

2024년 전망, 금융IT Innovation 컨퍼런스

AI를 위한 GPU팜 구축 시, 스토리지 설계 방안

IBM Technology
전상호 상무

GPU팜 구축 시, 스토리지 설계 방안



참조 문헌



NVIDIA DGX SuperPOD: Scalable Infrastructure for AI Leadership

Reference Architecture

Featuring NVIDIA DGX A100 Systems



NVIDIA DGX SuperPOD: Next Generation Scalable Infrastructure for AI Leadership

Reference Architecture

Featuring NVIDIA DGX H100 Systems

‘NVIDIA 백서’에서 제안하는 Storage 고려사항

1. Performance

- ✓ ML/DL에 지연이 발생되지 않을 스토리지 성능에 대한 검토 필요

2. Integration with other systems

- ✓ 데이터 수집, 분석, 저장 및 배포의 관리 프로세스 자동화

3. Cost

- ✓ 계층적 스토리지 구성을 통한 성능, 용량 및 비용의 균형

Performance

DGX SuperPOD



1 Scalable Units (SU)
A100 : 20 대
H100 : 32대



Storage with
Distributed File system

Performance

DGX SuperPOD



1 Scalable Units (SU)
A100 : 20 대
H100 : 32대

- **Storage Fabric은 InfiniBand를 사용하여 RDMA를 활용하십시오.**
 - ✓ Provide single node bandwidth more than 40 GBps.
 - ✓ Provides single-node bandwidth of at least 40 GBps to each DGX H100 system.
 - ✓ An IB-based fabric is essential to maximize bandwidth since the I/O per-node requirements for the DGX SuperPOD is to exceed 40 GBps.
 - ✓ Storage is provided over InfiniBand and leverages RDMA to provide maximum performance and minimize CPU overhead.



Storage with
Distributed File system

Performance

DGX SuperPOD

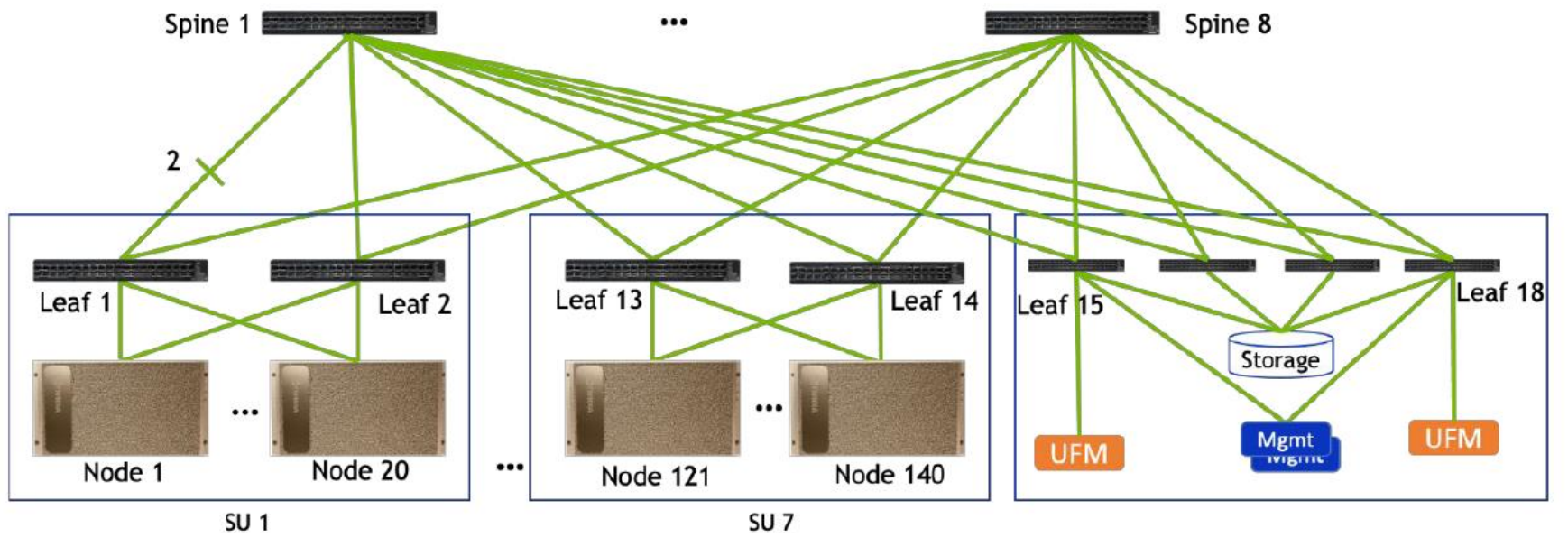


A100

Storage Fabric

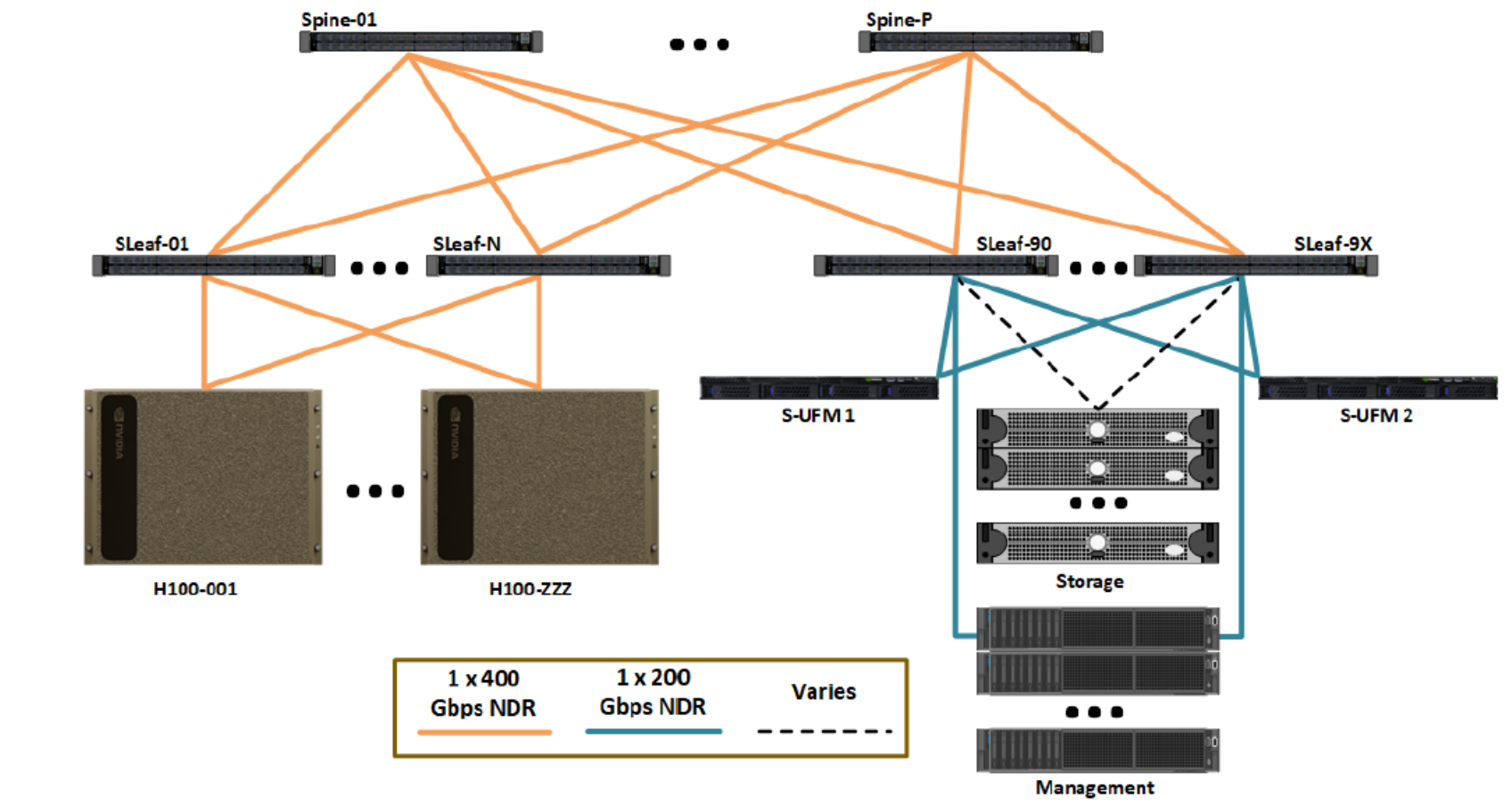
The storage fabric employs an IB network fabric (Figure 7). An IB-based fabric is essential to maximize bandwidth since the I/O per-node requirements for the DGX SuperPOD is to exceed 40 GBps. High-bandwidth requirements with advanced fabric management features, such as congestion control and adaptive routing, provide significant benefits for the storage fabric.

