

효율적인 데이터 관리와
보호대비를 위한 솔루션 세미나



한국 IBM
전상호 실장

비정형 데이터 관리의 최적 솔루션



IBM Spectrum **NAS**

IBM Software Defined Storage 전제품 포트폴리오

IBM Spectrum Storage

어플라이언스, 컨버전스 및 클라우드에서 제공되는 IBM 소프트웨어 정의 스토리지 제품군



스토리지
관리



스토리지
서비스



Scale-Out
스토리지



IBM 비정형 Data Container 솔루션 비교

구분	IBM Spectrum Scale	IBM Spectrum NAS	IBM Cloud Object Storage
워크로드	HPC (슈퍼컴퓨팅, AI) Big Data Analytics	VM환경에서의 파일 저장소	클라우드, Native Object Application
	초 고성능 파일 공유	Home 디렉토리 일반적 목적의 파일 저장소	컨텐츠 저장소 및 배포
	고성능 backup & restore	Microsoft 어플리케이션	백업 & 아카이브
특장점	High Performance	NFS, CIFS를 통한 범용적 파일 저장소	Multi-site Native Object Storage

IBM Storage Solution의 하드웨어 정책

IBM Spectrum Storage



Ready for IBM Storage

- 유연한 성능 및 용량 확장성
- TCO 최적화
- 신뢰성, 정합성, 안전성 극대화
- Data 보호 및 가용성 극대화
- 운영 효율 및 관리 편의성
- 구형 Architecture의 문제점 개선

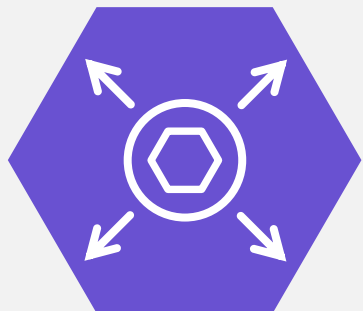


- Spectrum NAS 개요
- Object based File Storage
- 유연한 성능 및 용량 확장성
- TCO 최적화
- 신뢰성, 정합성, 안전성 극대화
- Data 보호 및 가용성 극대화
- 운영 효율 및 관리 편의성



IBM Spectrum NAS

엔터프라이즈 기업을 위한 Scale-out NAS 솔루션



단순한 설치 및 관리

- 물리서버 or VM에 간편한 설치
- 최상의 가용성
- self-balancing



Scale-out 파일 Storage

- 선형적인 성능 및 용량 증가
- 무 중단 증설
- 최대 노드 수 제한 없음



Cost 최적화

- 표준 X86 서버 및 IP Network
- 효율성, 성능, 데이터 보호 극대화 등 고객 요구 환경에 최적화



Enterprise NAS 기능 제공

- Native NFS, SMB 프로토콜
- 다양한 NAS 기능 제공
- 동기/비동기 복제

Enterprise NAS feature

IBM Spectrum NAS

Spectrum NAS 개요

Protocol Support



SMB 1 / 2.1 / 3.1.1

NFS 3 / 4.0 / 4.1

cross
protocol
access

DR



Authentication



Active Directory
LDAP
Kerberos KDC
NIS
Local database

File Policies



Tier



Snapshots



Quota



Encryption

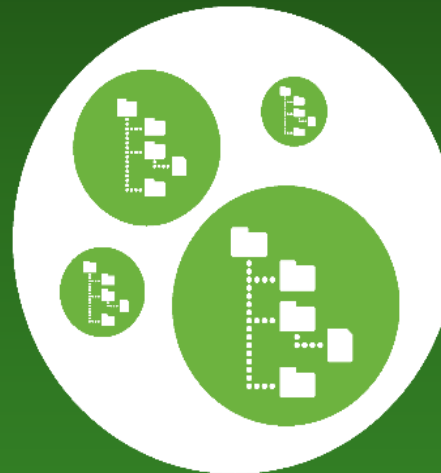


WORM



Retention

Multitenancy



Authorization



ACL
mode_t
SID
UID/GID

Antivirus



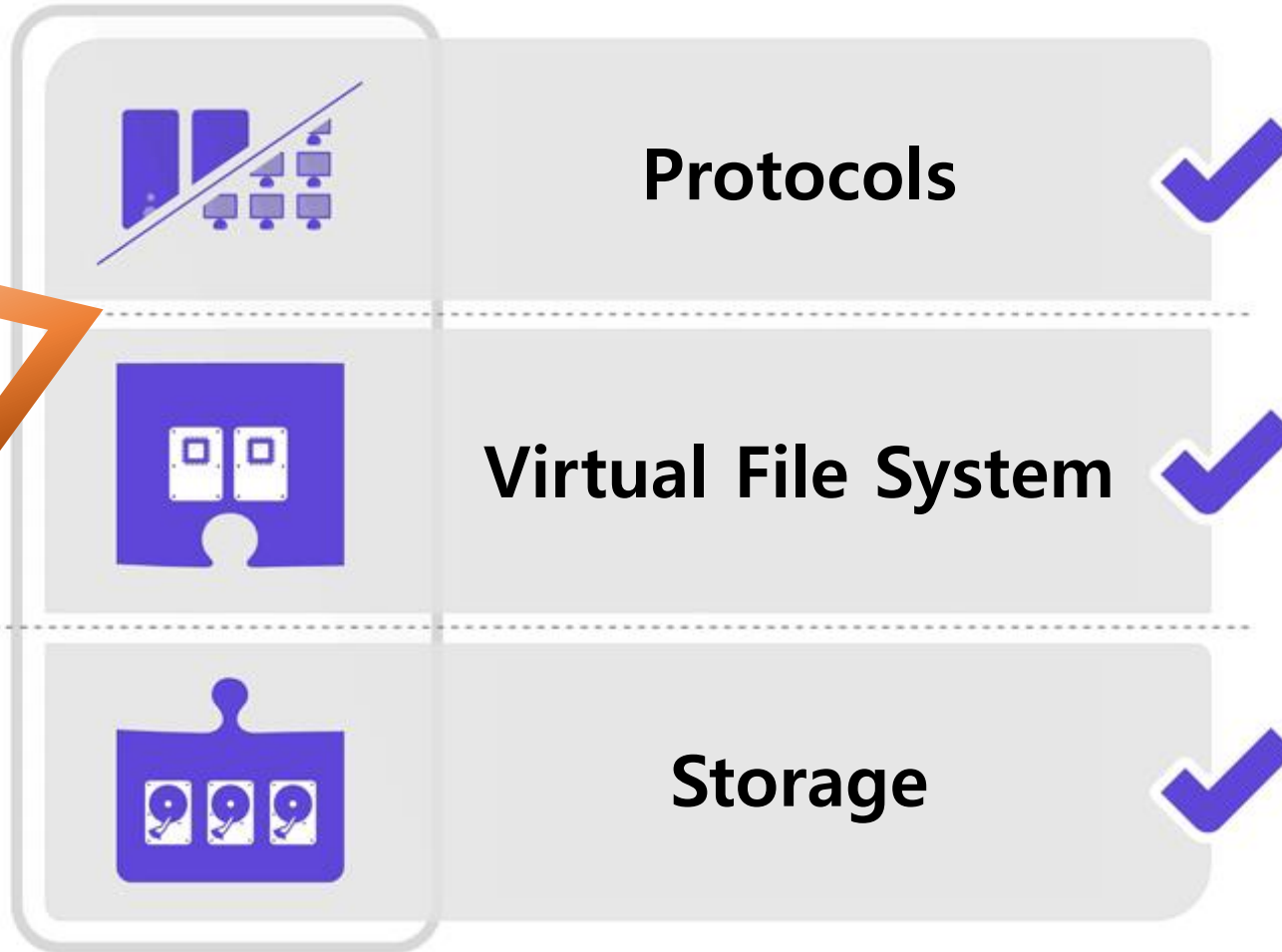
Integration

모든 Software Stack을 직접 제공

IBM Spectrum NAS

Spectrum NAS 개요

Latency
감소



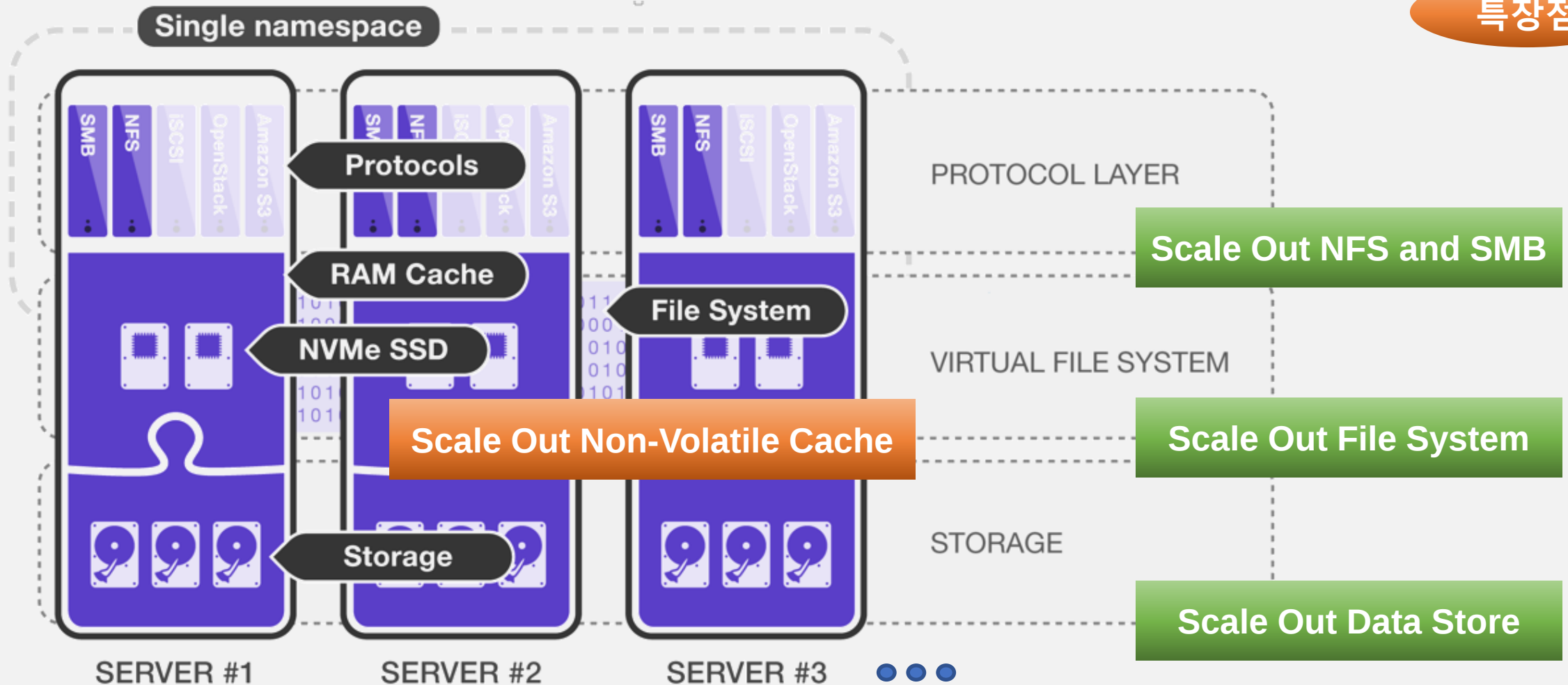
특장점

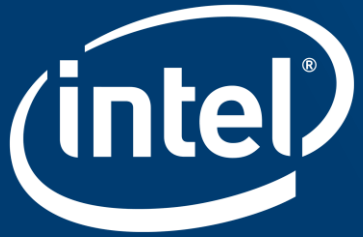
**NO
OPEN
SOURCE**

**NO
3rd Party
SOFTWARE**

Scale-out Storage (Node 증설형)

특장점





**Best Caching solution
for IBM Spectrum NAS,
Intel® Optane™ SSD**

Intel Innovations – Transforming the Data Tiering Hierarchy

Performance and Capacity for Every Need – Filling the Gaps!

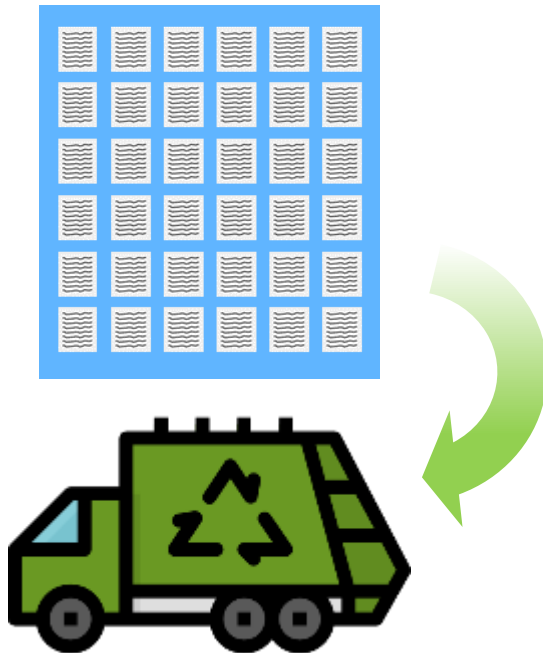


What is Intel® Optane™ DC SSD?

It's DIFFERENT...

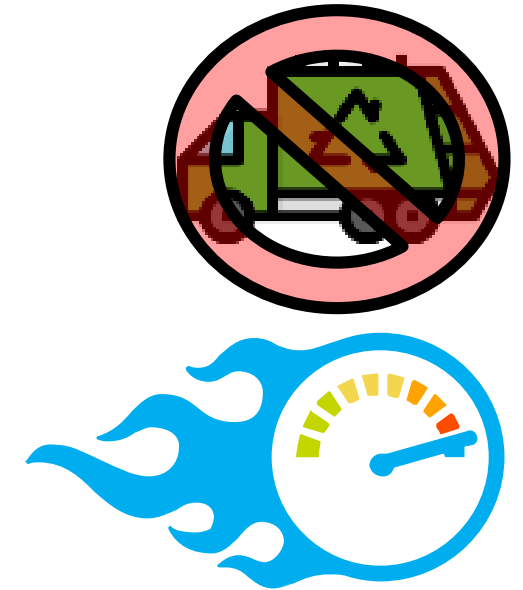
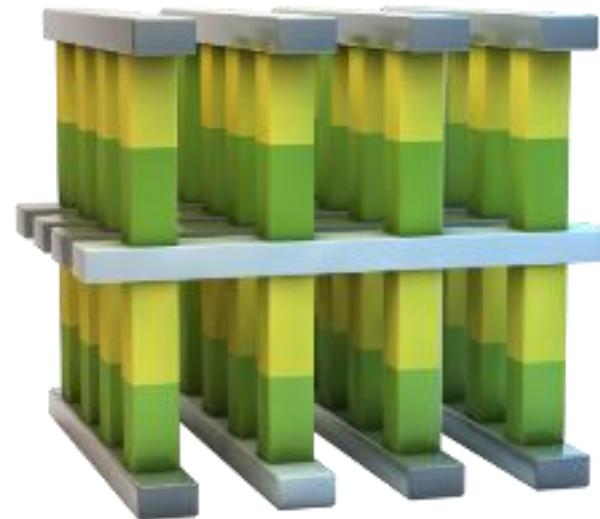
NAND

Read and write in PAGES
ERASE only in BLOCKS



Intel® Optane™ DC SSD

Media is addressable on a fine grain, so
No garbage to collect!



Permission to use garbage truck icon made by [monkik](#) from [www.flaticon.com](#) with reference

Intel® Optane™ SSD DC P4800X changes the storage and memory game

Breakthrough Performance

up to
6X
faster

Queue Depth 1
4K 70/30 RW IOPS¹

Predictably Fast Service

up to
60X
better

99% QoS²

Responsive Under Load

up to
63X
faster

Response Time³

high Endurance

up to
20X
more

Total Terabytes Written⁴

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks. **1.** Intel-tested: 4K 70/30 RW Performance at Low Queue Depth. Test and System Configuration: CPU: Intel® Xeon® Gold 6140 FC-LGA14B 2.3GHz 24.75MB 140W 18 cores CD8067303405200, CPU Sockets: 2, RAM Capacity: 32G, RAM Model: DDR4, RAM Stufng: NA, DIMM Slots Populated: 2 slots, PCIe* Attach: CPU (not PCH lane attach), Chipset: Intel C620 chipset BIOS: SE5C620.8 6B.00.01.0013.030920180427, Switch/ReTimer Model/Vendor: Cable - Oculink 800mm straight SFF-8611 to right angle SFF-8611 Intel AXXCBL800CVCR, OS: CentOS 7.5, Kernel: 4.14.50(LTS), FIO version: 3.5; NVMe* Driver: Inbox, C-states: Disabled, Hyper Threading: Disabled, CPU Governor (through OS): Performance Mode, EIST (Speed Step), Intel Turbo Mode=Disabled, and P-states = Enabled. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of August 7, 2018, and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. **2.** Source – Intel-tested: QoS: measures 99 percent QoS under 4K 70–30 workload at QD1 using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using FIO 3.1. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. SSDs tested were commercially available at time of test. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

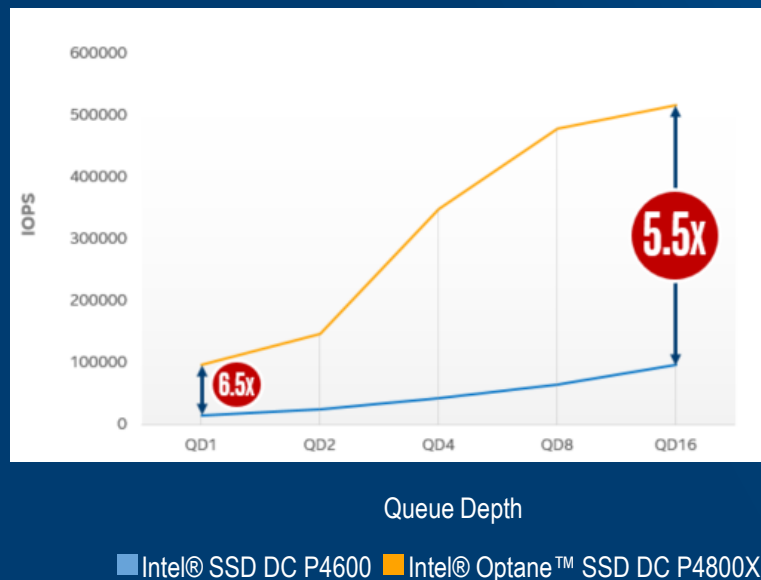
3. Source – Intel-tested: See configuration and notices in 2. above. Response Time refers to average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1.

