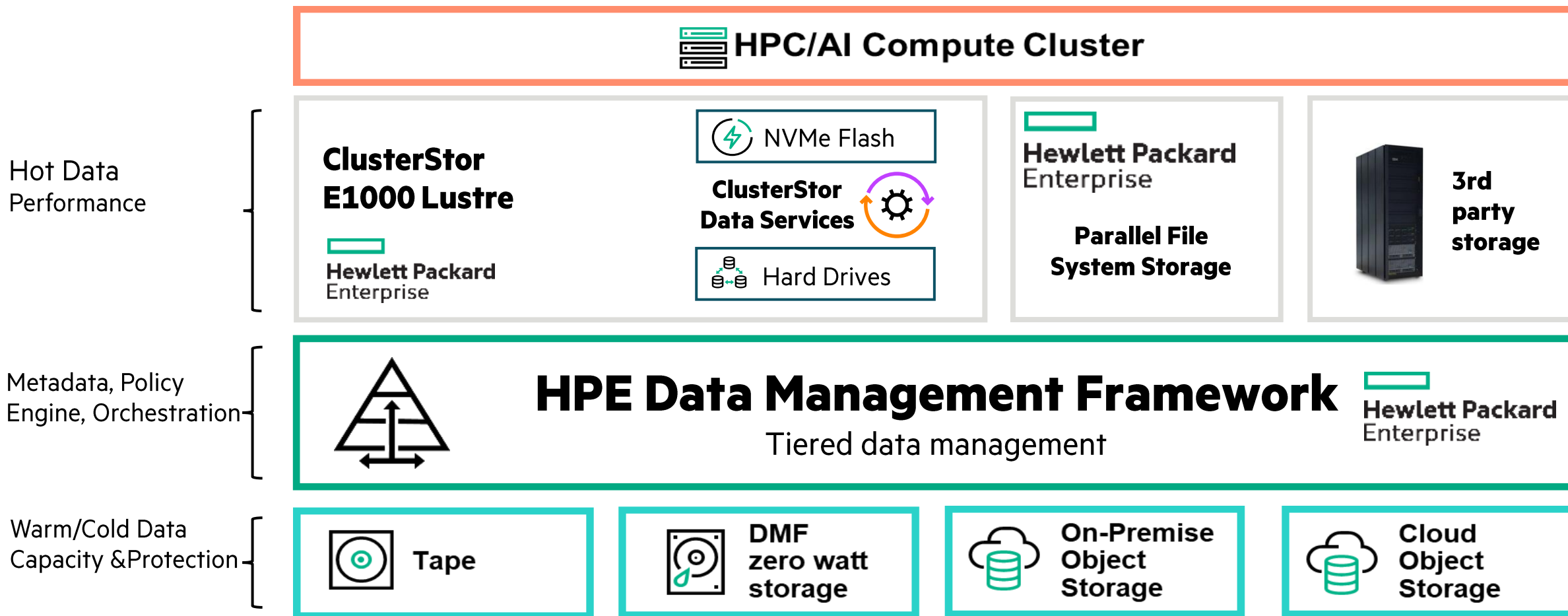


데이터 관리기법 적용

계층화된 스토리지 융합에 의한 데이터관리 자동화



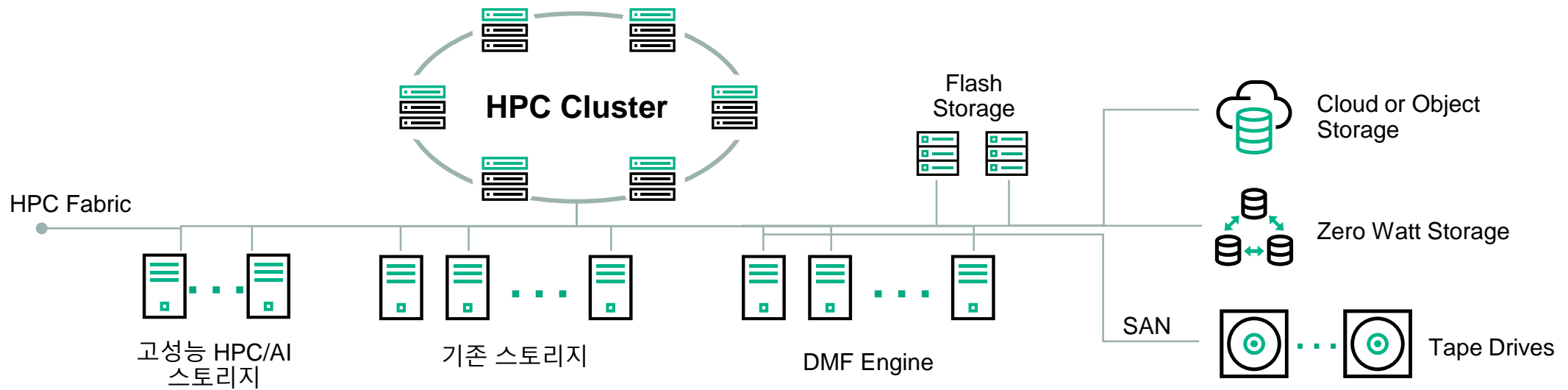
DATA MANAGEMENT > HPE DMF



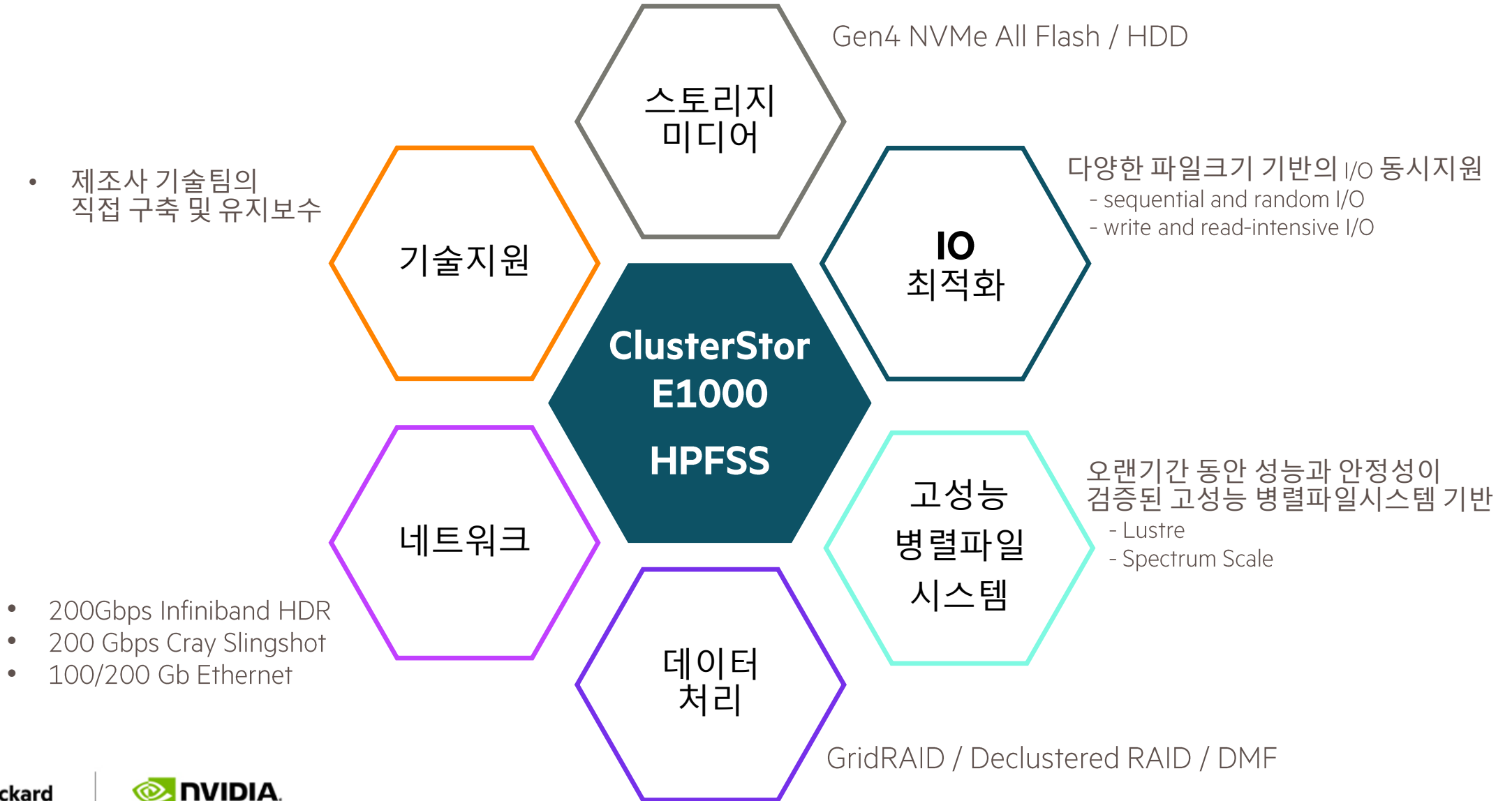
DATA MANAGEMENT > DMF 아키텍처

- ✓ HPC/AI Job 연계한 데이터 이동/저장 자동화
- ✓ Lustre / Spectrum Scale / 기존 스토리지를 효과적으로 연동
- ✓ 다양한 구성요소(Disk, Tape, Cloud...)와 연계된 통합데이터 관리체계

DMF Architecture



HPC/AI 전용 고성능 저장장치 요약



HPE STORAGE REFERENCE

OAK RIDGE
National Laboratory



~ 10 TB/sec
> 700 PB



> 4 TB/sec
30 PB (All Flash)

Lawrence Livermore
National Laboratory



2.5 TB/sec
~ 400 PB

Argonne
NATIONAL LABORATORY



1 TB/sec
~ 200 PB



Cray ClusterStor E1000

감사합니다

Contact: hpc@hpe.com