Figure 7. Storage fabric topology for 140-node system



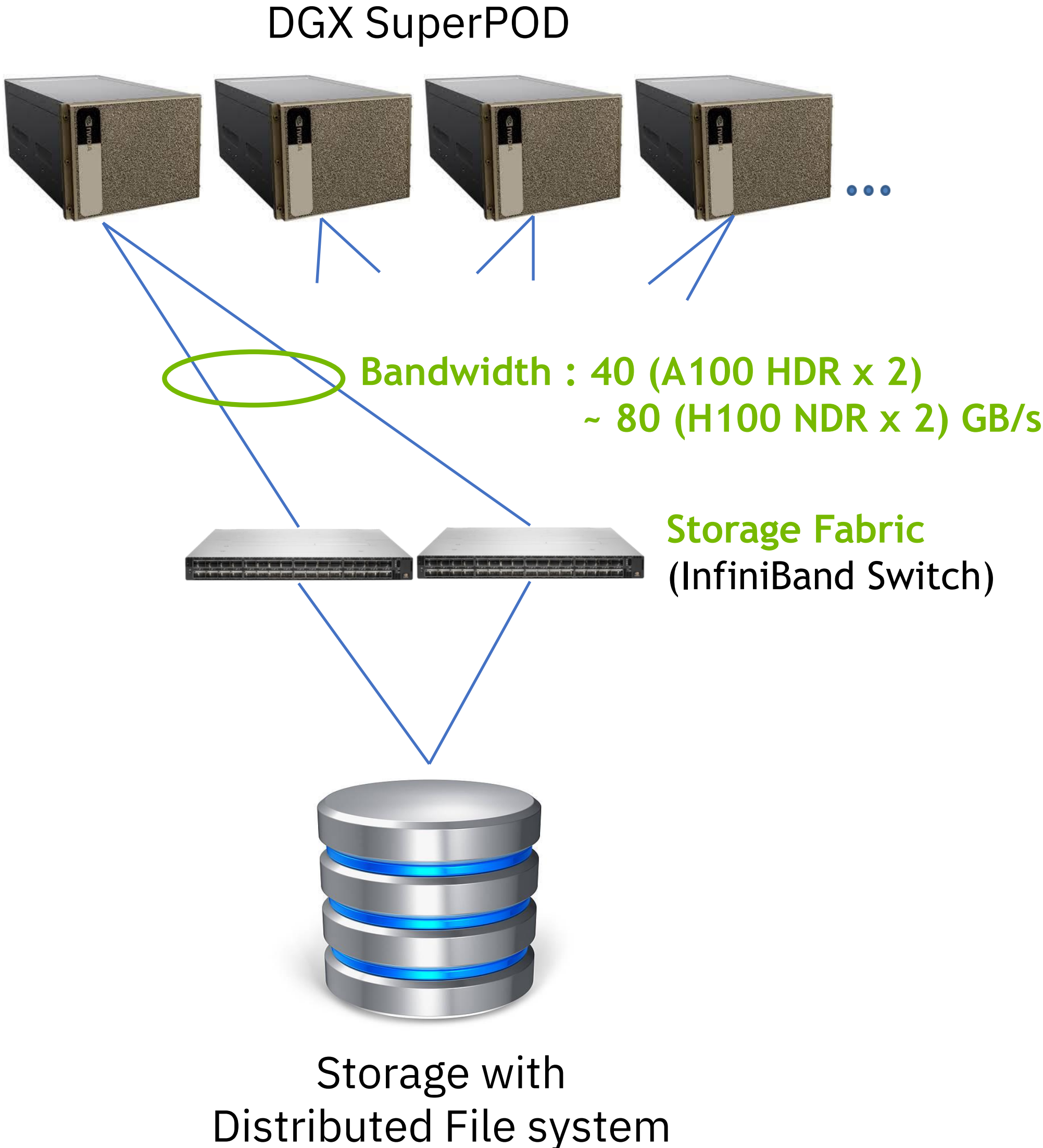
H100

Figure 6. InfiniBand storage fabric logical design



Storage with Distributed File system

Performance



Performance

DGX SuperPOD



- **스토리지가 병목이 되지 않도록 충분한 성능을 고려하십시오.**
 - ✓ Maximize storage access performance from a single SU.
 - ✓ **A100** : Allow for training of DL models that require peak I/O performance, exceeding 16 GBps (**2 GBps per GPU**) directly from remote storage.
 - ✓ **H100** : Use cases in automotive and other computer vision-related tasks, **4 GBps per GPU** for read performance is needed. ... Even for the best category above, it is desirable that the single node read performance is **closer** to the maximum network performance of **80 GBps**.
 - ✓ If you are looking for best end-to-end training performance, do not ignore I/O operations for **checkpoints**.