4. Source – Intel: Endurance ratings available at <https://www.intel.com/content/www/us/en/solid-state-drives/optane-ssd-dc-p4800x-brief.html>

Intel® Optane™ SSDs Deliver Breakthrough storage Performance

Breakthrough Performance

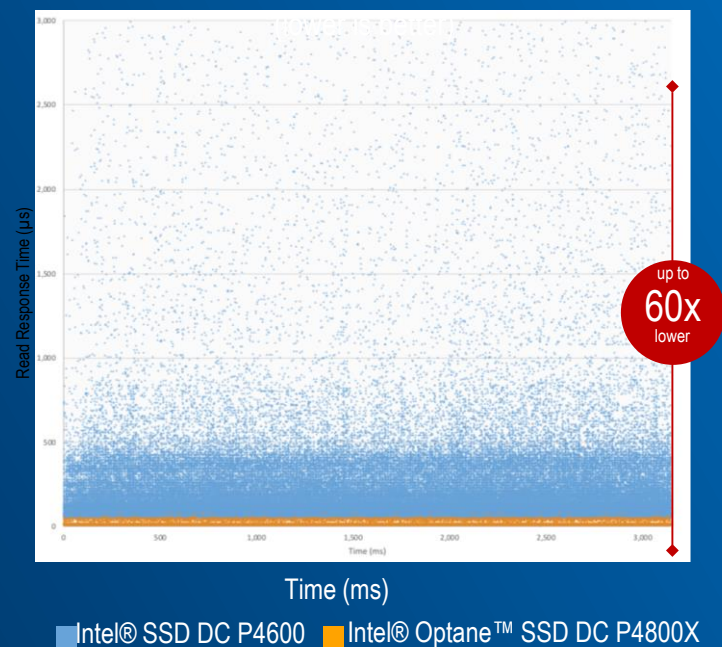
4K 70/30 RW Performance at Low Queue Depth¹
(higher is better)



5-6x faster at low QD¹

Predictably Fast Service

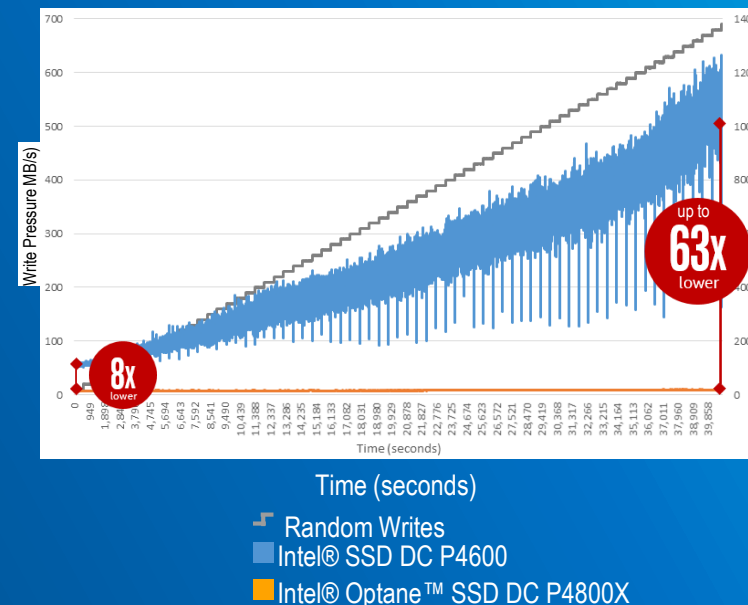
Read QoS in Mixed Workload^{1,2}



up to 60x better 99% QoS²

Responsive Under Load

Average Read Latency under Random Write Workload^{1,3}
(lower is better)



up to 63x faster under load³

¹ Source – Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using FIO 3.1. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. SSDs tested were commercially available at time of test. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

² Source – Intel-tested: See configuration and notices in 1. above. QoS: measures 99 percent QoS under 4K 70–30 workload at QD1 using FIO 3.1.

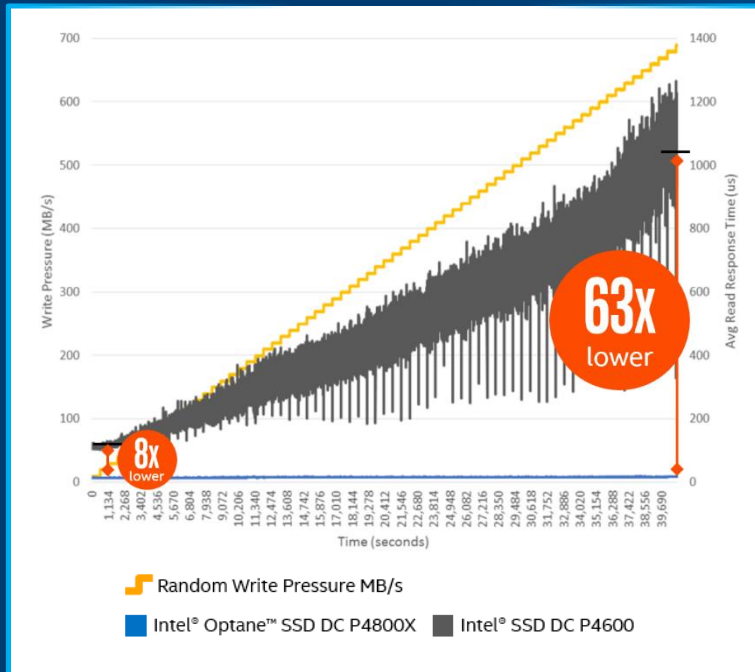
³ Source – Intel-tested: See configuration and notices in 1. above. Response Time refers to average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1.

Caching: Intel® Optane™ SSD DC P4800X. the ideal caching solution.



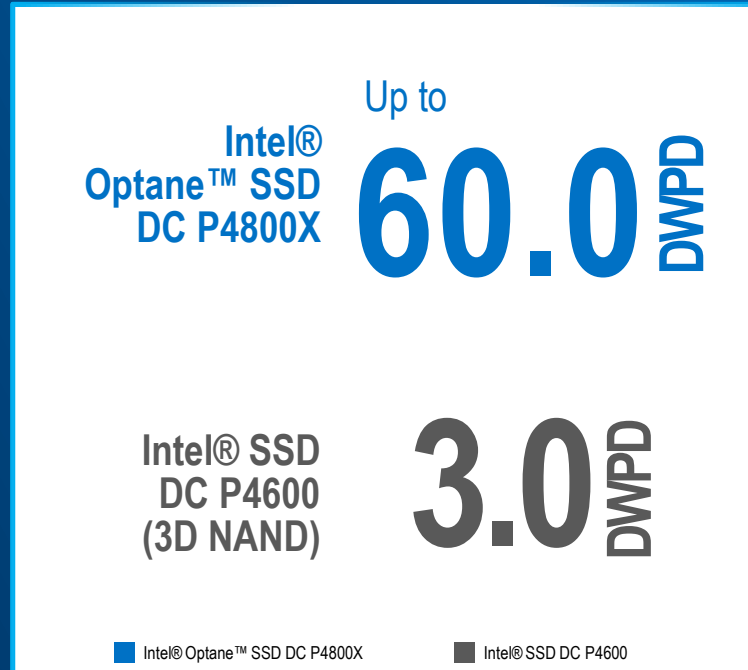
lower & more consistent latency

Average Read Latency under Random Write Workload¹



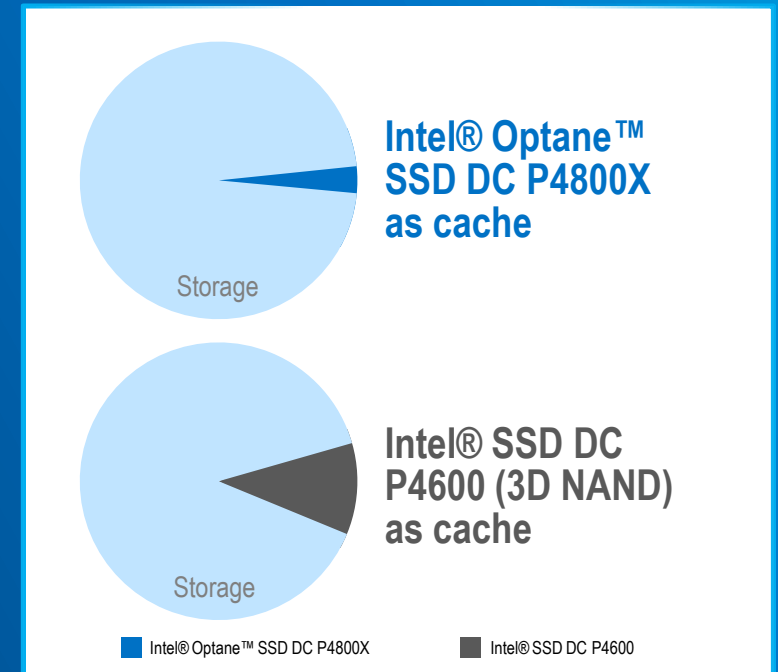
higher endurance

Drive Writes Per Day (DWPD)²



more efficient

Cache as a % of Storage Capacity³



Low latency + high endurance = **greater SDS system efficiency**

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

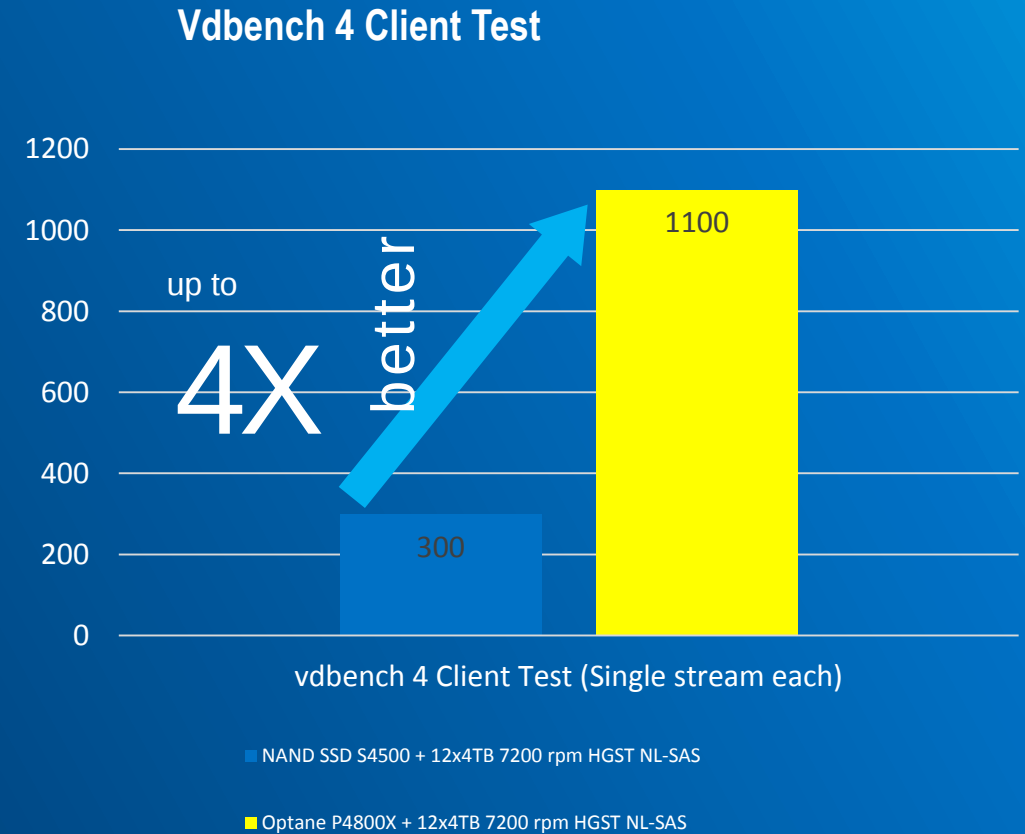
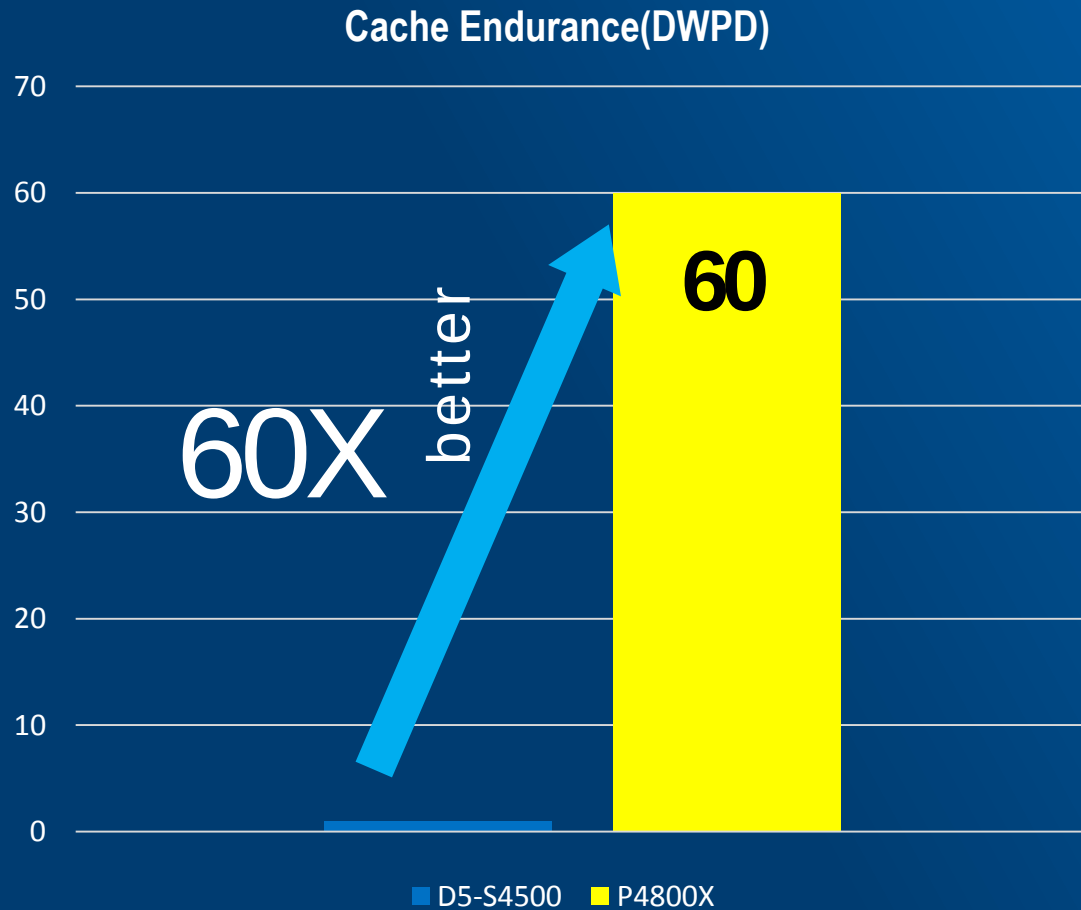
1. Source – Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using FIO 3.1. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. SSDs tested were commercially available at time of test. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

2. Source – Intel: Endurance ratings available at <https://www.intel.com/content/www/us/en/solid-state-drives/optane-ssd-dc-p4800x-brief.html>

3. Source – Intel: General proportions shown for illustrative purposes.



A Clear Advantage by adding Intel® Optane™ Cache for IBM spectrum NAS



Tests by DSNG. SuperMicro SuperServer x3650 M5, model 8871-AC1/CPU : 2 * Intel Xeon Silver 4114 (13.75 MB cache @ 2.2GHz, 10 cores, 2400MHz Memory Clock * 6 channel, max 48 PCIe lane)/https://ark.intel.com/products/123550/Intel-Xeon-Silver-4114-Processor-13_75M-Cache-2_20-GHz/Memory: 96GB (Configured 2400 MHz, Physically 2666MHz) 8GB * 12 channel for maximum channel configuration/Network: 10GbE, Mellanox Dual Port SFP+ ConnectX-3 Pro based adapter (two NIC with two ports each)/Drives: 12x 4TB 7200 rpm HGST NL-SAS; 1x Intel Optane P4800X 375GB PCIe NVMe / 1x Intel SSD D3-S4510 /80GB/<https://www.intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds/optane-dc-p4800x-series.html>/2+1 Erasure Code (1 fault resiliency) : 67 % utilization (33 % parity)/1-cache Replication @ Cluster config/Read/Write Cache with RAM Cache Enable @ Node config / Gateway/Disable disk write cache @ Node config / Storage/NFSv4 test only/CentOS 7.3 /Custom Build Kernel 4.9.76