Storage with
Distributed File system

Performance

DGX SuperPOD

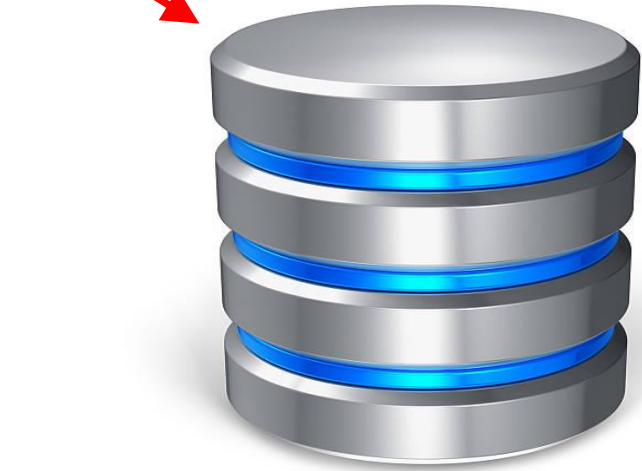


Bandwidth : 40 (A100 HDR x 2)
~ 80 (H100 NDR x 2) GB/s

DGX 1식당 요구 성능

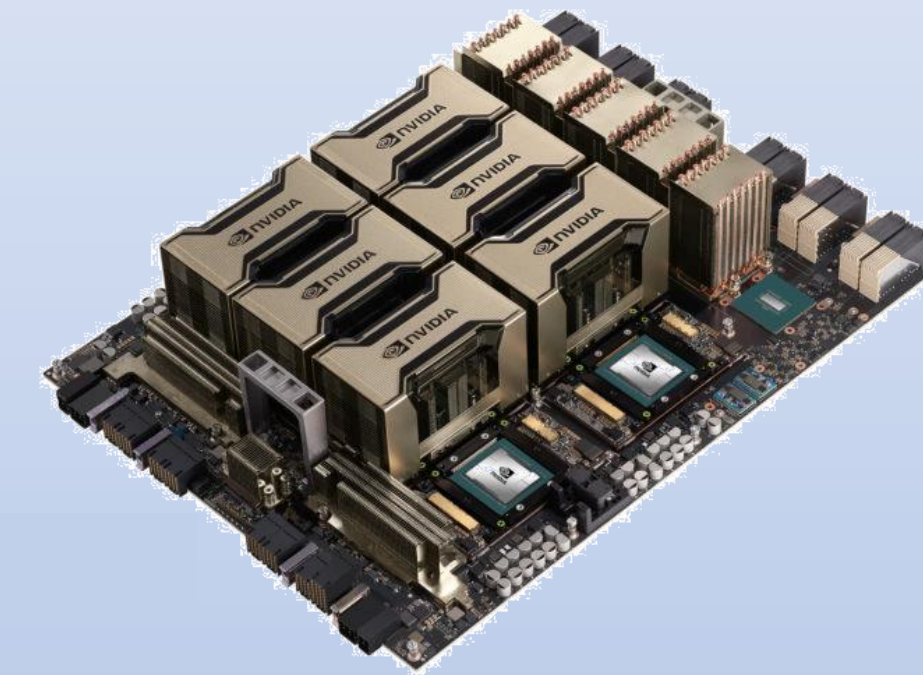
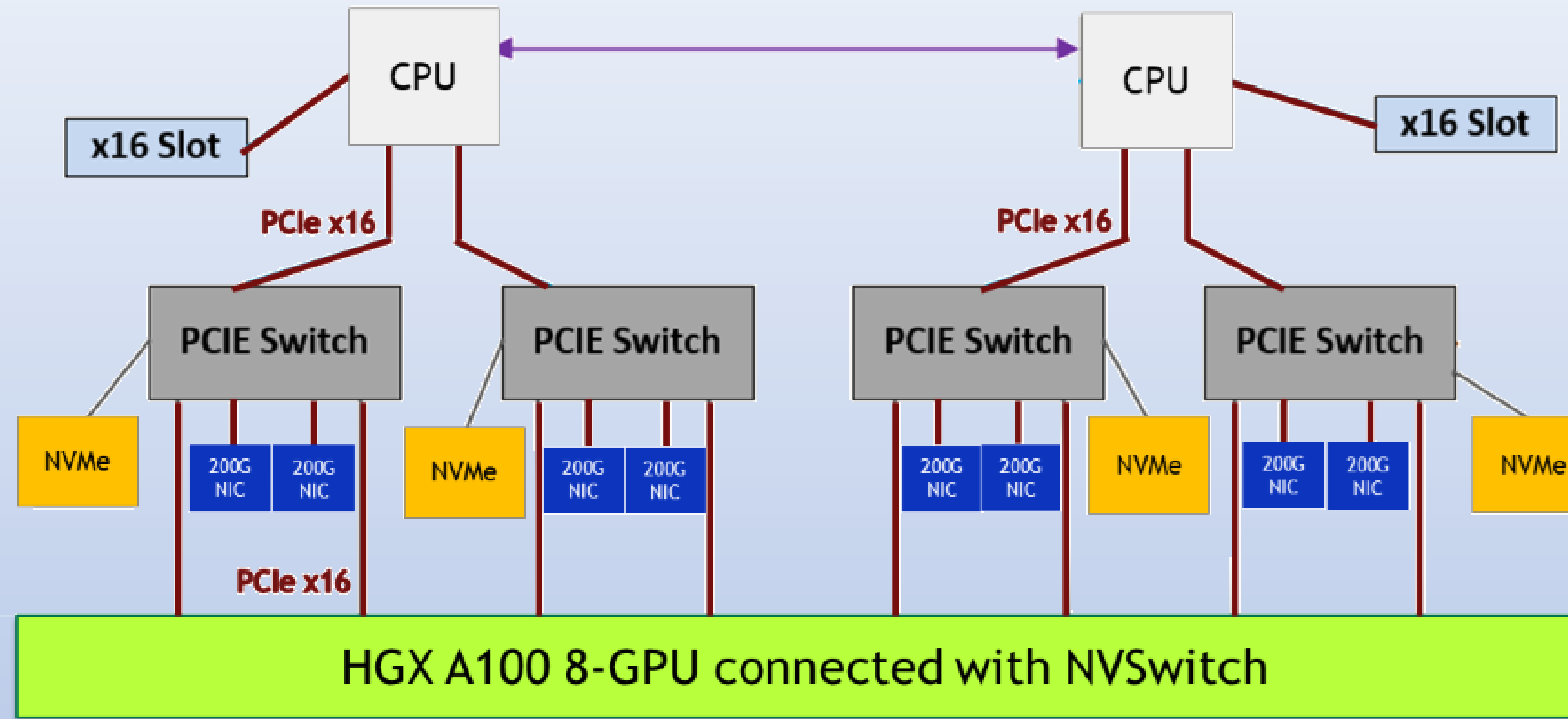
Use RDMA.
A100 : Peak I/O performance > 16 GB/s
H100 : Peak I/O performance > 32 GB/s
→ 80GB/s

Storage Fabric
(InfiniBand Switch)



Storage with
Distributed File system

참고) HGX A100 board



PCI 버전	날짜		인코딩	데이터 전송률	대역폭				
	발표	적용			1레인 (x1)	2레인 (x2)	4레인 (x4)	8레인 (x8)	16레인 (x16)
1.0~1.1	2003년	2004년 6월	8b/10b	2.5 GT/s	250 MB/s	500 MB/s	1 GB/s	2 GB/s	4 GB/s
2.0~2.1	2007년 1월	2007년 8월	8b/10b	5 GT/s	500 MB/s	1 GB/s	2 GB/s	4 GB/s	8 GB/s
3.0~3.1	2010년 11월	2011년 7월	128b/130b	8 GT/s	984.6 MB/s	1.969 GB/s	3.938 GB/s	7.877 GB/s	15.754 GB/s
4.0	2017년 6월	2018년	128b/130b	16 GT/s	1.969 GB/s	3.938 GB/s	7.877 GB/s	15.754 GB/s	31.508 GB/s
5.0	2019년 5월	2020년	128b/130b	32 GT/s	3.938 GB/s	7.877 GB/s	15.754 GB/s	31.508 GB/s	63.015 GB/s
6.0	2022년 1월	미정	1b/1b	64 GT/s	7.877 GB/s	15.754 GB/s	31.508 GB/s	63.015 GB/s	126.031 GB/s

Performance

DGX SuperPOD

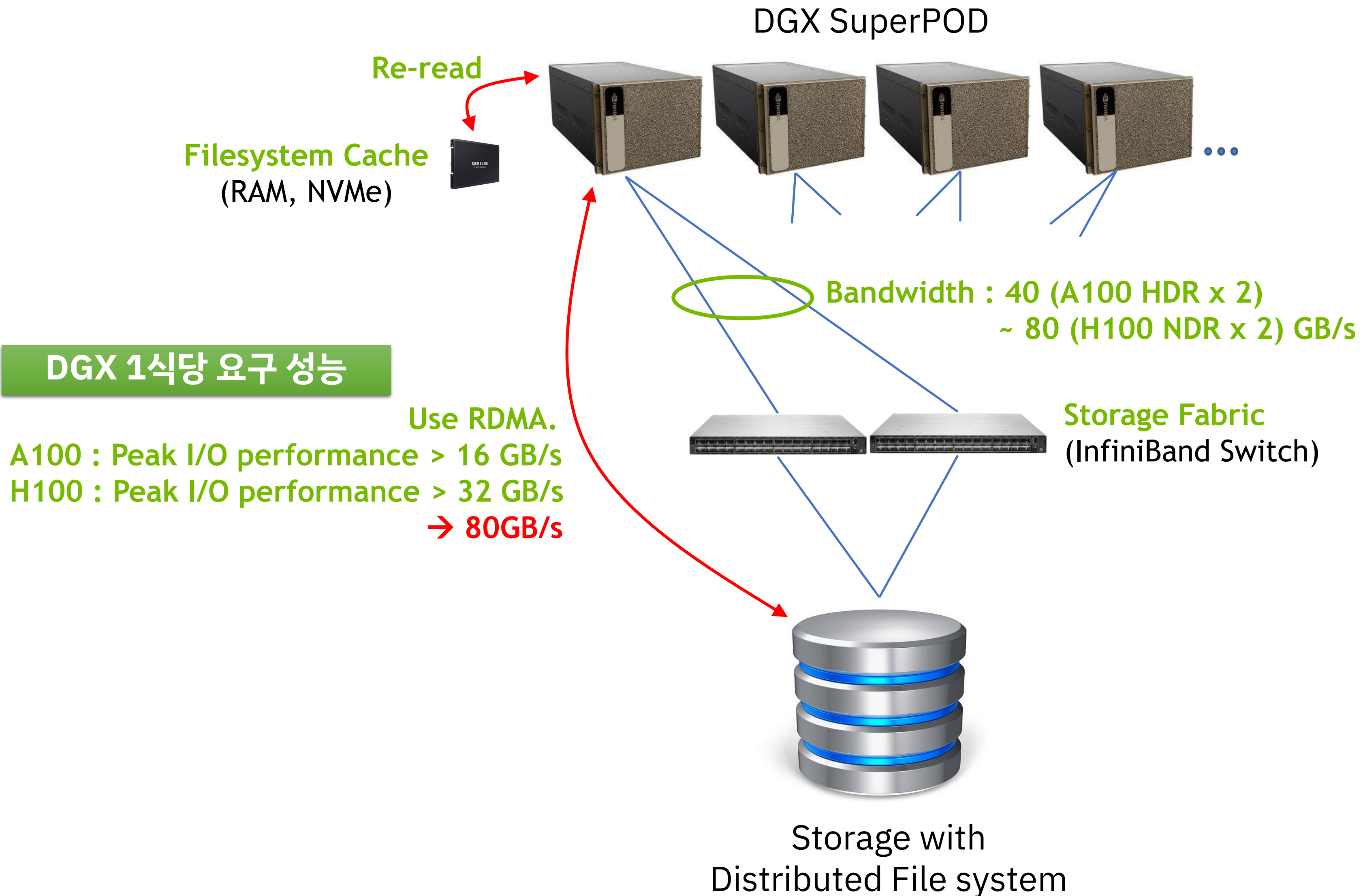
- **GPU 서버에서 데이터 Cache read hit%를 높일 수 있는 구성을 검토하십시오.**
 - ✓ The key to performance is the ability to read data multiple times.
 - ✓ The closer the data are cached to the GPU, the faster they can be read.
 - ✓ Caching data in local RAM provides the best performance for reads. This caching is transparent once the data are read from the filesystem. However, the size of RAM is limited and less cost effective than other storage and memory technologies. Local NVMe storage is a more cost-effective way to provide caching close to the GPUs. 로컬 디스크를 자동으로 활용하는 방법(예: NFS 파일 시스템용 `cachefilesd`)이 있지만, 모든 네트워크 파일 시스템이 그렇게 할 수 있는 방법을 제공하는 것은 아닙니다.