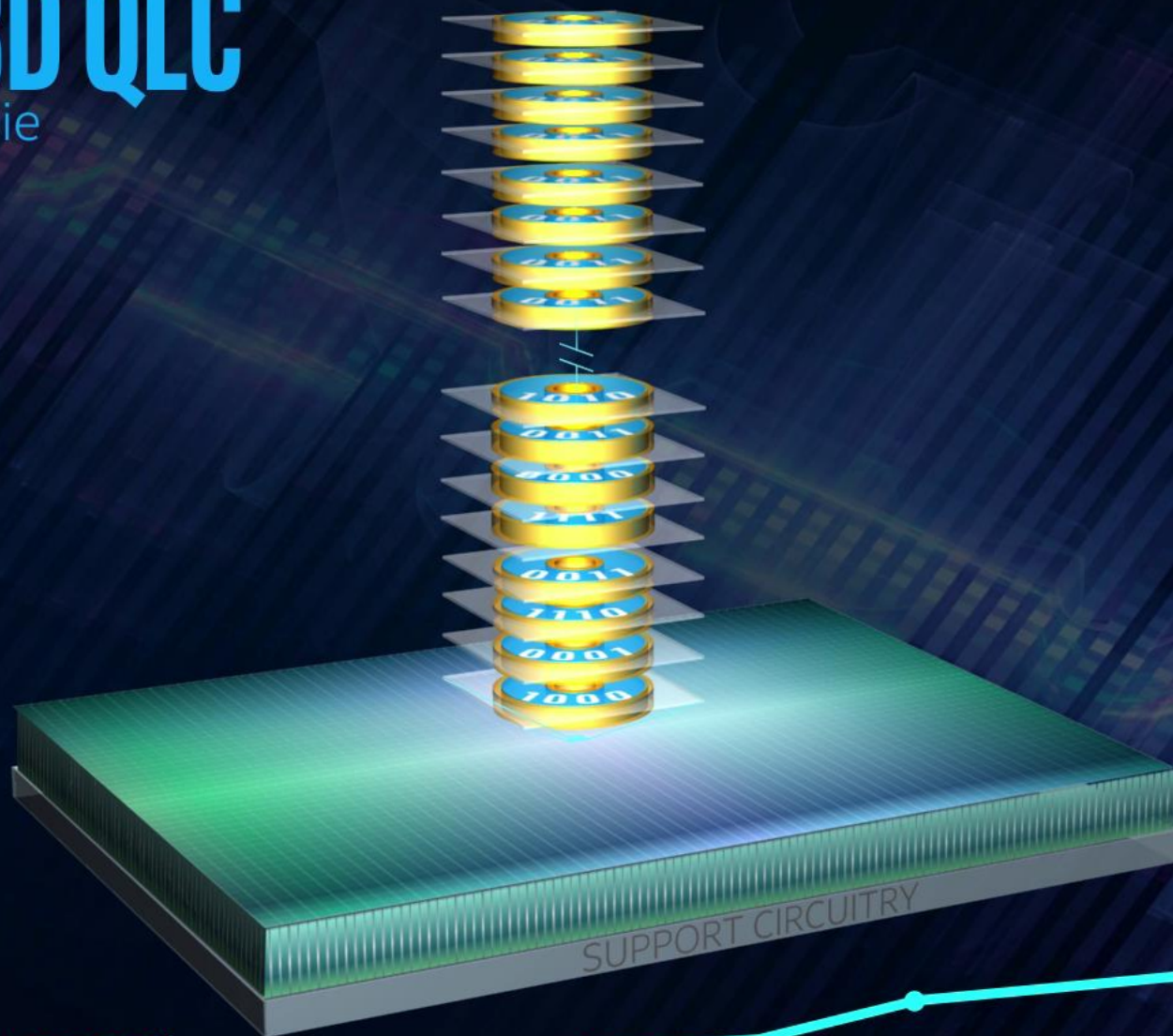
Best Caching solution for IBM Spectrum NAS, Intel® Optane™ Technology



2018 3D QLC

1024 Gb /Die

64 LAYERS



256X
INCREASE IN AREAL DENSITY

2D SLC
4 Gb

AREAL DENSITY

2D MLC
128 Gb

3D TLC
384 Gb

3D QLC
1024 Gb

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

intel® ssd D5-P4320, D5-P4326 series store more. Save more. Do more.

U.2 15mm



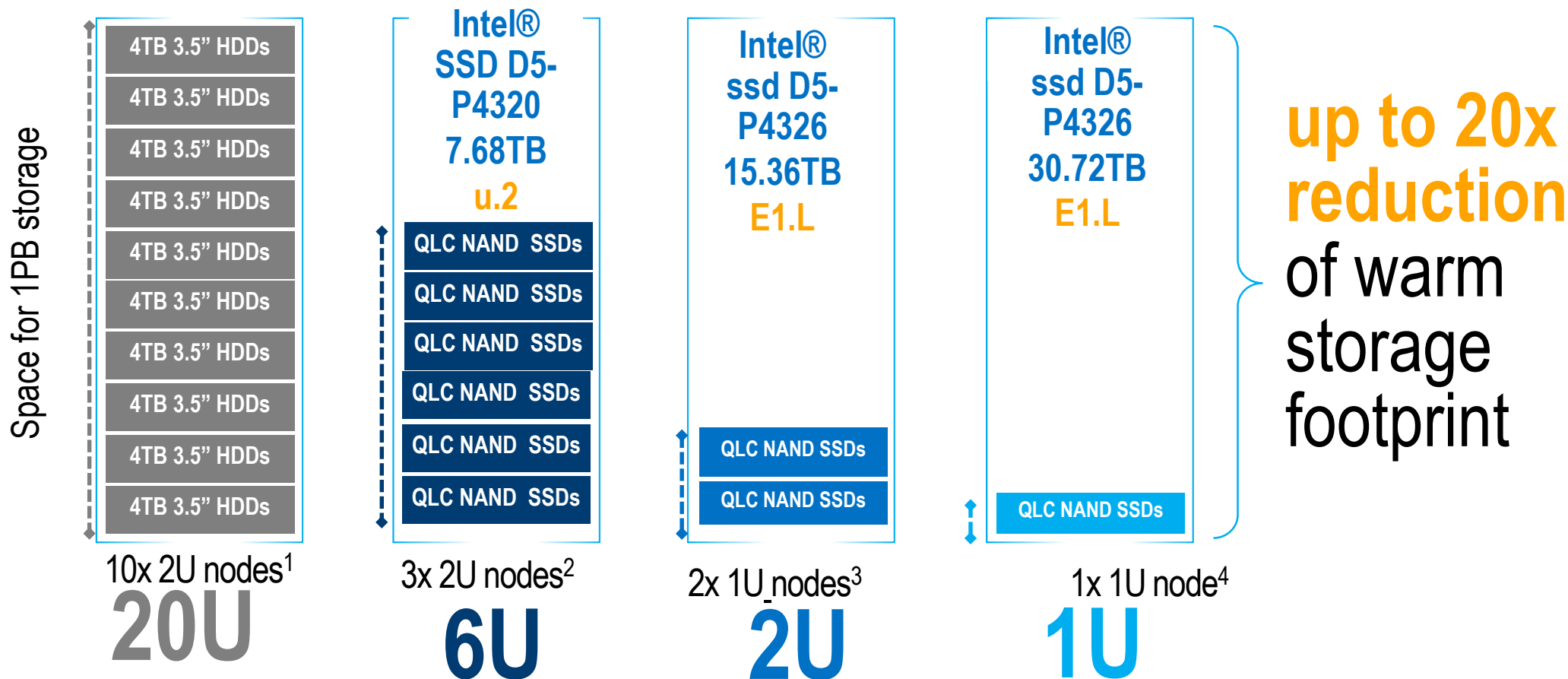
E1.L



The industry's 1st QLC PCIe* SSD¹, the Intel® SSD D5-P4320 and D5-P4326 Series **delivers big, affordable PCIe drives to warm storage**. As data centers look to cost-effectively unleash the value of stored data, these read-optimized SSDs enable users to **store more, save more and do more** than legacy solutions. Packing 33% more bits/cell² than the prior generation the drives **consolidate storage footprints up to 20x³**. Consolidation reduces operations expenditures, helping to **save cost vs. legacy storage⁴**. High IOPS/TB enabled by PCIe ensure you get the most out of storage as capacities scale.

* Other names and brands may be claimed as property of others.
See appendix for footnotes.

Massively consolidate storage footprint

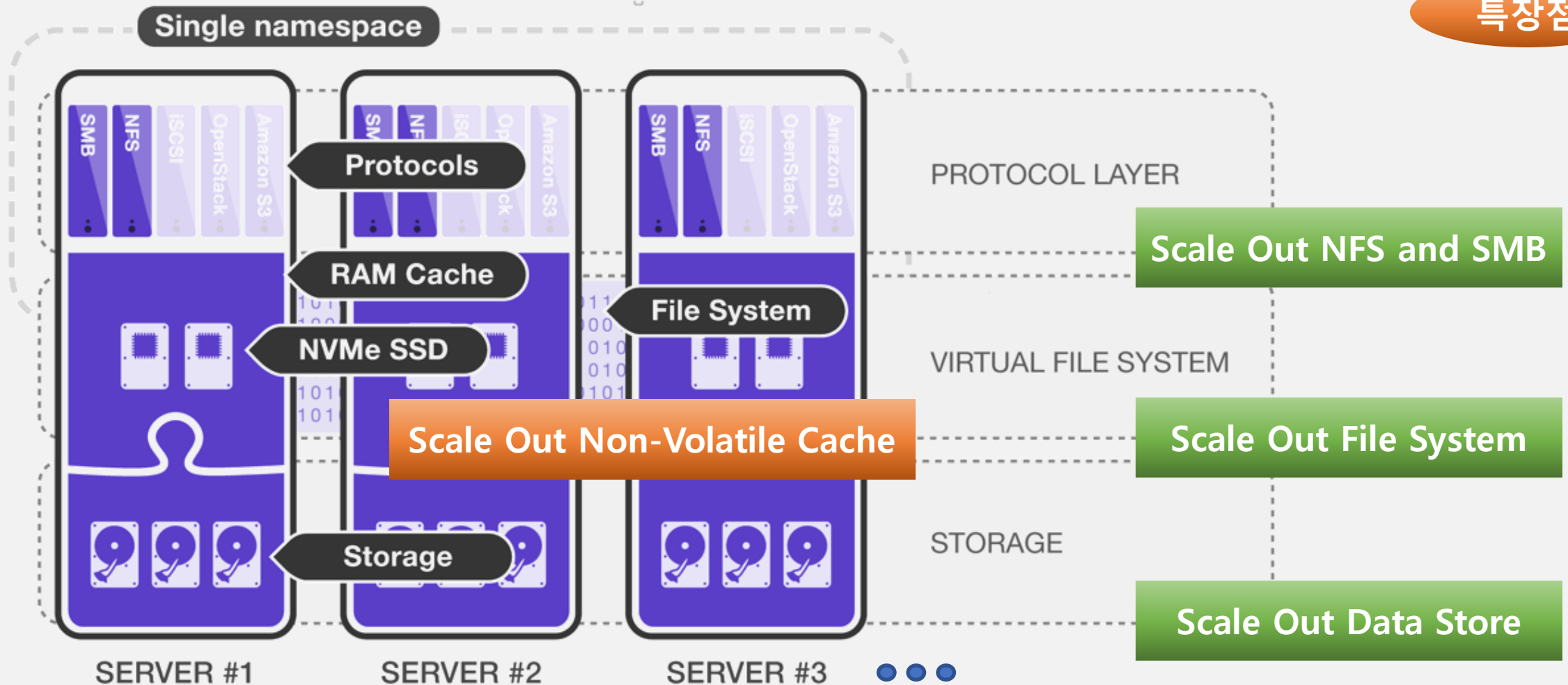


*Other names and brands may be claimed as property of others.

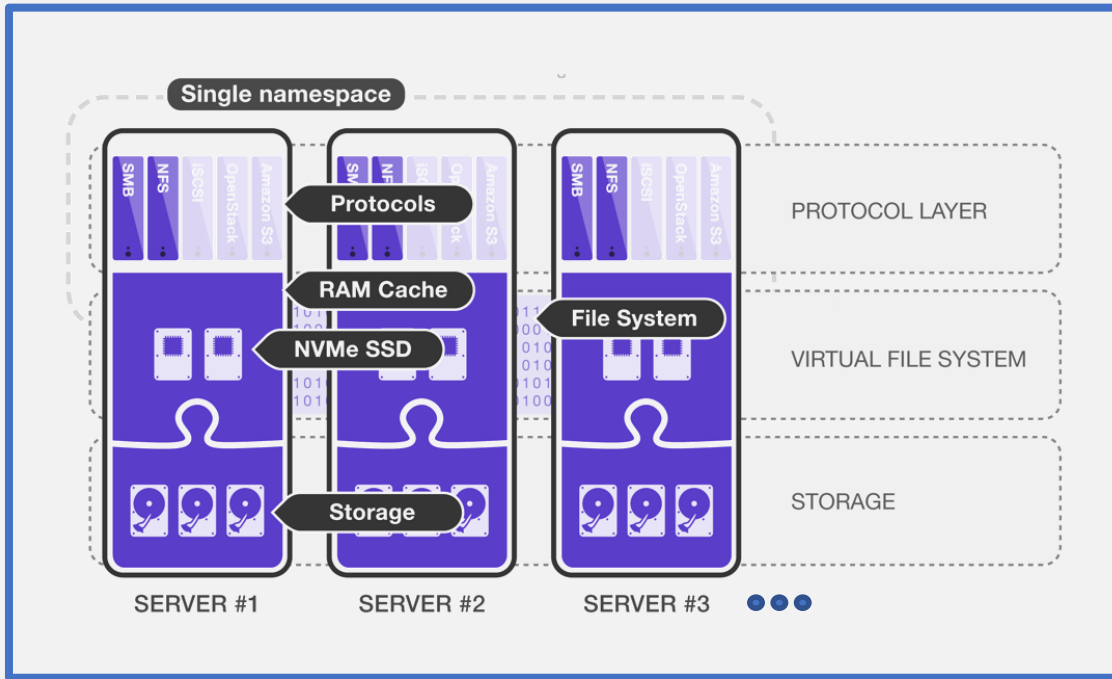
1. 4TB 3.5" HDDs - 10 2U nodes per 1PB - 960TB total based on 24 3.5" HDDs per 2U. Note that 4TB HDDs are used in this comparison as we are targeting "warm" storage meaning that a certain capacity is needed, but performance is also important and 4TB HDDs perform much better than larger capacity HDDs. Based on 4TB 3.5" HDD - WD Gold TB Enterprise class 7200 RPM- https://www.newegg.com/Product/Product.aspx?Item=N82E16822235059&nm_mc=KNC-MSNSearch-PC&cm_mmc=KNC-MSNSearch-PC--pla--Hard+Drives--Western+Digital--22235059&msclkid=db39c4b23332181f75ba9f9095adacd2&gclid=CPCN99fX09sCFRiFXQld5DQAeA&gclid=ds
2. 8TB 2.5" U.2 SSDs - 3 2U nodes per 1PB - 1,106TB total using 144 7.68TB SSDs; 48 2.5" SSDs per 2U node using 2.5" U.2 from factor; 3 2U nodes for 6U total. Based on 7.68TB Intel® D5-P4320 QLC SSD
3. 16TB E1.L SSDs - 2 1U nodes per 1PB - 983TB total using 64 15.36TB SSDs; 32 SSDs per 1U node using E1.L form factor; 2 1U nodes for a total of 2U. Based on 15.36TB Intel® D5-P4326 QLC SSD available at a future date
4. 32TB E1.L SSDs - 1 1U node per 1PB - 983TB total using 32 30.72TB SSDs; 32 SSDs per 1U node using E1.L form factor; 1 1U node for a total of 1U. Based on 30.72TB Intel® D5-P4326 QLC SSD available at a future date

Scale-out Storage (Node 증설형)

특장점



Scale-out Storage (Node 증설형)



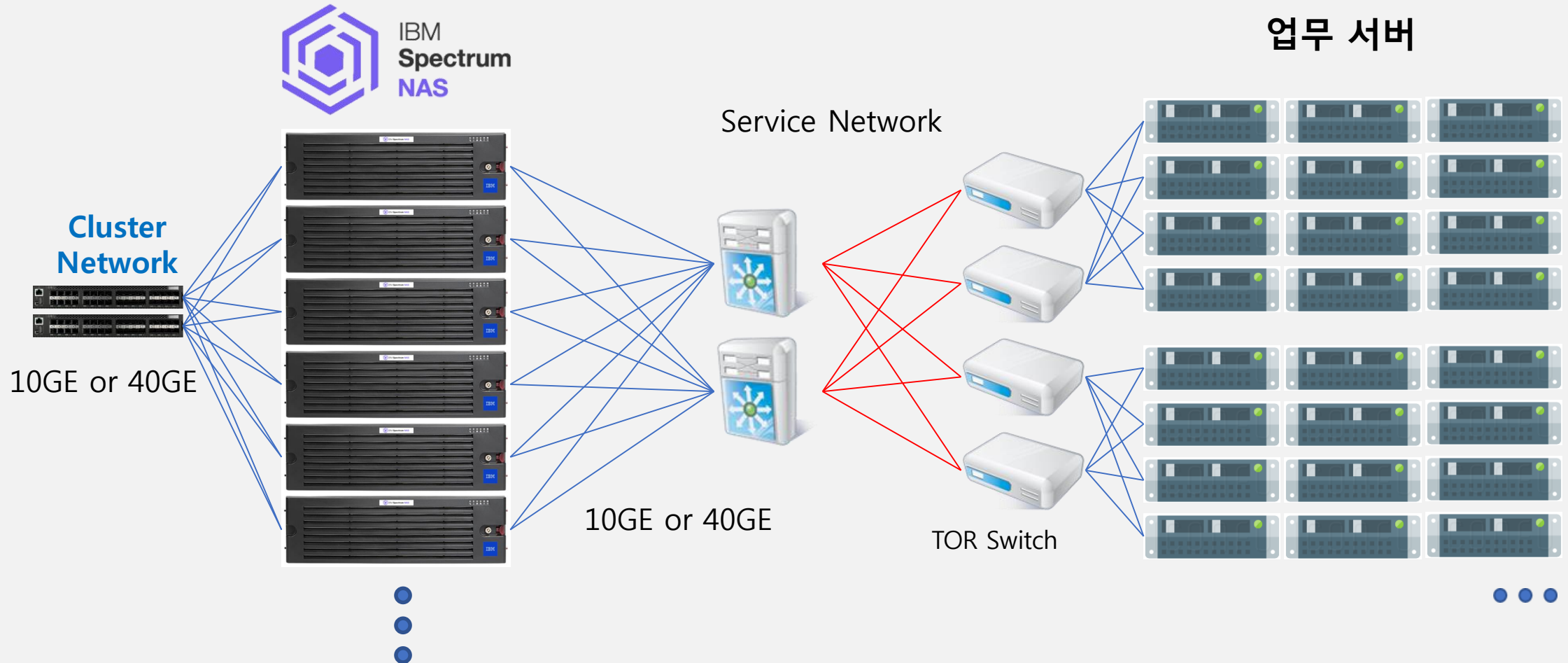
True Scale-Out

- Symmetric Architecture
 - ✓ 각 노드는 동일한 역할이 배정됨
- 핫 스팟 No ! Bottlenecks No!
- 파일은 모든 노드의 캐시에서 Read/Write 됨
- 최소 4 node 부터 시작

- 데이터 저장소가 Erasure Coding을 실행하고 Load Balancing을 담당합니다.
- Self-Healing Cluster → 노드 장애시에도 데이터의 가용성을 보장합니다.
- 수백 대 이상의 노드로 증설 가능합니다.