→ 80GB/S



Storage with
Distributed File system

Performance



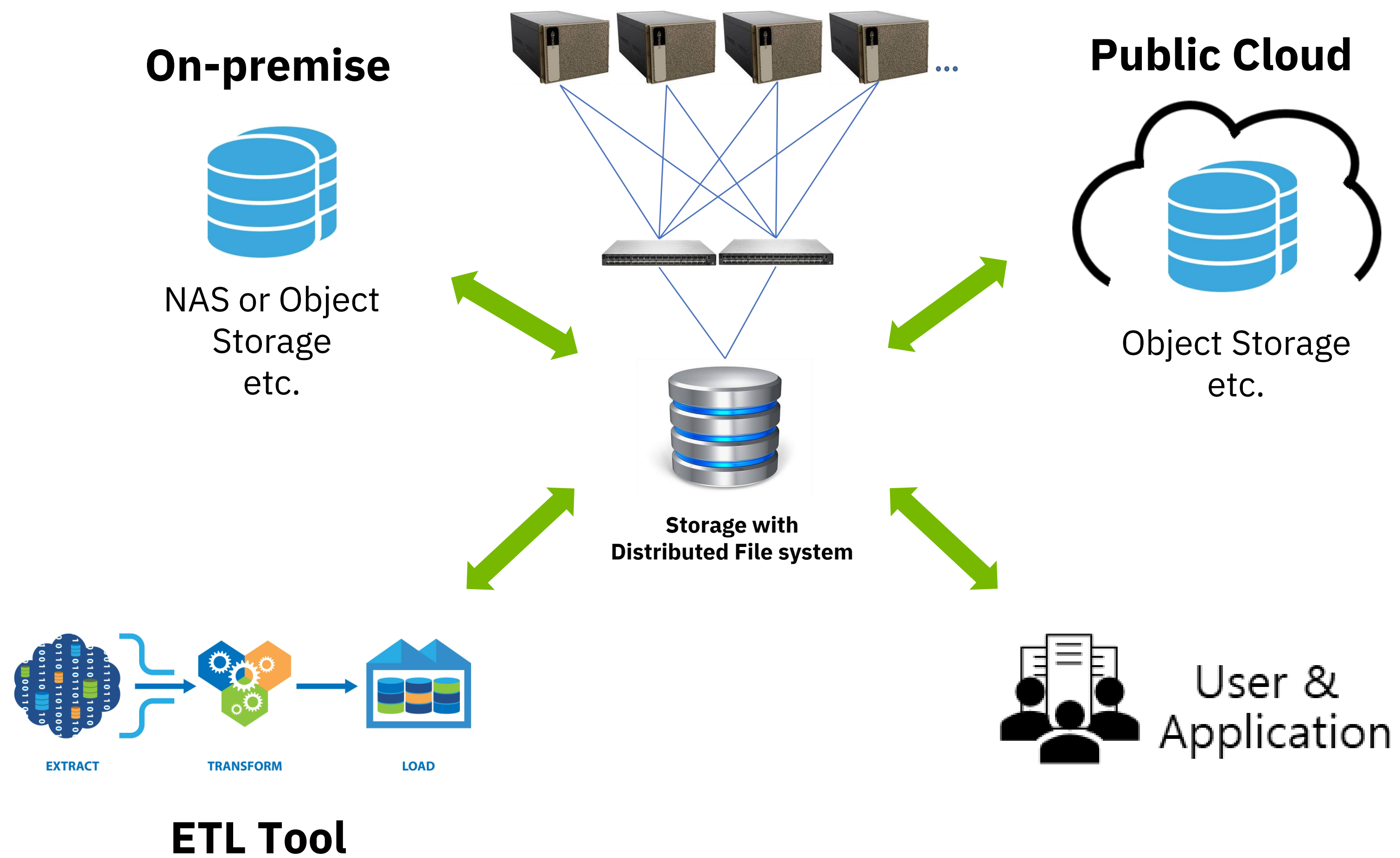
DGX 1식당 요구 성능

Use RDMA.
A100 : Peak I/O performance > 16 GB/s
H100 : Peak I/O performance > 32 GB/s
→ 80GB/s

Integration with other systems

- **타 시스템 or Storage와의 연동 방안을 고려하십시오.**
 - ✓ Provide additional connectivity to shared storage between the DGX SuperPOD and other resources in (out of) the data center.

Integration with other systems



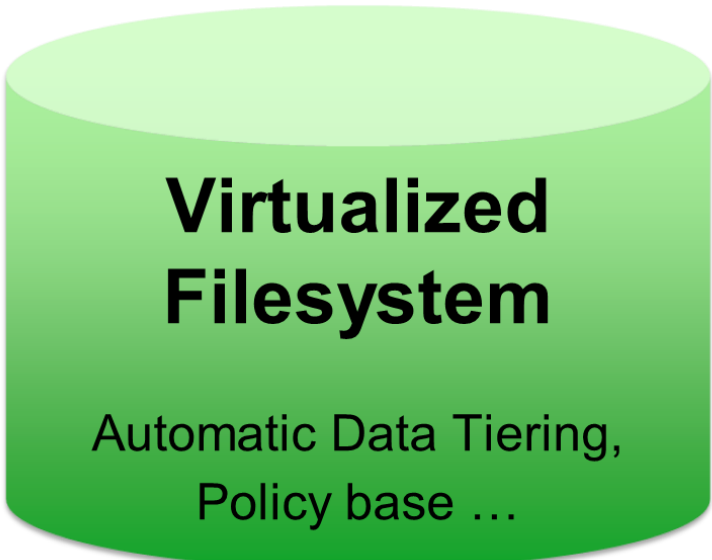
Cost

- **Storage의 요구 성능, 용량 및 비용 최적화를 위해 Data Tiering 방안을 고려하십시오.**
 - ✓ Storage architecture and design must consider the hierarchy of different storage technologies, either persistent or non-persistent, to balance the needs of performance, capacity, and cost.

Cost



Storage with Distributed File system



Public Cloud



Object Storage etc.



NVMe U.2 Disk



NL-SAS Disk



Tape

GPU팜 스토리지 고려사항

1. Performance

- ✓ ML/DL에 지연이 발생되지 않을 스토리지 성능에 대한 검토 필요

2. Integration with other systems

- ✓ 데이터 수집, 분석, 저장 및 배포의 관리 프로세스 자동화

3. Cost

- ✓ 계층적 스토리지 구성을 통한 성능, 용량 및 비용의 균형

IBM ESS3500



전면



후면



Throughput

Read 277GByte/s
Write 155GByte/s

IOPS

13M 4KB random reads/sec

Controller H/W spec

Dual AMD EPYC Genoa 48C processor

- ✓ AMD EPYC 9454 48-Core Processor

Memory

- ✓ 768GB or 1.5TB

Host Interface Card x 2개

- ✓ NVIDIA CX7 supported cards:
- ✓ 400Gb single port (IB only) x16 Gen5
- ✓ 200Gb VPI dual port (IB/ETH) x16 Gen5

NVMe Enclosure

24 or 48 NVMe

- ✓ NVMe SSDs (3.84TB, 7.68 TB, 15.36 TB or 30.74TB)
- ✓ NVMe FCM (19TB or 57TBe and 38TB or 114TBe)