Spectrum NAS의 일반적 구성도

IBM Spectrum NAS
Spectrum NAS 개요



Spectrum NAS Offering

구분	All Flash	High IOPS	High Capacity & Throughput	Data Archive	I/O Booster
모델명	Flash	SAS	NLSAS	Archive	Booster
용도	- 초고성능 - Random Access	- Workload : High IOPS가 필요한 업무 - 다수의 Small File - 저용량 But 높은 Throughput	- Workload : High Throughput이 필요한 업무 - Sequential Access - 다수의 Large File	대용량 Archive	- Access Booster - Backup Booster
CPU	Xeon Gold 6126 (2ea)	Xeon Gold 6126 1ea (2ea optional)	Xeon Silver 4114 1ea (2ea optional)	Xeon Silver 4110 (1ea)	Xeon Gold 6126 (2ea)
Memory	192GB (384GB optional)	192GB (384GB optional)	96GB (192GB optional)	48GB	192GB (384GB optional)
2차 Cache	NVMe 2ea (4ea optional)	NVMe 1ea (2ea optional)	NVMe 1ea (2ea optional)	SSD (1ea)	NVMe 1ea (2, 4ea optional)
Network	40GE	10GE (40GE)	10GE	10GE	40GE (10GE)

Spectrum NAS Offering

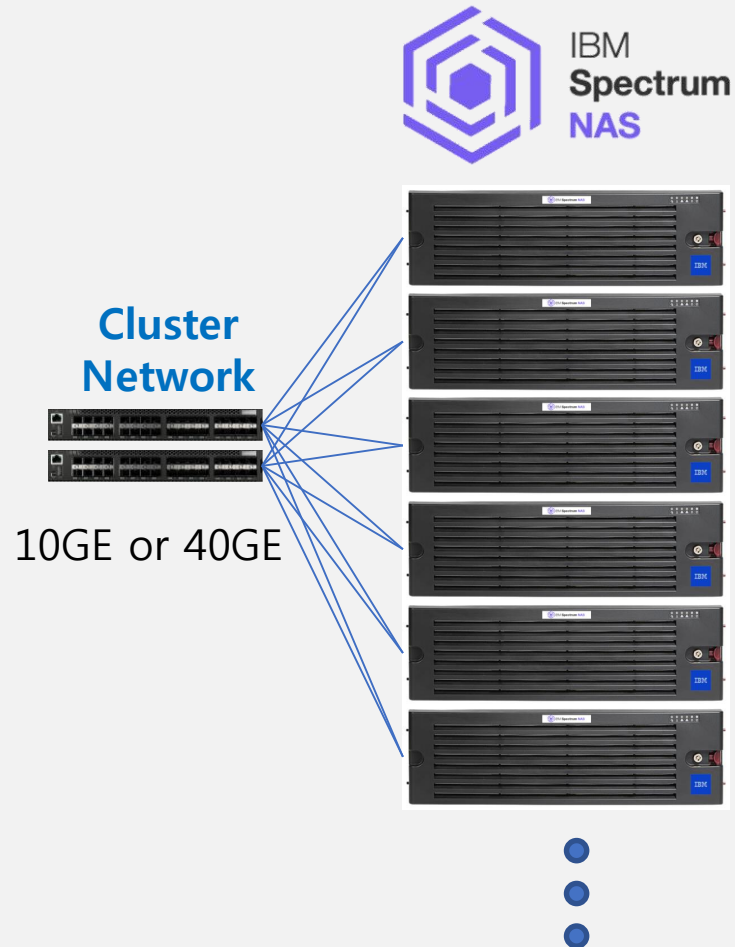
구분	All Flash	High IOPS	High Capacity & Throughput	Data Archive	I/O Booster
모델명	Flash	SAS	NLSAS	Archive	Booster
Disk	NVMe / SSD 전모델	2.5" 10Krpm SAS Disk 전 모델	3.5" 7.2Krpm NL-SAS Disk 전 모델	3.5" 7.2Krpm NL-SAS Disk 전 모델	-
Disk Slot	24 bay	24 bay	12, 16, 24 bay	36, 45, 90 bay	-
Node 수	최소 4 node ~ 최대 제한없음. (1 node 단위 증설)				1 node 단위 증설
지원 Protocol	SMB (1 / 2.0 / 2.1 / 3.11), NFS (3 / 4.0 / 4.1)				
File Coding	- Copies : 3 or 5 copies - Erasure coding : 2+1 2+2 3+1 3+2 4+1 4+2 5+1 5+2 6+1 6+2 8+1 8+2				

각 Model은 고객 I/O의 특성에 따라 최적화 구성으로 Modify 가능합니다.

Spectrum NAS의 하드웨어 외형

IBM Spectrum NAS

Spectrum NAS 개요



NS-SAS
16Bay 모델

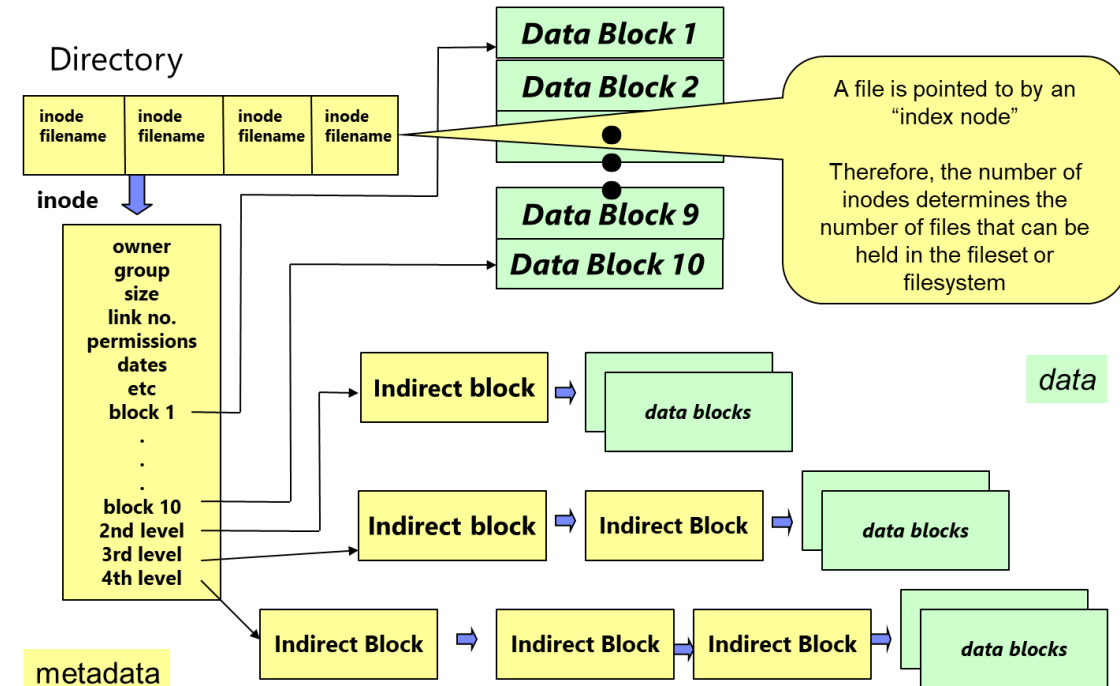
일반적인 파일시스템의 Meta data 저장 구조

일반적인 File System의
Meta Data 저장 Architecture

Database의 Table 구조와 유사
(다수의 Field로 구성된 Big Table)

파일 개수가 많아진다
= Table의 Record 수가 많아진다
= Record 조회성능이 떨어진다
= Meta data Access 성능이 떨어진다
Table 크기에 제한이 있다
= 최대 파일 개수에 제한이 있다

The inode in a UNIX/Linux file system



Spectrum NAS의 Meta data 저장 구조

일반적인 File System의
Meta Data 저장 Architecture



Database의 Table 구조와 유사
(다수의 Field로 구성된 Big Table)



파일 개수가 많아진다
= Table의 Record 수가 많아진다
= Record 조회성능이 떨어진다
= Meta data Access 성능이 떨어진다
Table 크기에 제한이 있다
= 최대 파일 개수에 제한이 있다

Spectrum NAS의 File System
Meta Data 저장 Architecture



Object Storage 구조와 유사
(Meta Data를 Object로 저장)



- 파일 개수가 많아지는 것과 Meta Data조회 성능은 직접적인 연관 없음.
- Meta Data 영역을 별도로 확보하지 않으므로 그 크기에 제한이 없음 → 파일시스템 내의 최대 파일 개수 제한 없음

노드 추가에 따라 용량 및 성능 확장

IBM Spectrum NAS
Object Based File Storage

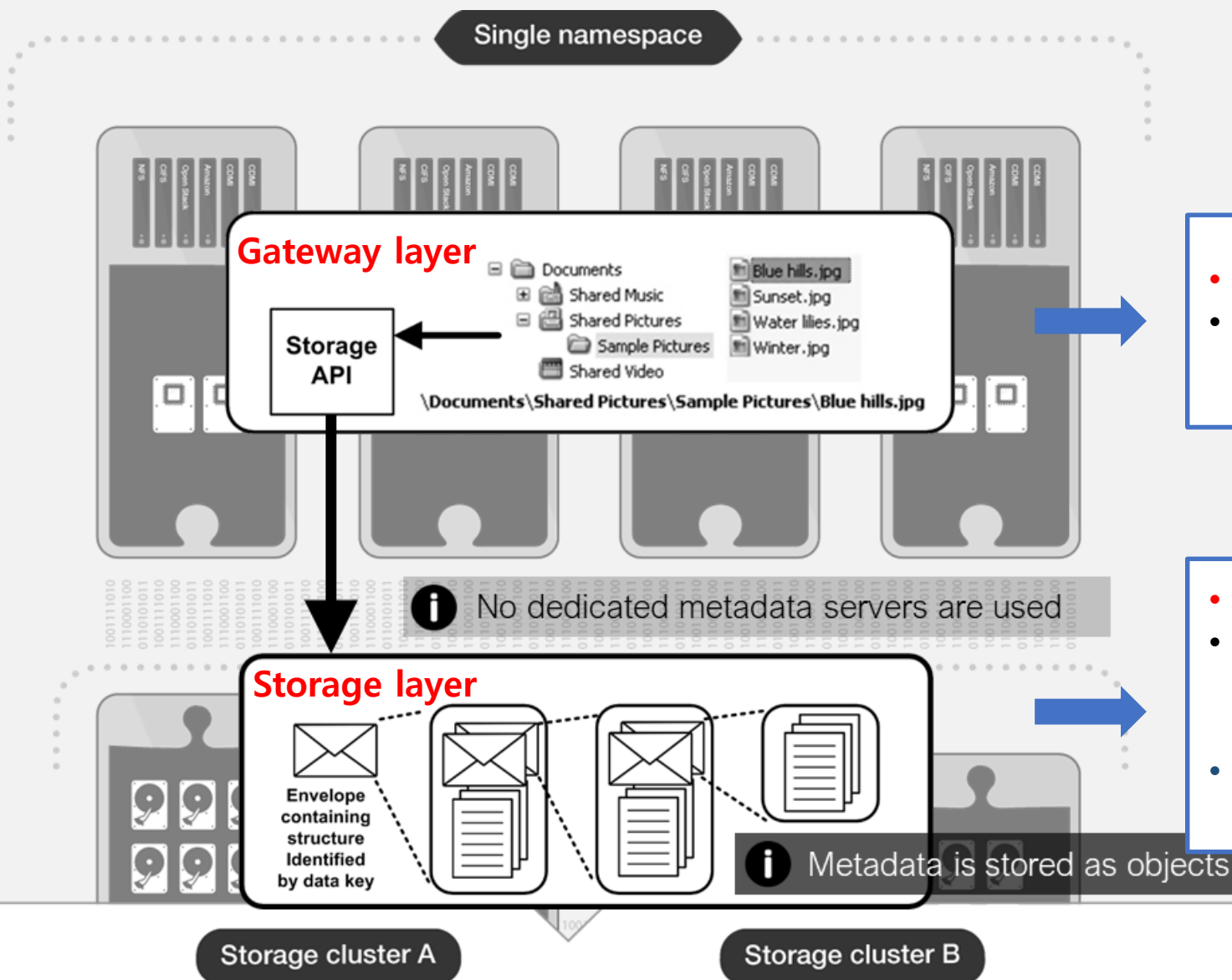


- 최대 노드 수 제한 없음.
- 파일 시스템 최대 크기 제한 없음. 파일 시스템당 최대 파일 개수 제한 없음.

Meta-Data Operation

IBM Spectrum NAS
Object Based File Storage

How is structure defined



- **사용자 인터페이스 영역**
- 사용자에게 친숙한 Directory 계층을 표현

- **데이터 저장 영역**
- 디렉토리 및 그 하위 개체의 Meta Data를 Envelope에 담아 하나의 Object로 저장
- **각 Envelope에 설정된 특정 Attribute는 하위로 계승됨**



(Directory 별) 가변 Erasure Coding

특장점

- Storage Cluster 내의 동일 File System내에서 Sub Directory 별로 별도의 Eraser Coding 가능
- 파일의 크기, 성능, 중요도 등에 따른 최적화된 Storage 구성

> ls /mnt/

./app ← 6 + 2

./home ← 2 + 2

./temp ← 8 + 1

8 + 2
(Data) (Parity)



Copies : 3, 5 Copy

Erasure Coding :	2+1	3+1	4+1	5+1	6+1	8+1
	2+2	3+2	4+2	5+2	6+2	8+2

Spectrum NAS의 Meta data 저장 구조 (계속)

일반적인 File System의 Meta Data 저장 Architecture

DB Table의 Field 수가 많아지면...
= Record 조회성능이 더 많이 떨어진다
= Meta data Access 성능이 더 많이
떨어진다
∴ Field의 수를 최대한 줄여야 한다



개별 파일의 특성에 대한 다양한 설정을
Meta Data를 통해서 할 수 없다.

Spectrum NAS의 File System Meta Data 저장 Architecture

Meta Data를 Object로 저장
= 하나의 Object의 크기와 다른 Object의
읽기 성능과는 관련이 없다
= 각 파일의 Meta Data에 다양한 정보를
기록할 수 있다

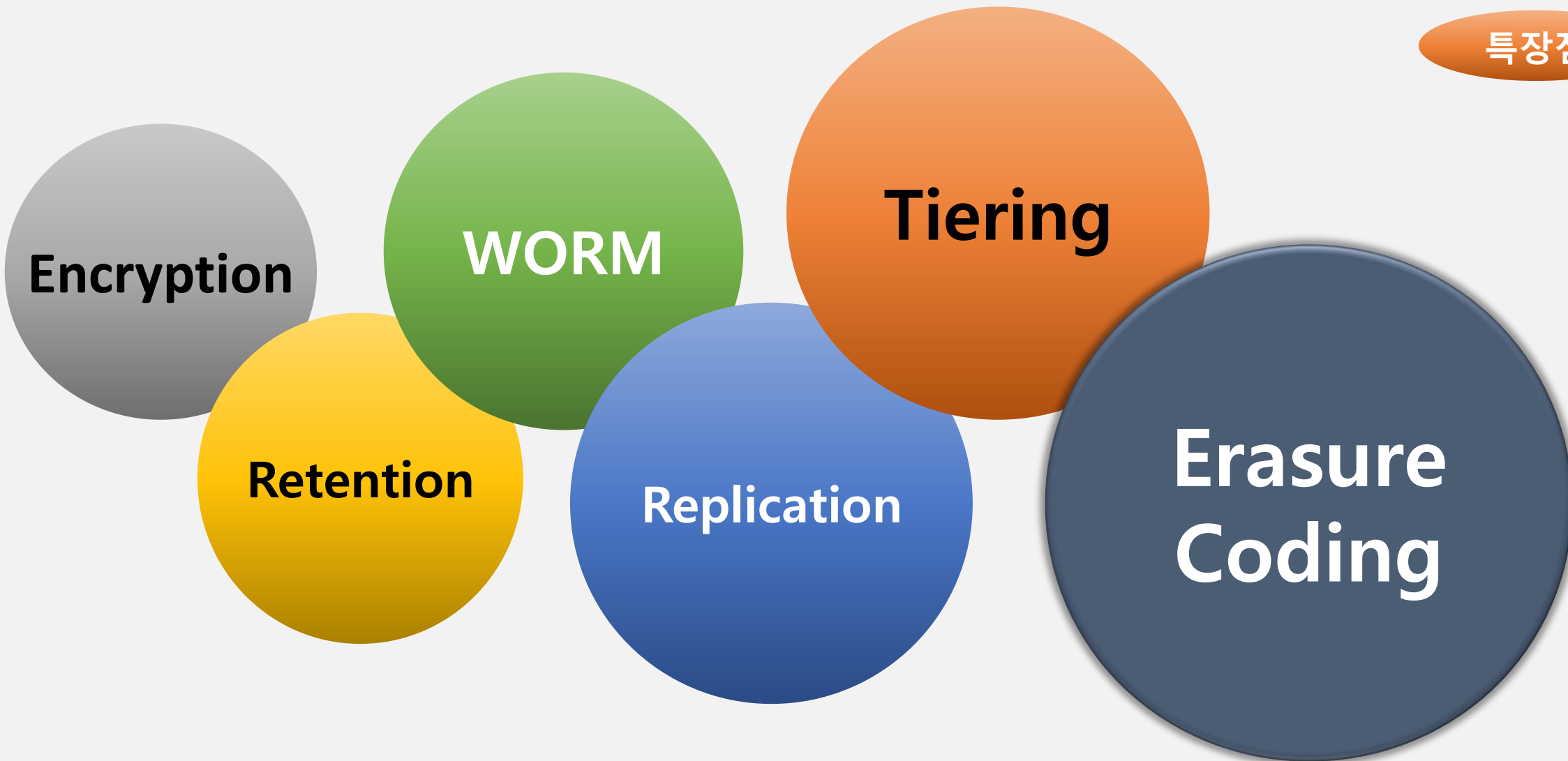


• 개별 파일의 특성에 대한 보다 다양한
설정을 Meta Data를 통해서 할 수 있다.