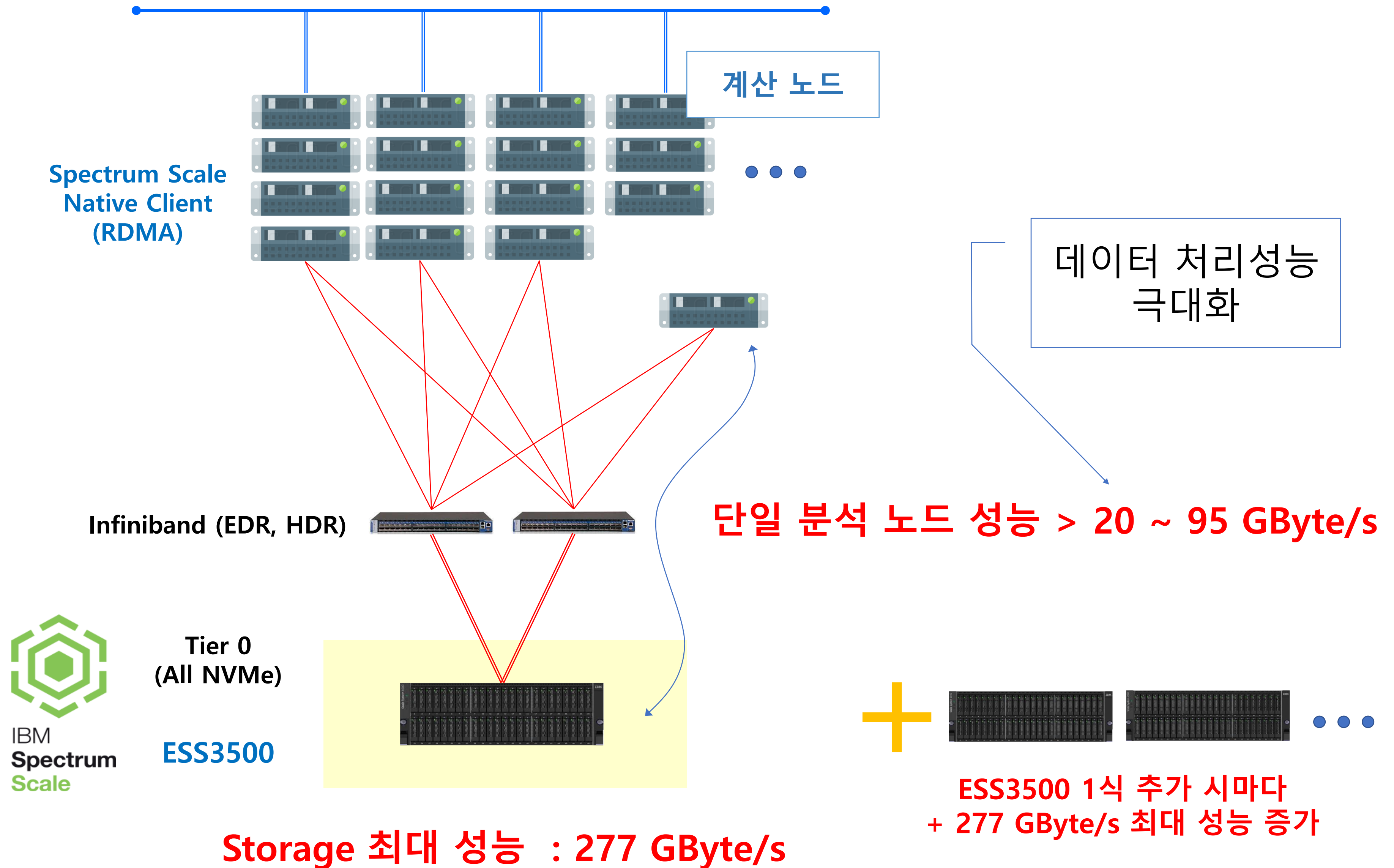
**IBM
ESS3500**

Performance





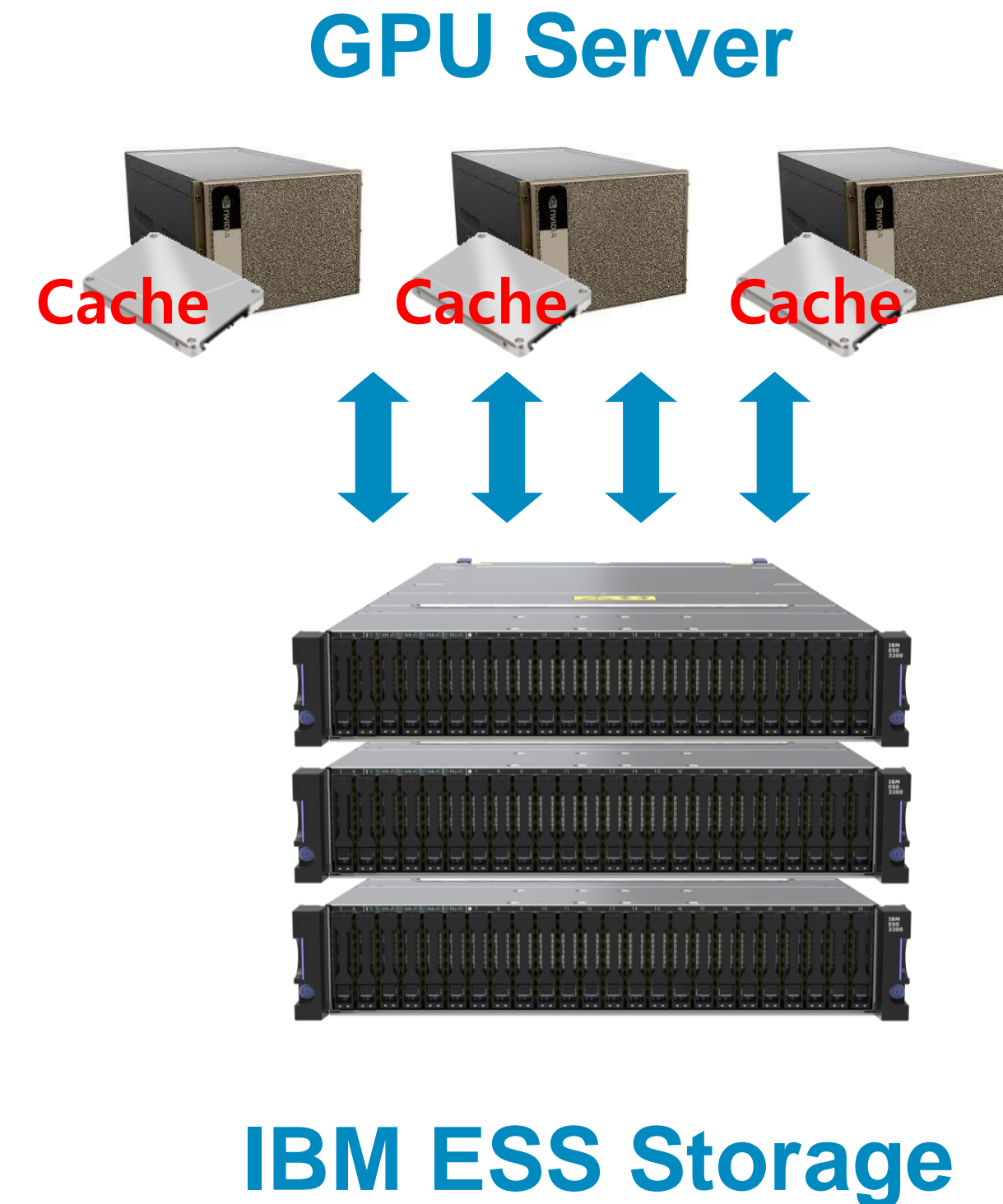
ESS3500의 성능



FS Cache 확장 > Flash Local Read Only Cache (LROC)

- SSD Disk를 Spectrum Scale Client Node에 직접 설치하여 File System Cache로 사용.
- CPU의 Wait I/O를 줄여 I/O 성능을 최대 6배까지 향상시켜 줌.
- 또한, 전체적인 Network 부하를 감소시키고, 노드간 통신 성능을 향상시킴.
- 공유 스토리지의 모든 관리상 장점을 유지하면서 Application 성능을 향상시킴.
- Cache 일관성은 Spectrum Scale 토큰에 의해 보장됨
- Data는 checksum 및 읽기검증에 의해 보호됨
- Spectrum Scale는 코드 수정 없이 Low latency로, Application 에 투명하게, Cache 관리를 자동으로 처리 합니다.

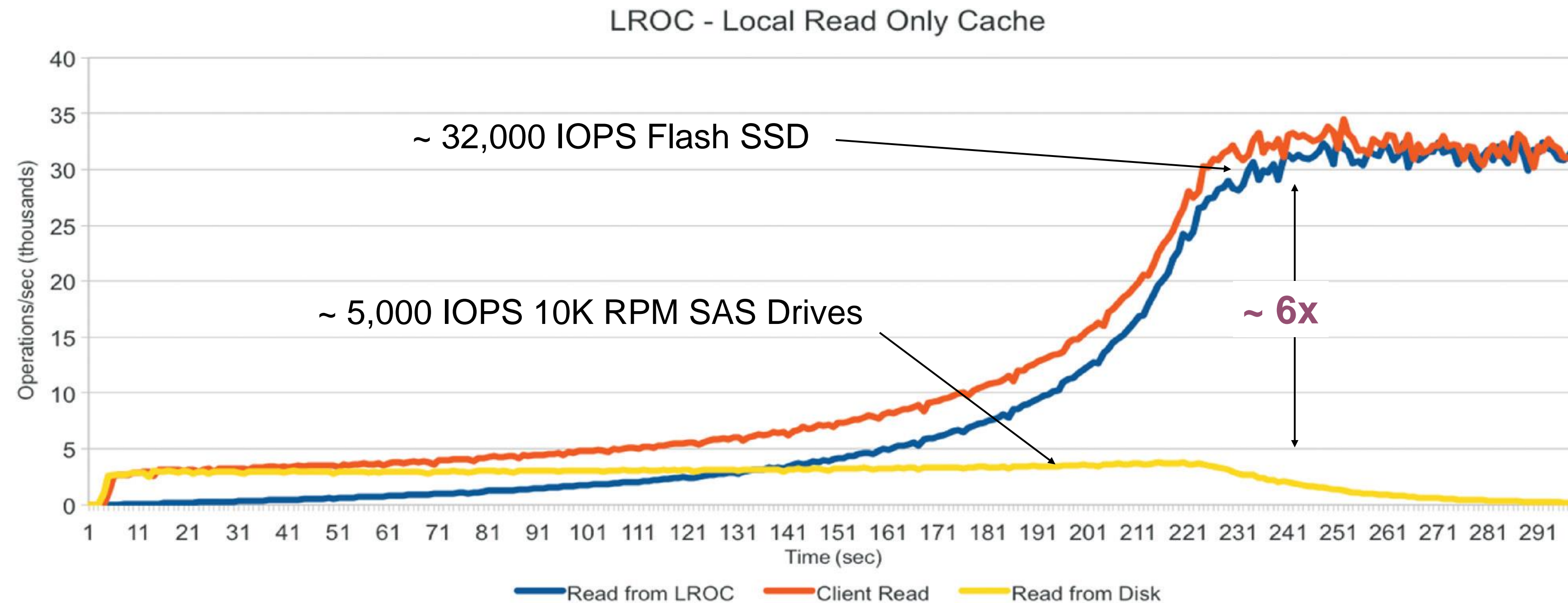
Flash LROC NVMe (복수 개 구성 가능)





FS Cache 확장 > Flash Cache 성능 향상 Test 결과

- Two consumer grade 200 GB SSDs cache a forty-eight 300 GB 10K SAS disk Spectrum Scale storage system



- Initially, with all data coming from the disk storage system, the client reads data from the SAS disks at ~ 5,000 IOPS
- As more data is cached in Flash, client performance increases to 32,000 IOPS while reducing the load on the disk subsystem by more than 95%