파일별로 개별 설정 할 수 있는 항목

IBM Spectrum NAS
Object Based File Storage

특장점



파일별로 개별 설정 할 수 있는 항목

IBM Spectrum NAS
Object Based File Storage



File Policy

Filters:

Actions applies to all files immediately

☐ Pattern:

☐ Age:

Actions:

☐ File coding:

☐ Tier:

☐ Data encryption

☐ Retention Modified ago:

Accessed ago:

☐ WORM:

Save

- Spectrum NAS 개요
- Object based File Storage
- 유연한 성능 및 용량 확장성
- TCO 최적화
- 신뢰성, 정합성, 안전성 극대화
- Data 보호 및 가용성 극대화
- 운영 효율 및 관리 편의성

노드 추가에 따라 용량 및 성능 확장

IBM Spectrum NAS

유연한 성능 및 용량 확장성



- 매 노드 증설 시 마다

- ✓ 용량 : 내장 Disk의 용량에 Erasure Coding 적용 후 가용 공간 만큼 증가
- ✓ 성능 : 노드 모델에 따라 600MB/s ~ 5GB/s 의 Read 성능 증가

- 최대 노드 수 제한 없음.
- 파일 시스템 최대 크기 제한 없음. 파일 시스템당 최대 파일 개수 제한 없음.



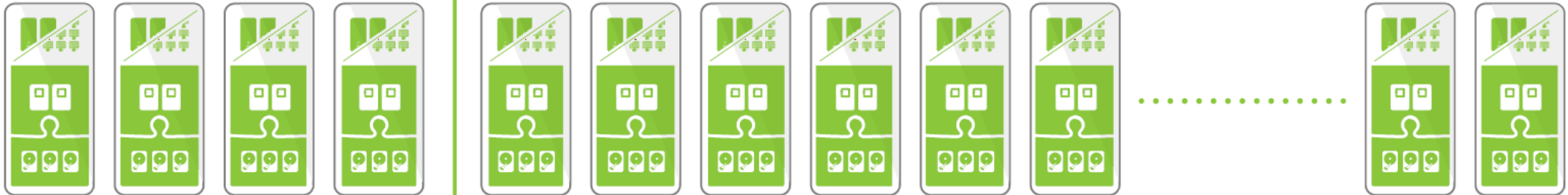
IBM Spectrum NAS

Commodity X86 Server를 사용하여 Cost Saving



- 표준 X86 서버 및 IP Network
- Low grade CPU 및 일반적인 서버 구성 사용

Setup starts with four nodes



Linear scaling up to hundreds of nodes

성능, 효율성, 데이터 보호 극대화 등 고객 요구 환경에 최적화

Performance

- Latency 최적화
- Read/Write 성능 최적화
- 유연한 NIC 구성

Protection

- Sync. & Async. Replication
- Snapshot
- WORM, Encryption
- **Directory 별 Parity 비율 설정**

Efficiency

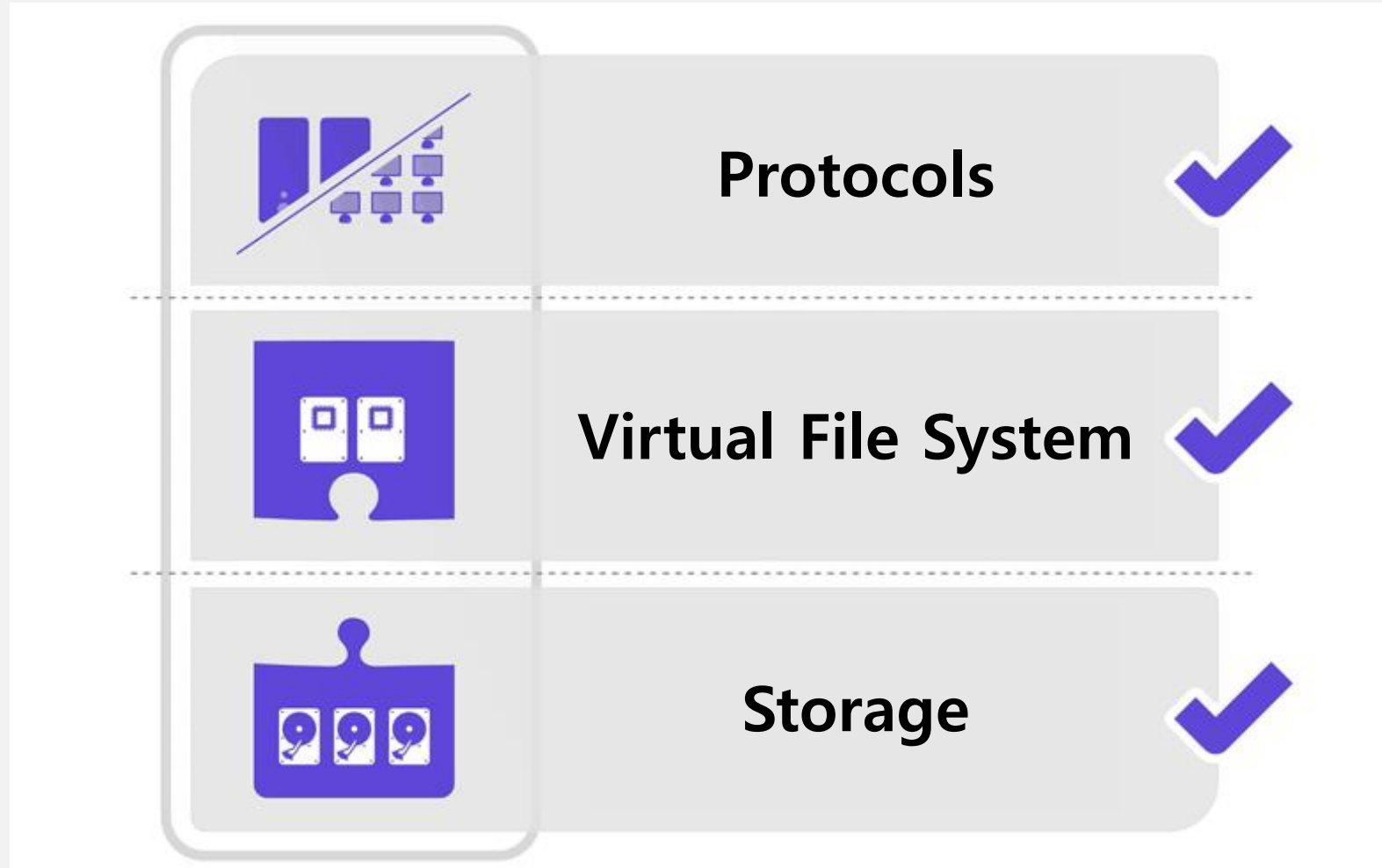
- Storage Tiering
- Quota
- Multitenancy

**비용 Vs. 효용의
Micro-Adjustment
(미세조정)**

SW 스택에 통합된 Clustered Protocol

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화



**NO
OPEN
SOURCE**

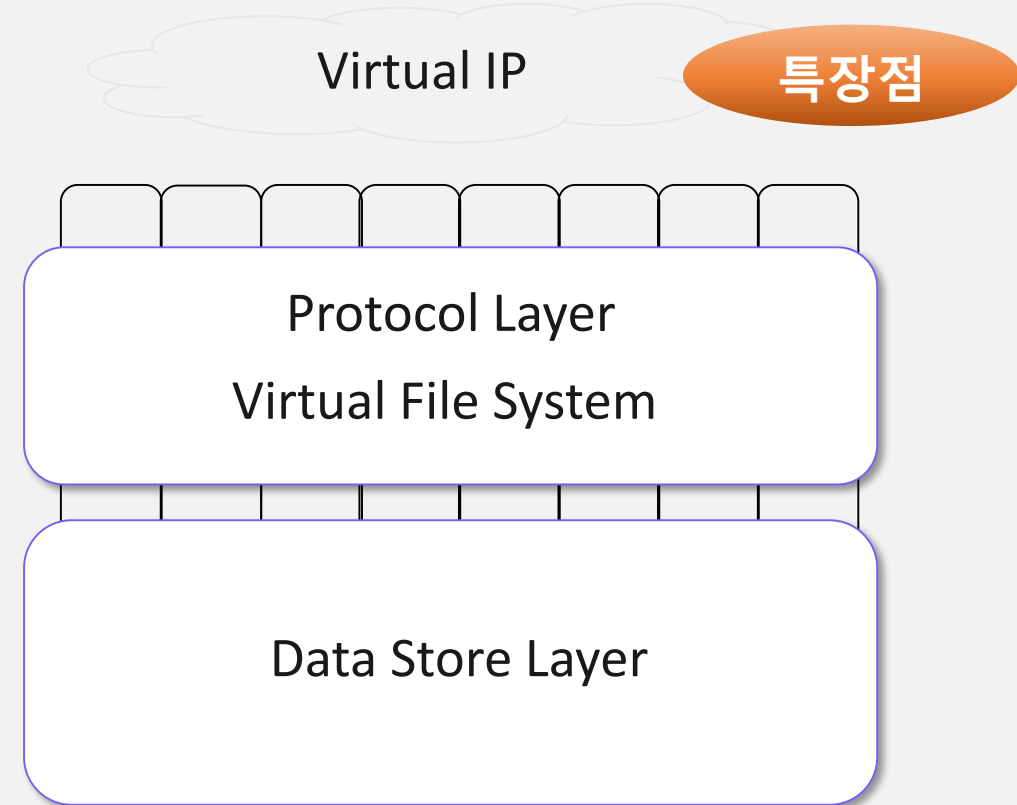
**NO
3rd Party
SOFTWARE**

SW 스택에 통합된 Clustered Protocol

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화

- **Virtual FS와 Clustered Protocol은 하나의 Stack으로 코딩되어 서로의 상태를 모니터링합니다.**
 - ✓ SMB 상태는 가상 파일 시스템에 저장되어 노드 장애 시 원활한 호스트 재 연결을 가능하게 합니다.
 - ✓ 오픈 소스 코드가 없는 프로토콜 계층은 새로운 노드 추가 시 Scaling Protocol Function을 실행합니다.
 - ✓ Write cache는 1개 또는 2개의 복제본으로 보호됩니다.
 - Default : 1 개의 복제본 (= 2 copy)
 - ✓ Read 또는 Write cache 영역은 Access pattern에 따라 동적으로 조정됩니다.
 - ✓ Client connection failover를 위한 Virtual IP 지원

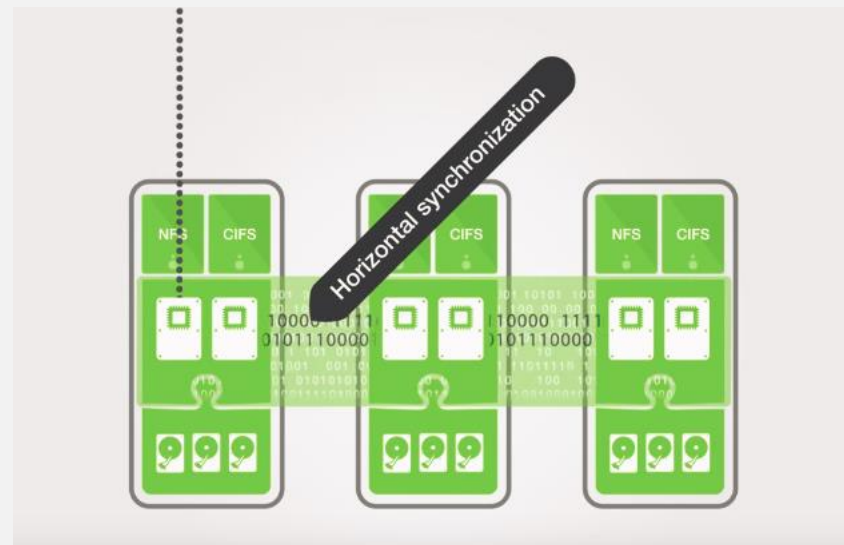
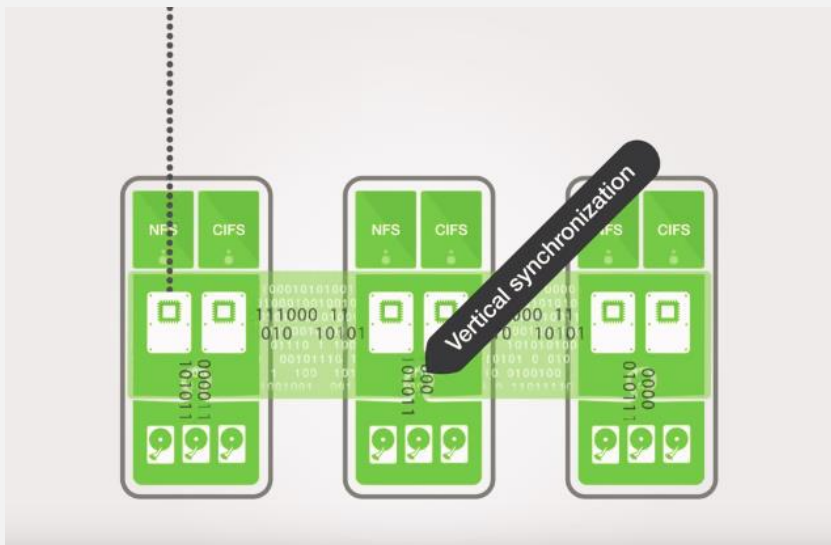


오픈 소스가 없이 자체 개발하였기 때문에, 프로토콜 계층이 소프트웨어의 나머지 부분과 완전히 통합되어 있습니다. 이로 인한 특징으로 한가지 예를 들면, NFS/SMB 프로토콜은 파일 수신 시, Write cache에 쓰기와 다른 노드의 cache 복제를 동시에 수행합니다.

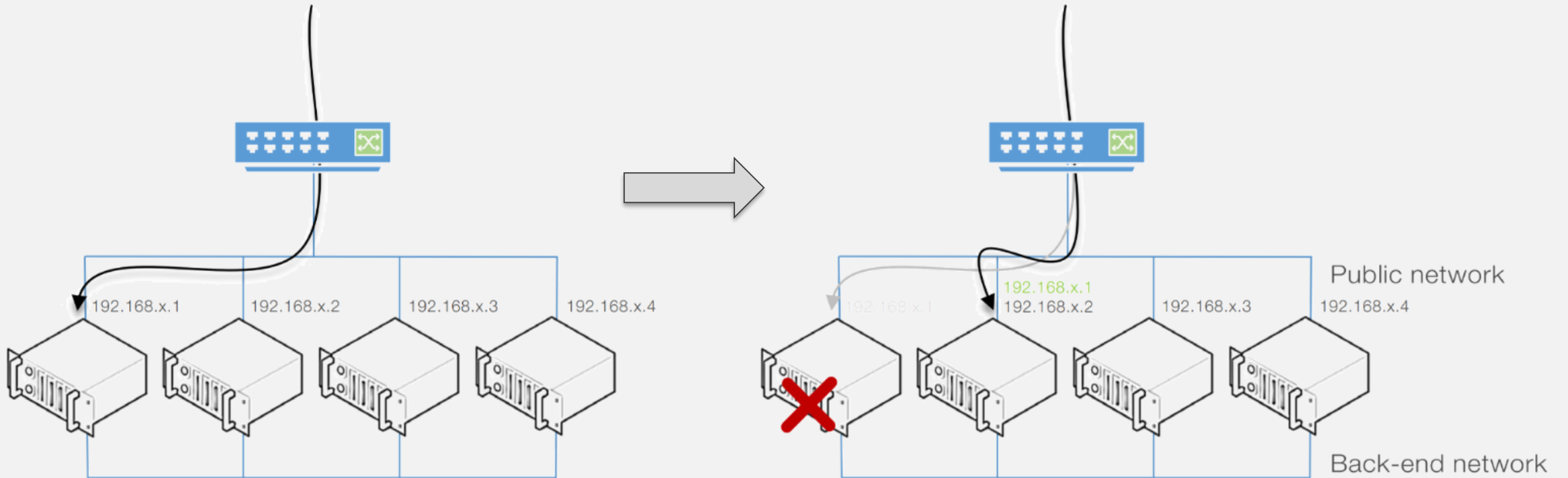
Write Operation – Cache Sync

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화



- Cache Mirroring 지원 (2 Copy, 3 Copy)
 - ✓ File별 Cache 복제 파트너 선정
 - ✓ Cache Mirroring → Off 가능
 - ✓ 지정된 시간 만큼 디스크로 내려쓰기 지연 가능
- 하나의 파일을 위한 다중 서버 읽기 Cache 구성 가능

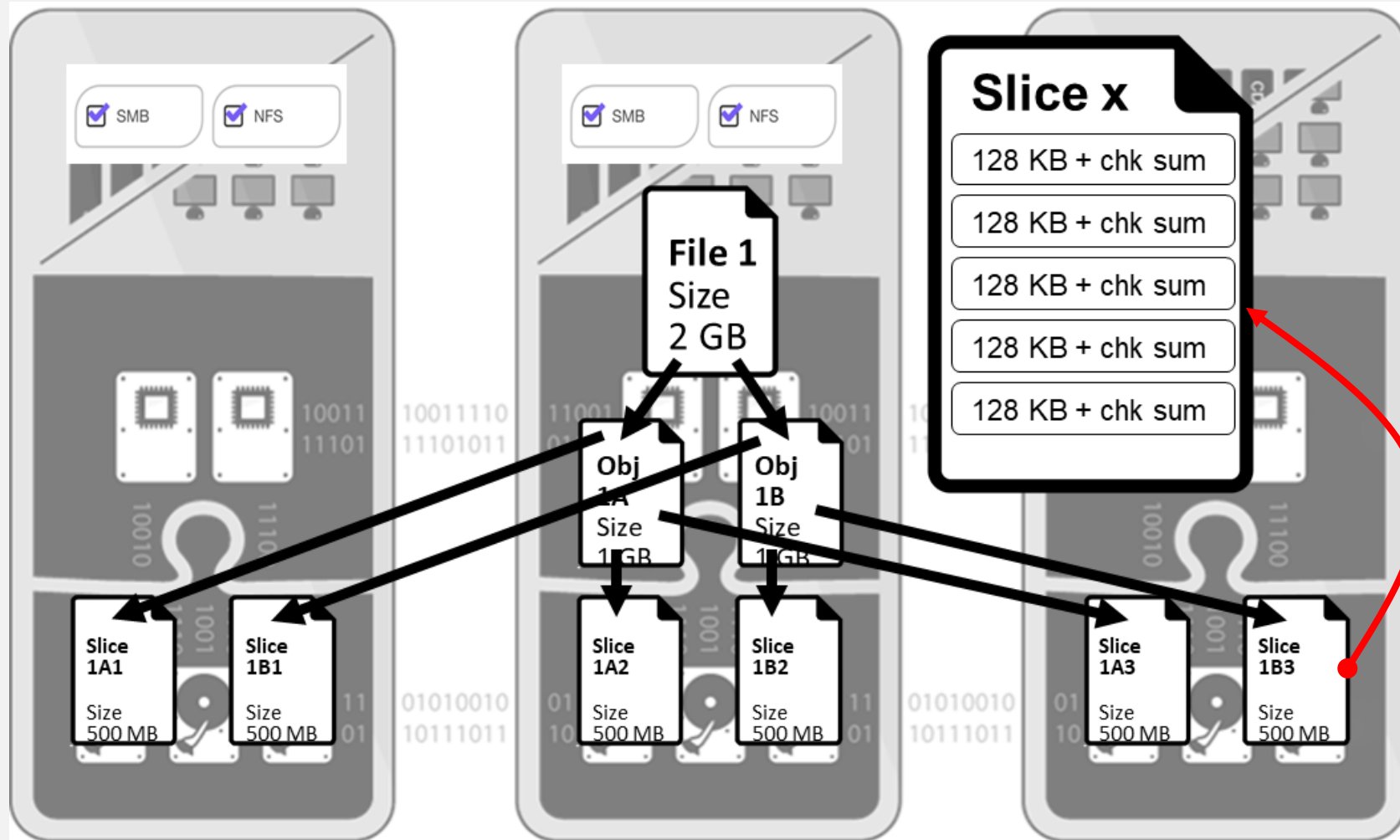


- 각 노드마다 정의된 가상 IP는 노드 장애 시 정상 노드로 이전됩니다.
- 온라인 중, S/W 등을 Upgrade 할 수 있습니다.

File Encoding

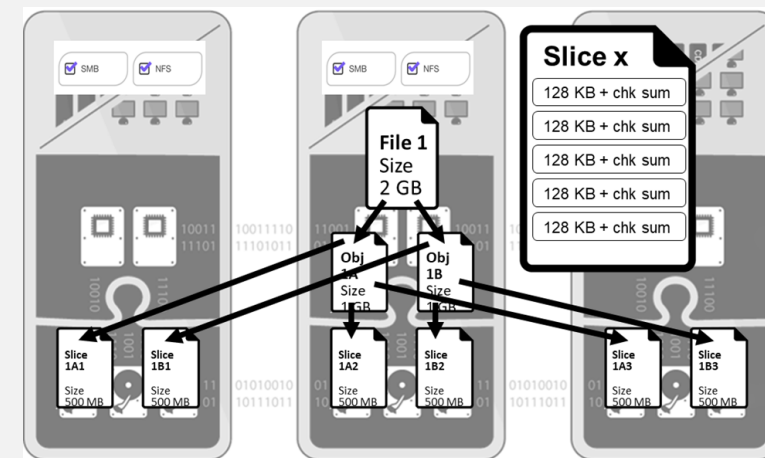
IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화



File encoding: 2+1 Extent (object) size: 1 GB

- File이 1GB보다 크면 Data 저장소에 1GB 크기의 여러 Data Object로 분할 저장됨
- 각각의 Data Object는 Policy에 정의된 File Erasure Coding 정책이 적용됨
- Data Objects는 각 노드의 상태(빈 공간, 성능 등)에 따라 최상의 노드가 선택되어 최적화된 저장이 이루어짐
- Erasure Coding 정책에 의해 만들어진 각 조각에는 128KB block마다 Check Sum을 추가해 안정성을 극대화함
- 각 Envelope들은 해당 Directory에 포함된 모든 콘텐츠에 대한 정보를 저장함 (files and other directories + their metadata). 하나의 Envelope는 하나의 Data Object로 저장됨
- 노드 또는 HDD에 장애가 발생하면 다른 모든 노드가 rebuilding에 참여함



File encoding: 2+1 Extent (object) size: 1 GB

질문 : df 명령어에 대한 결과는?



파일 시스템 초기 생성시에는 개별 Disk의 전체 용량이 보임

- 개별 Disk 전체 용량을 2진수 용량으로 보여줌 (GiB, TiB, PiB)
- Parity, Hot Spare 영역이 별도로 사전 할당되지 않음
 - ✓ 특히, Hot Spare Disk에 대한 별도의 지정 없음



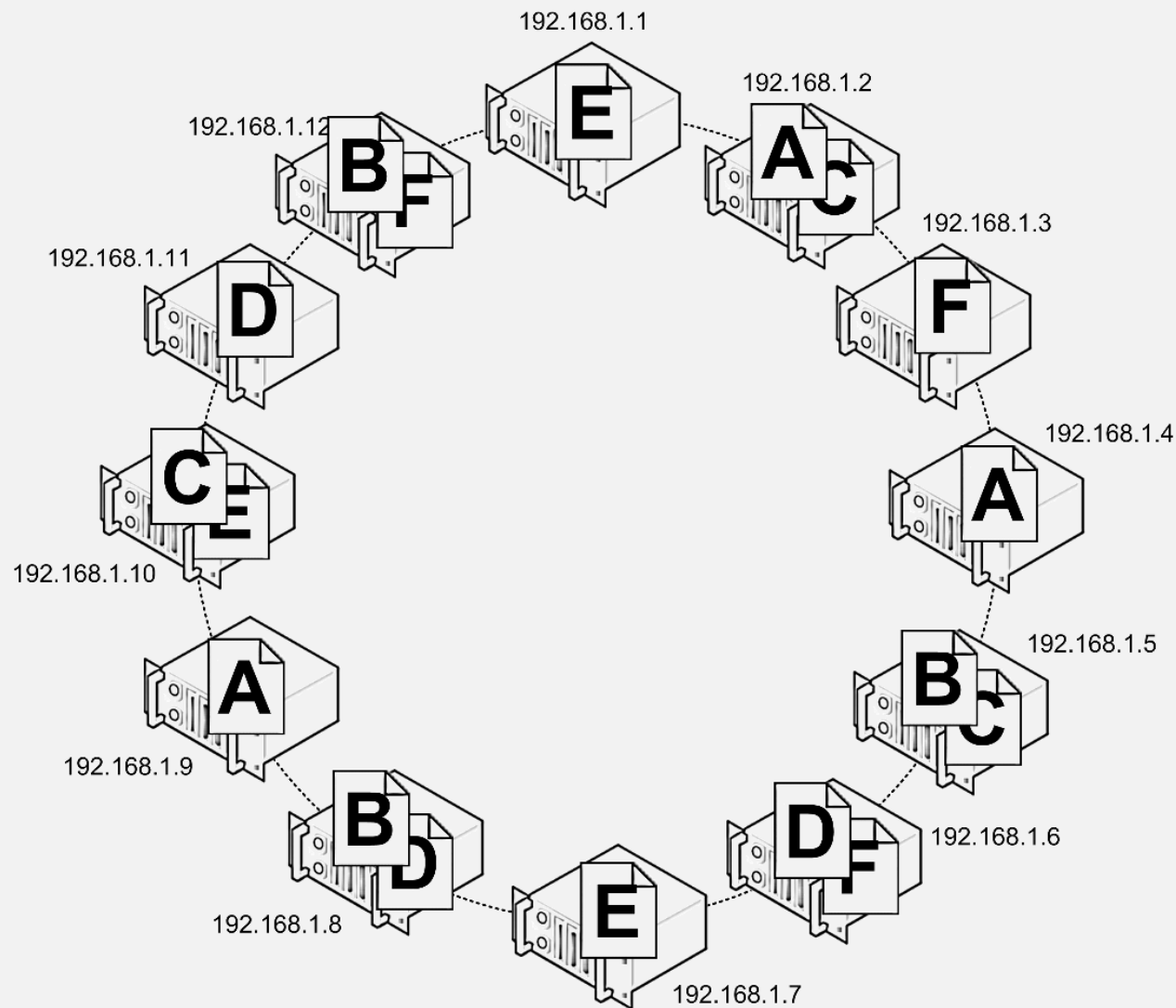
Data 저장 시 Parity 영역이 함께 할당되는 방식임

- Erasure Coding 2+1 설정에서 1GiB 파일을 저장하면 1.5GiB의 영역이 할당됨
 - ✓ Disk 전체 용량에서 1.5GiB가 줄어 듦
 - ✓ 나머지 미 할당 전체 영역은 노드 또는 Disk 장애 시 복구를 위한 공간으로 쓰임

클러스터 내에서의 Data 배분 방식

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화

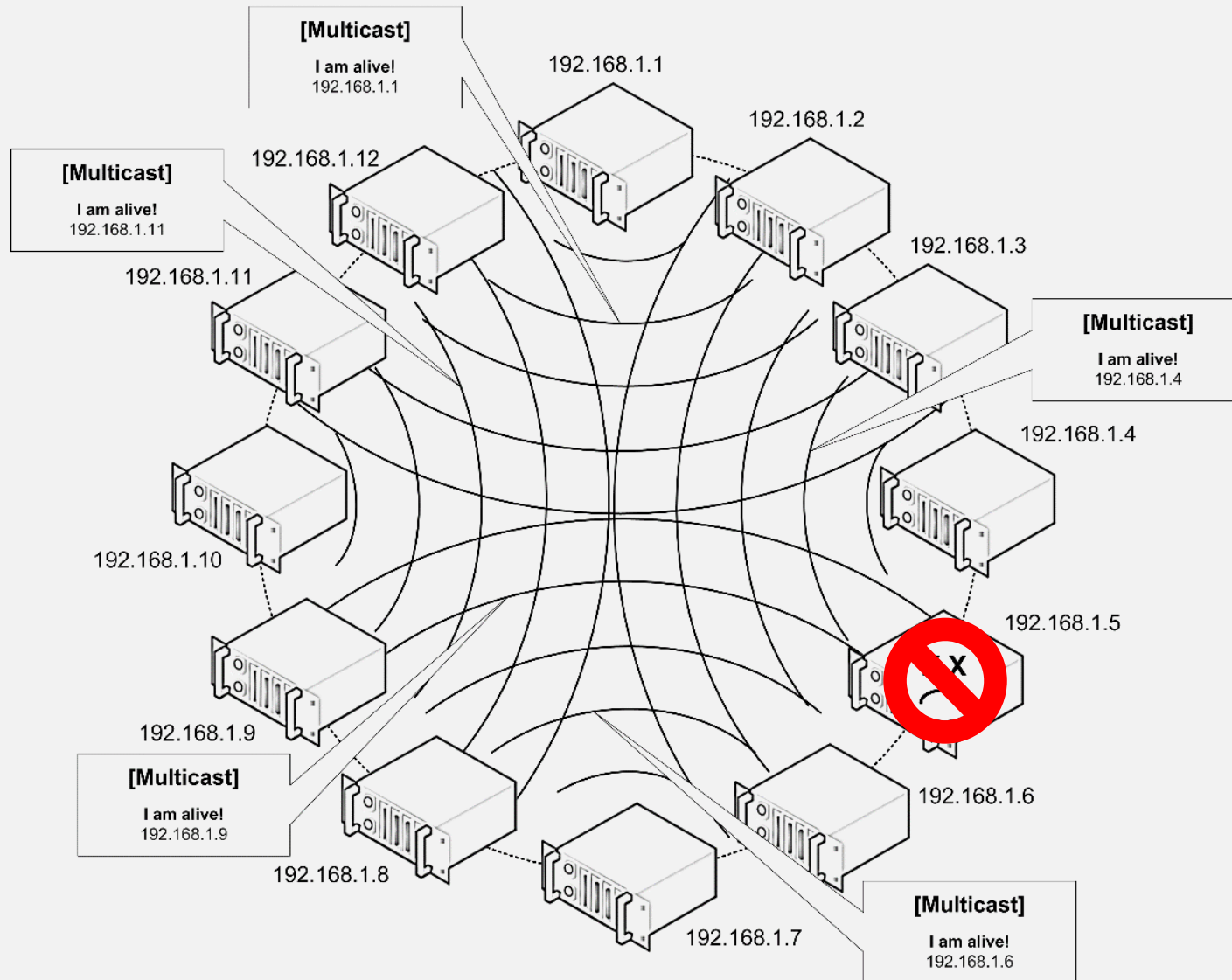


- 노드간 1대1 매핑 없음
- Data 위치 선택 조건 :
 - ✓ 노드의 빈 공간
 - ✓ 지리적 위치
 - ✓ 현재 부하량
 - ✓ Tier 정책
 - ✓ Random

노드의 Health Check

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화



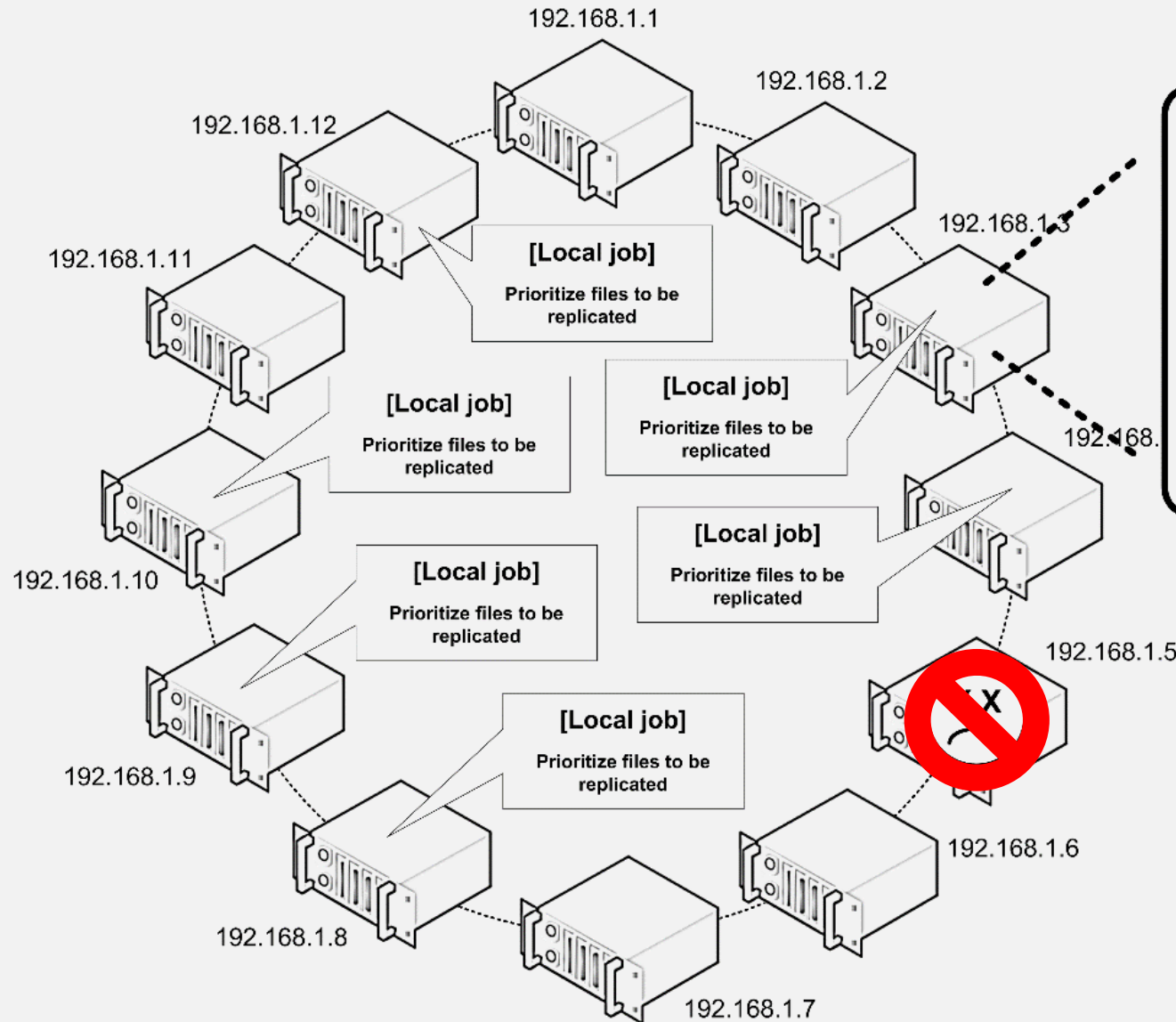
- 모든 노드는 동일한 역할을 하므로 전체 Node의 Health Check를 하는 관리 노드 없음
- 각 노드들은 주기적으로 자기 상태를 알리는 Multicast Packet을 Cluster Network에 전송

Self Healing

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화

특장점



Storage node Replication priority lists

High	Med	Low
A4099325-FD-4D... 6A206642-EB-47...	252212D8-87-4D... 717B9C3C-C2-47... 5A8E6CA8-2B-4C...	613DB68B-0A-41... 867230C1-88-4E... CBDB0010-42-45... 070326D9-03-41... E7A51F85-59-42...