IBM
ESS3500

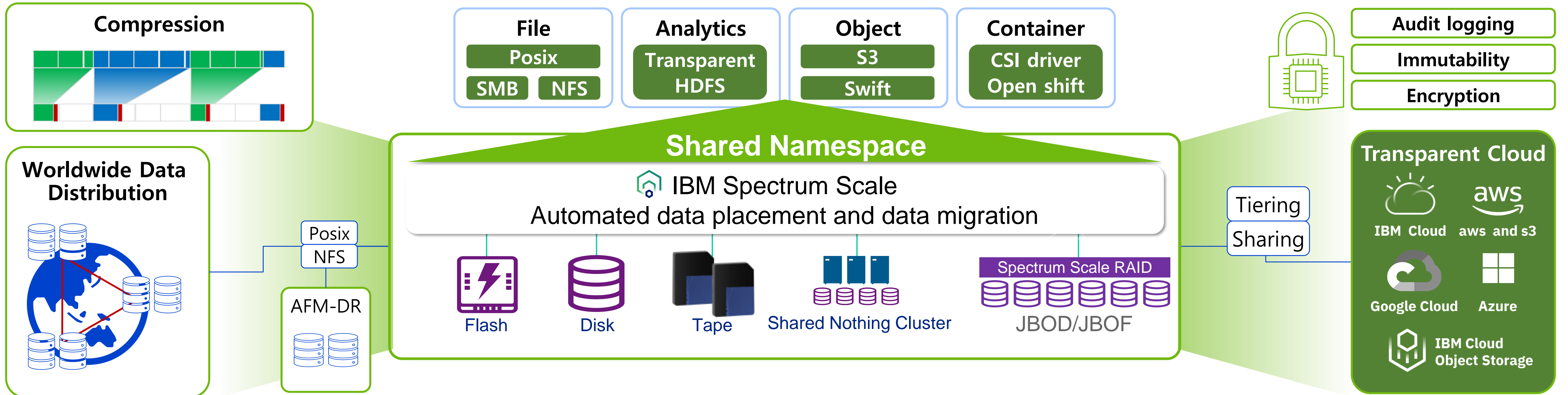
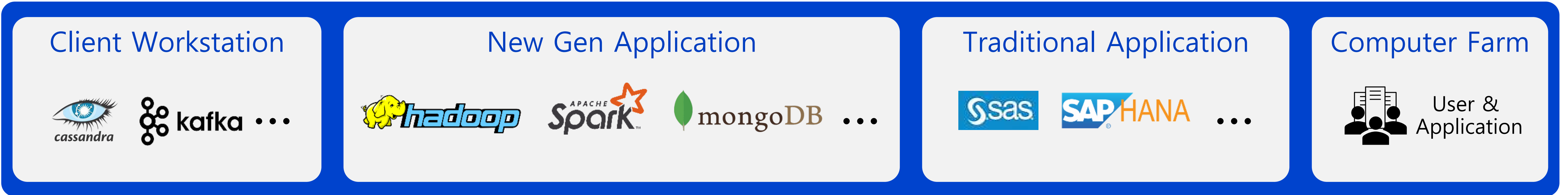
Function





Spectrum Scale 기능 요약

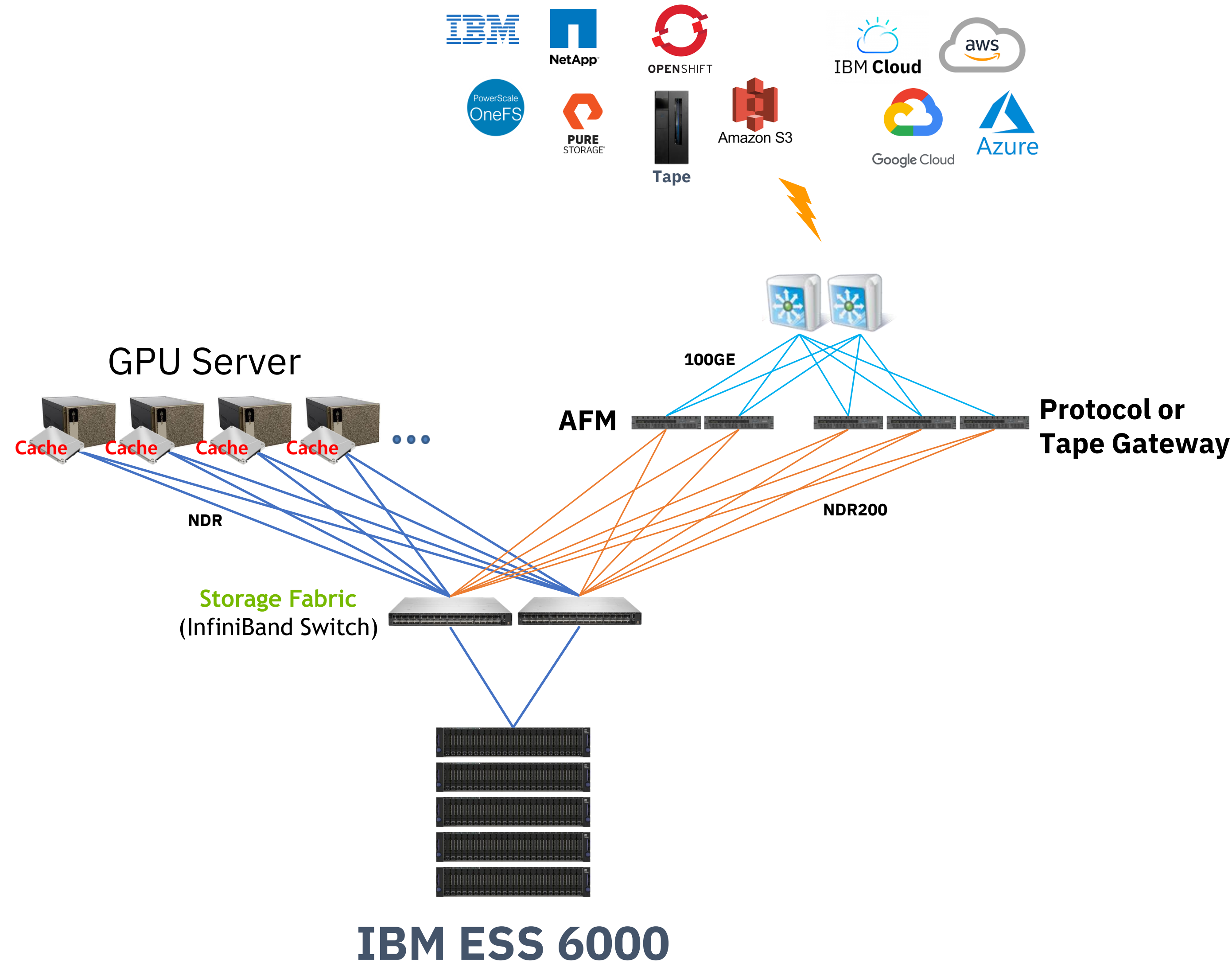
IBM ESS3500 스토리지는 다양한 New 어플리케이션 요구를 최적으로 제공하는 통합 파일 시스템 솔루션이며, 단일 파일 시스템에서 **File (POSIX, NFS, SMB), Object (S3, Swift), HDFS, Container**를 모두 지원하는 소프트웨어 정의 스토리지 솔루션 입니다.



**IBM
ESS3500**

제안 구성도





워크로드	사용 사례
 AI/ML/HPC	NVIDIA, TensorFlow, Caffe, Pytorch, IBM Cloud Pak® for Data, ML Kit, AWS Lambda, AWS SageMaker ...
 분석	Cloudera, Hadoop, Apache Spark™, SAS®, Tableau, Python, Power BI, ...
 협업	미디어 개발, AI 모델 개발, 컴퓨터 지원 엔지니어링
 크리티컬 앱	데이터 요구사항이 증가하고 있는 비즈니스 크리티컬 애플리케이션
 하이브리드 클라우드	Red Hat® OpenShift® 워크로드, 백업 및 아카이브, 데이터를 이동하거나 온프레미스 및 클라우드에 액세스하는 애플리케이션
 통합	데이터 레이크, 데이터 오션, 빅데이터

IBM