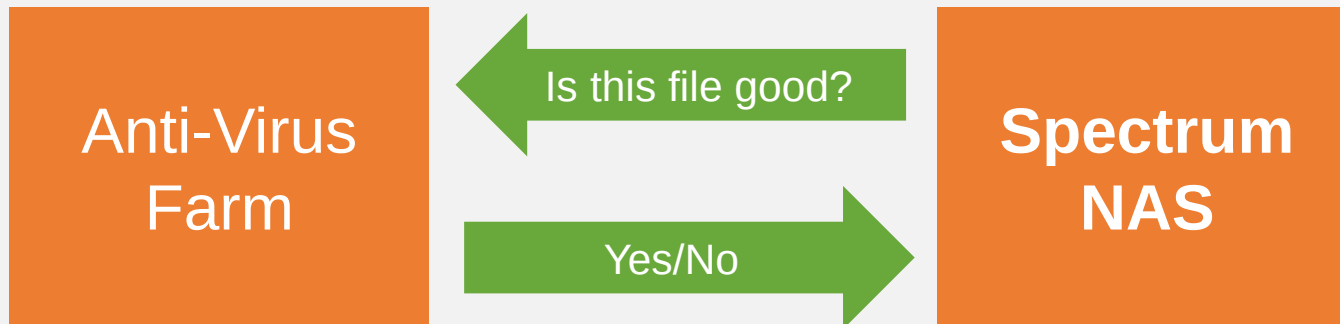
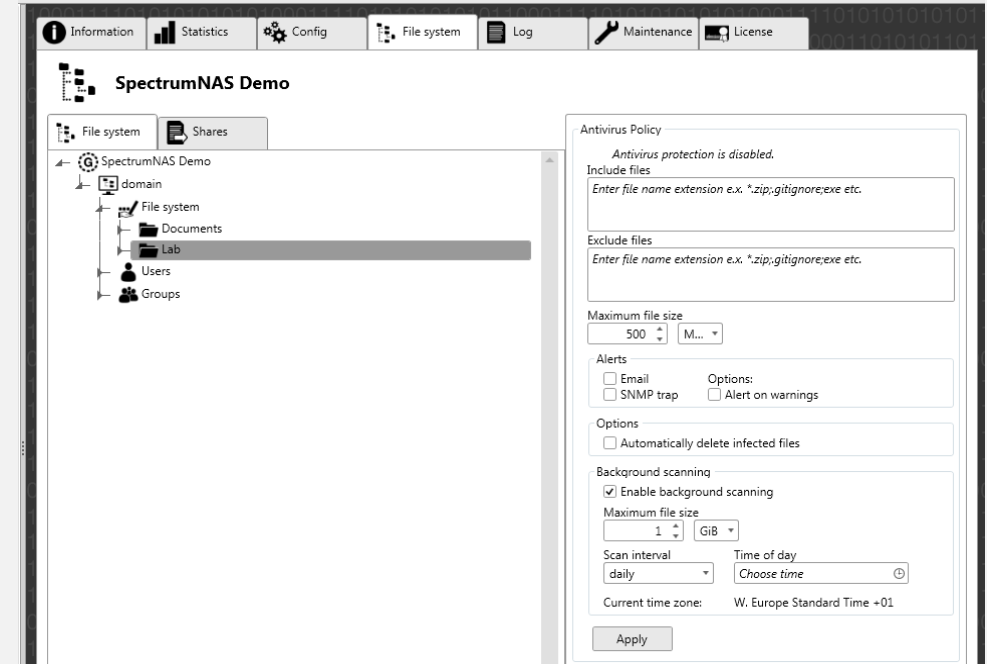
- Erasure Coding의 특징
 - ✓ Hot Spare 영역이 별도로 지정되어 있지 않고 Free Block이 Hot Spare를 대신 함.

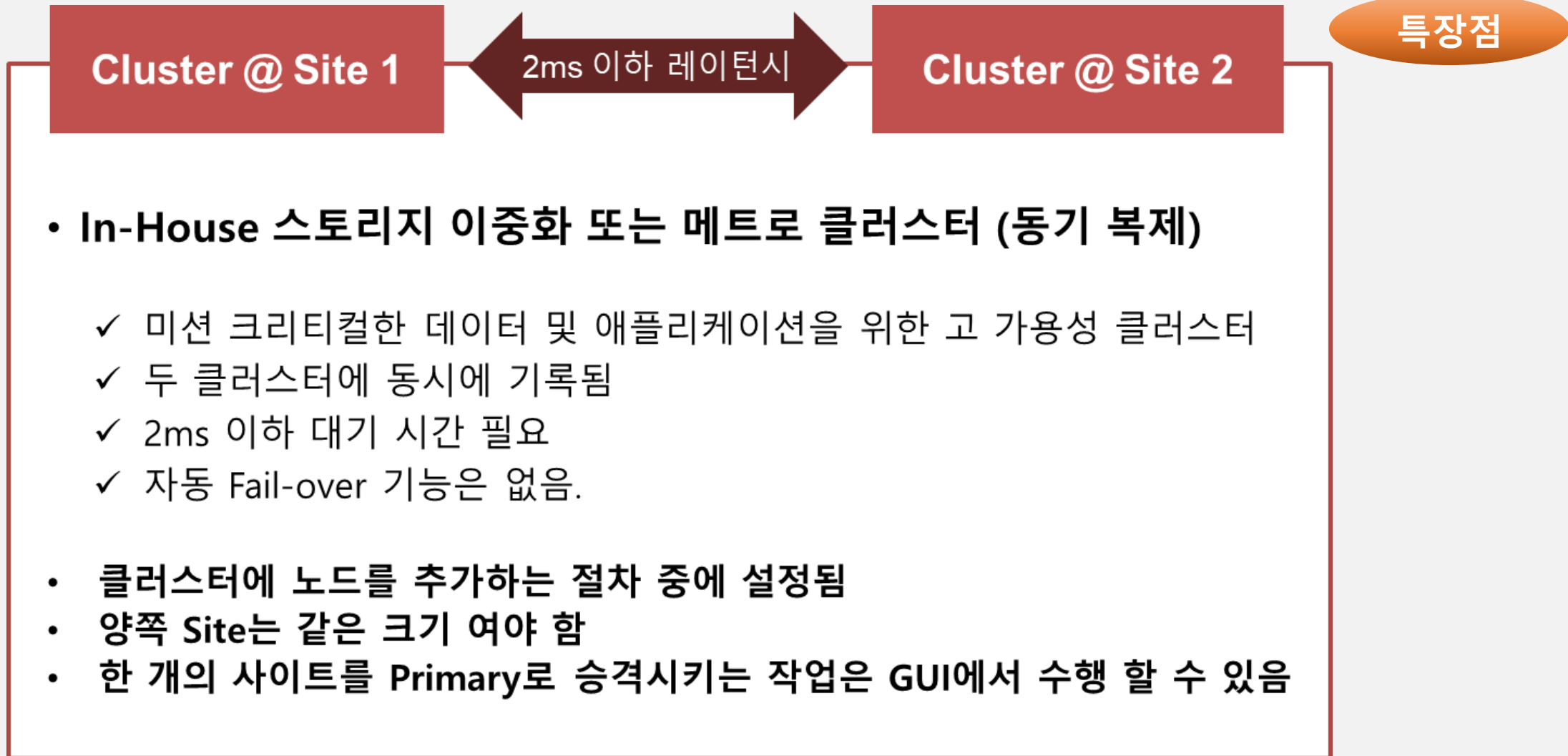
Anti-virus support for files

IBM Spectrum NAS

신뢰성, 정합성, 안전성 극대화

- ICAP protocol을 통해 지원
- 현재는 **Symantec Protection Engine Anti-Virus server** 만 지원
- 파일을 읽을 때, Anti-virus server에 바이러스 체크 의뢰가 이루어 짐
- 파일이 아직 Cache에 있고 Anti-virus definition Update가 없다면 파일을 다시 체크하지 않음
- 새로운 Anti-Virus Server의 적용을 위한 추가 작업이 진행 중임





Spectrum NAS – Async. Replication

IBM Spectrum NAS

Data 보호 및 가용성 유지

- **Hybrid-Cloud (Async.-Replication)**

- ✓ Snapshot based
- ✓ 이전 Snapshot과 비교 후 달라진 부분을 복제
- ✓ 복수의 Target으로 복제 가능

Create Hybrid-Cloud

Hybrid-Cloud Name to Create

myHybridCloud

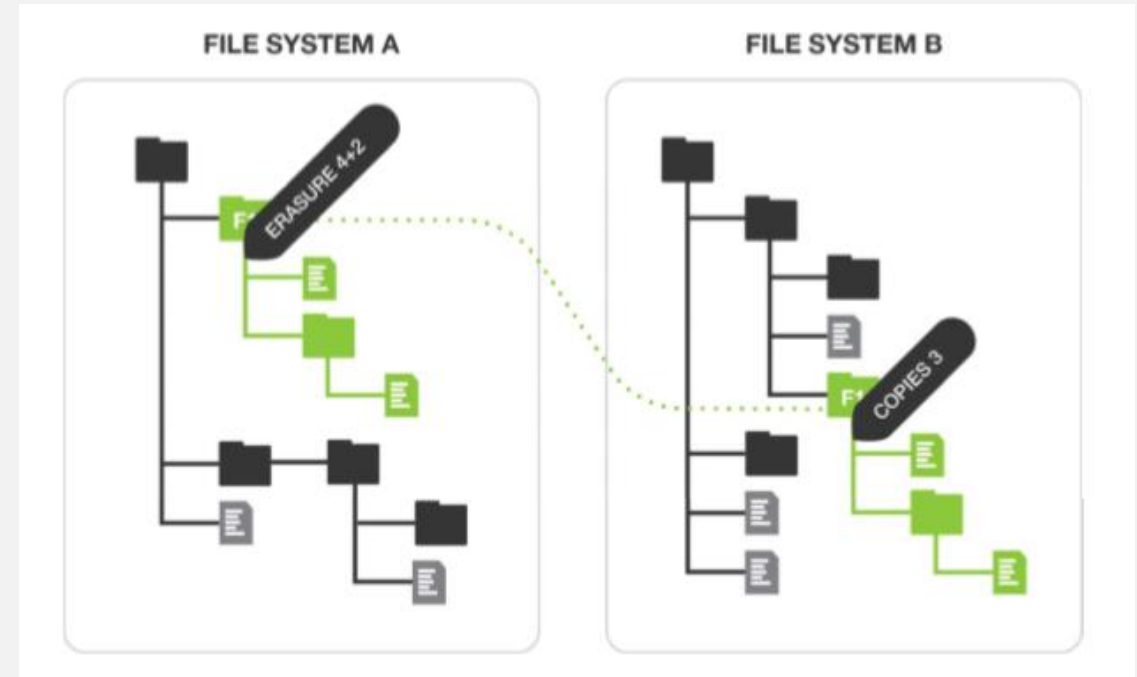
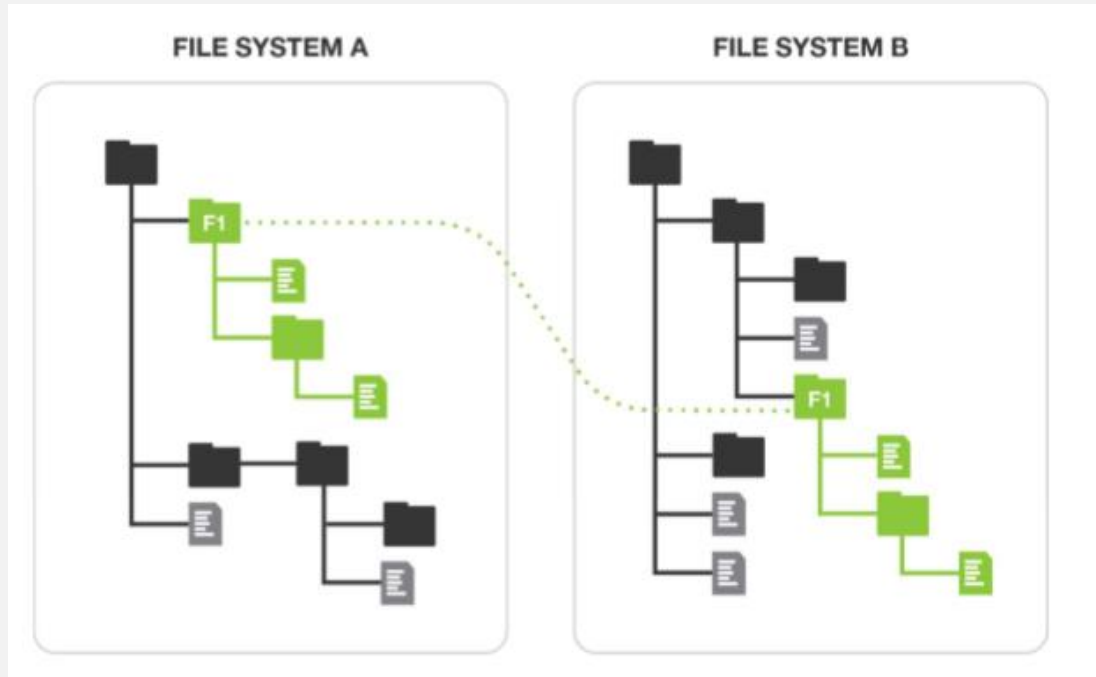
Local Data Center Name

DataCenter

Local Data Center URL

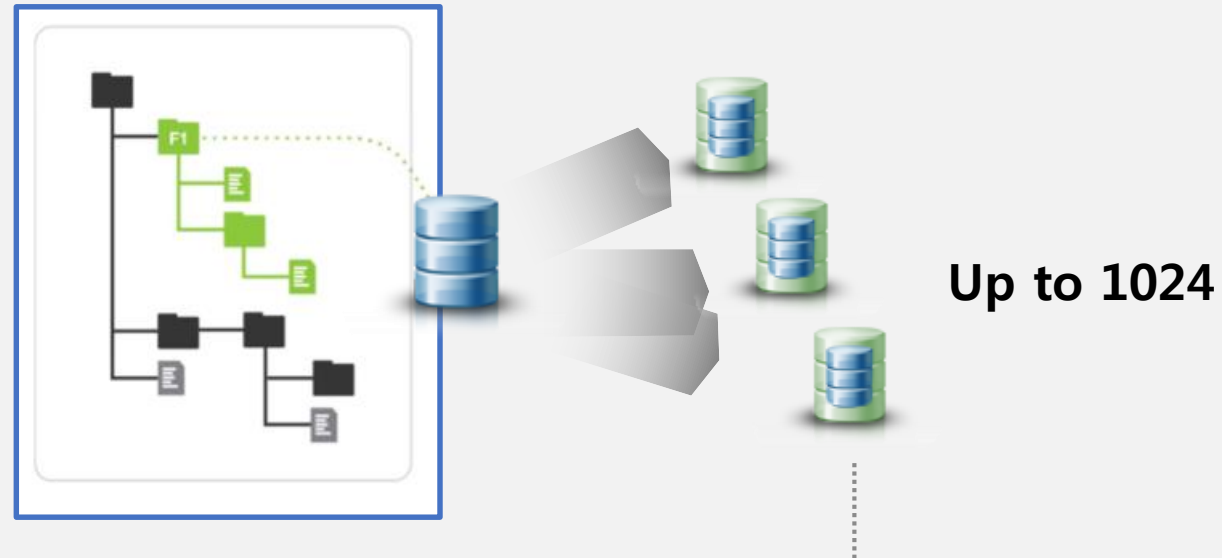
http://

Apply



✓ (Directory 별) Snapshot 설정

- 각 **Sub Directory** 별로 **1,024개**의 Snapshot 생성 가능
- Snapshot 정책이 폴더에 할당되면 모든 하위 폴더 및 파일에 유효

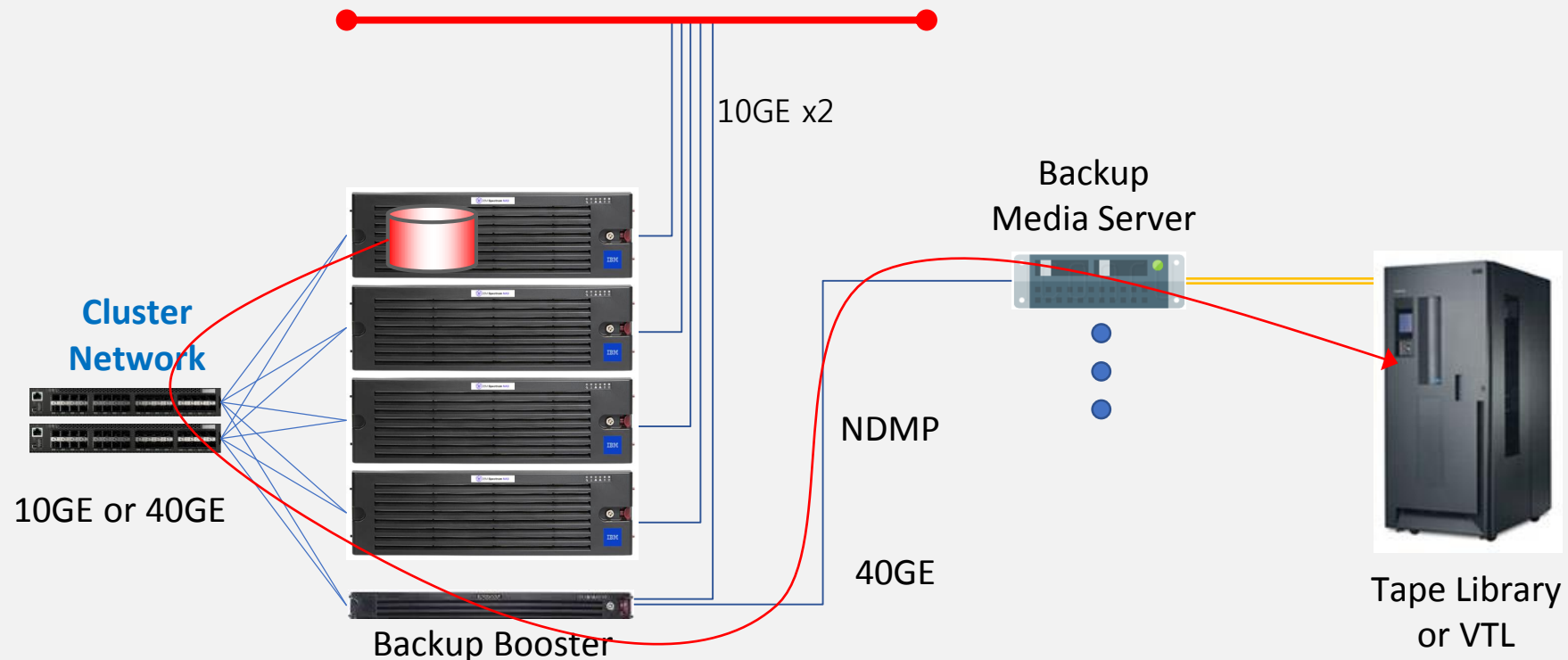


Spectrum NAS Backup options

IBM Spectrum NAS

Data 보호 및 가용성 유지

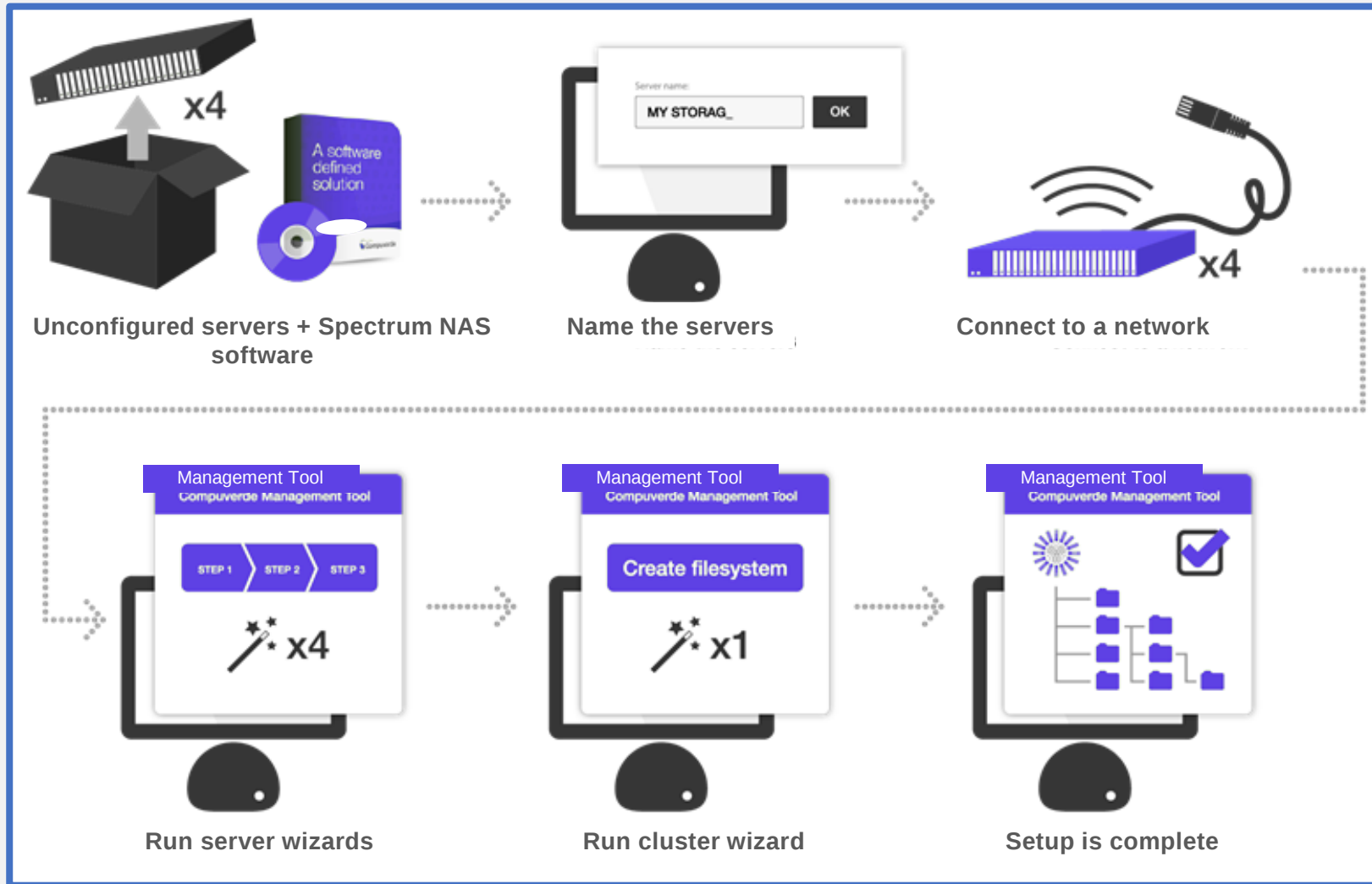
- ✓ NDMP with Spectrum Protect and CommVault is announced.
- ✓ Spectrum NAS NDMP is NOT same as traditional NDMP
- ✓ Spectrum NAS NDMP uses Snapshot diff to find changed files
- ✓ Traditional NDMP walks through file system to find changed file which is very slow



Easy to Use and Manage

IBM Spectrum NAS

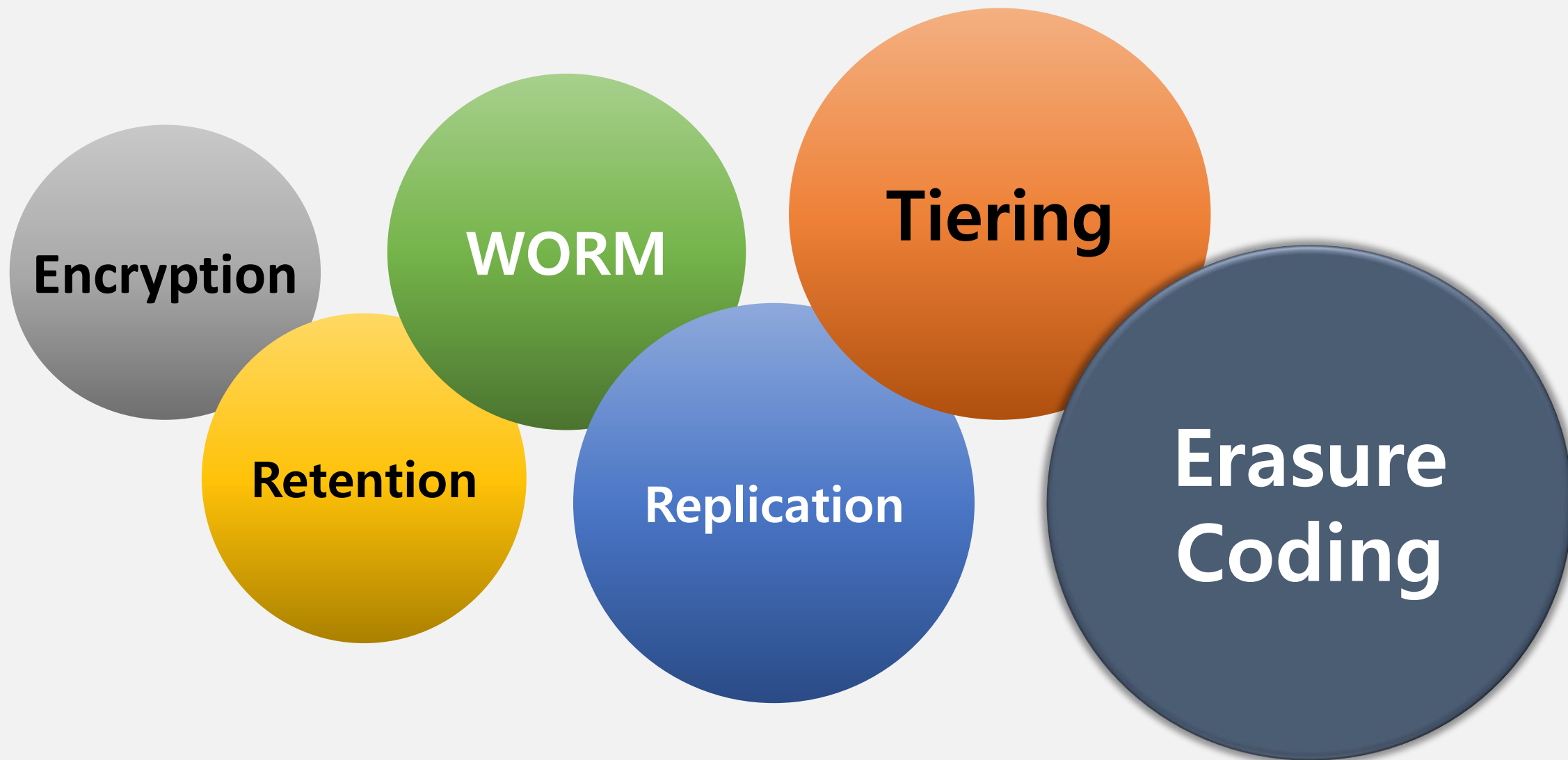
운영 효율 및 관리 편의성



파일별로 개별 설정 할 수 있는 항목

IBM Spectrum NAS

운영 효율 및 관리 편의성



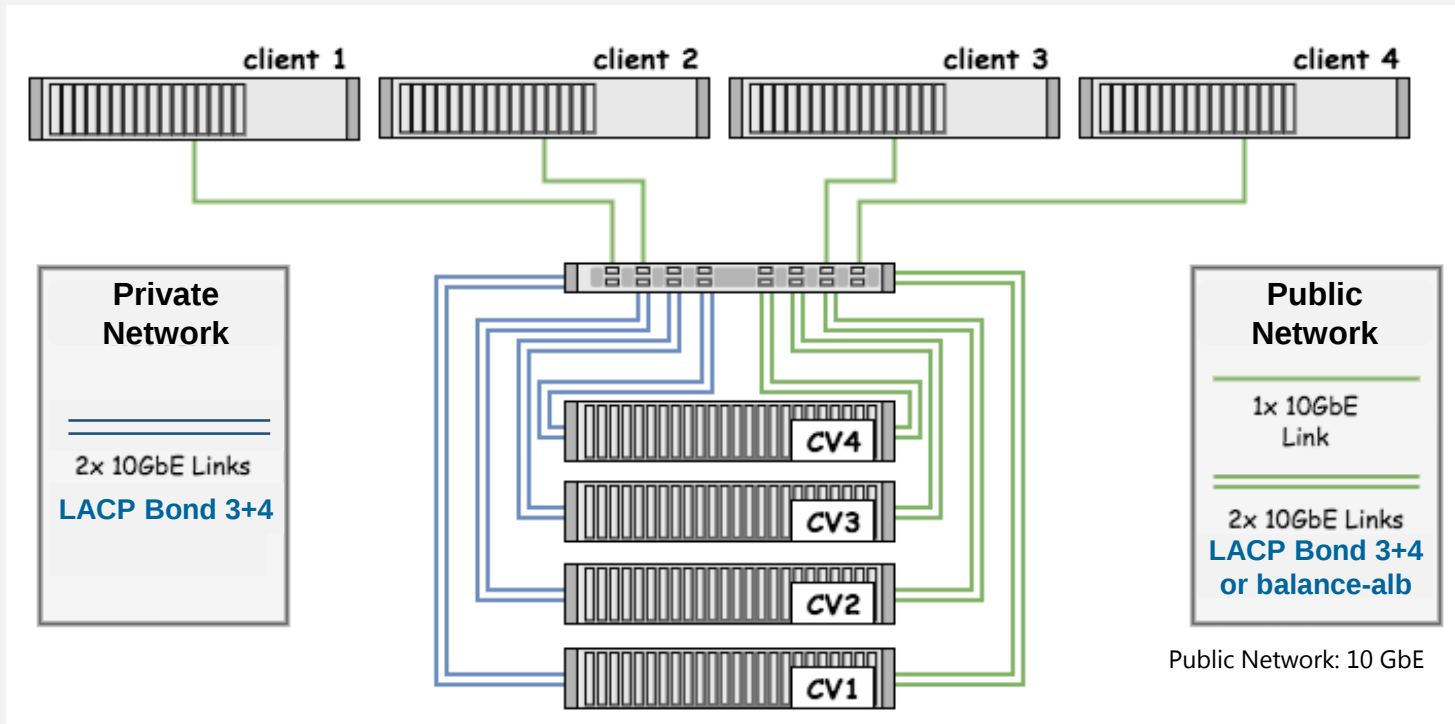
Spectrum NAS - Multitenancy

- 한 개의 스토리지 클러스터에 복수개의 분리된 Domain 및 File System 생성 가능
- 각 Domain과 File system은 전용 IP, 인증 방식, 프로토콜 집합 등을 다르게 구성할 수 있다.
- CPU, Cache, Storage 저장공간 등은 모든 Domain들이 공용으로 사용할 수 있다.

Multi-tenancy Feature	Spectrum NAS	Comments
Separate File Systems	✓	Separates data between tenants
Separate Networks	✓	Separate networks – NICs or VLANs
Separate Authentication	✓	Separate AD, LDAP, etc

Performance Benchmarks

Performance POC Environment



Each node is configured with:
12 HDDs, 4 TB each
2 NVMe drives, for cache
375 GB each
Data protection: 2+1
Total cluster capacity = 170 TB

SuperMicro SuperServer x3650 M5, model 8871-AC1

CPU : 2 * Intel Xeon Silver 4114 (13.75 MB cache @ 2.2GHz, 10 cores, 2400MHz Memory Clock * 6 channel, max 48 PCIe lane)

https://ark.intel.com/products/123550/Intel-Xeon-Silver-4114-Processor-13_75M-Cache-2_20-GHz

Memory: 96GB (Configured 2400 MHz, Physically 2666MHz) 8GB * 12 channel for maximum channel configuration

Network: 10GbE, Mellanox Dual Port SFP+ ConnectX-3 Pro based adapter (two NIC with two ports each)

Drives: 12x 4TB 7200 rpm HGST NL-SAS; 2x Intel Optane P4800X 375GB PCIe NVMe

<https://www.intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds/optane-dc-p4800x-series.html>

- 2+1 Erasure Code (1 fault resiliency) : 67 % utilization (33 % parity)
- 1-cache Replication @ Cluster config
- Read/Write Cache with RAM Cache Enable @ Node config / Gateway
- Disable disk write cache @ Node config / Storage

- NFSv4 test only

1 Client Test

4 GB file 16 개

Write : 762 MB/sec (726 MiB/sec)

Read : 1209 MB/sec (1153 MiB/sec)

4 Client Test (Single stream each)

각 노드 당 4 GB file 16 개, 총 64 개 파일

Write : 3.0 GB/sec (2.8 GiB/sec)

Read : 4.8 GB/sec (4.5 GiB/sec)

- **32개 Job Test (8개 Job / 4 node 각각)**

각 Job 당 120 MB file 270개, 총 8640개 파일, 총 1TB

- - Write : 2225 MB/sec
 - (3228.5 MB/sec – no replication, cache mode : write only)
- - Read : 4967 MB/sec
- - Mixed
 - Read 60% : 2583 MB/sec
 - Read 80% : 3899 MB/sec
 - Read 90% : 4424 MB/sec

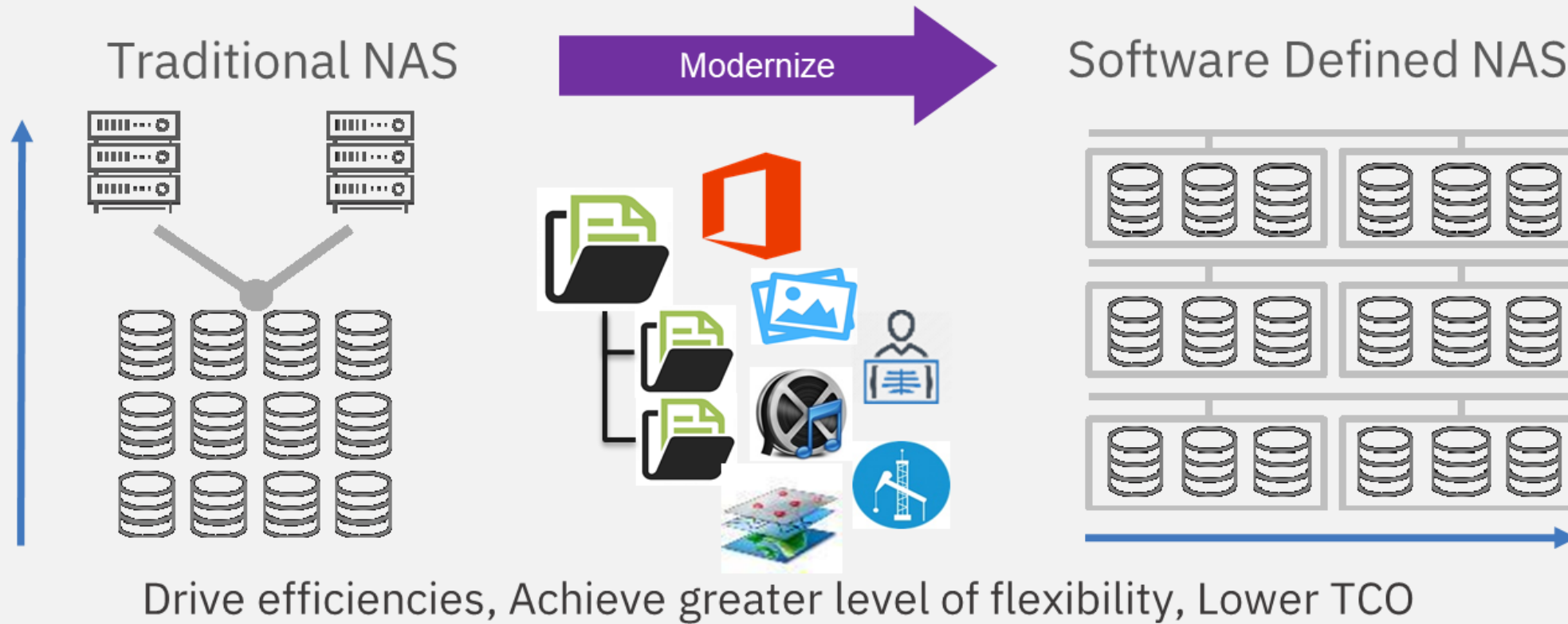
Spectrum NAS 세미나 정리

- 구형 Architecture의 문제점 개선

Modernize and Transform

IBM Spectrum NAS

구형 Architecture의 문제점 개선



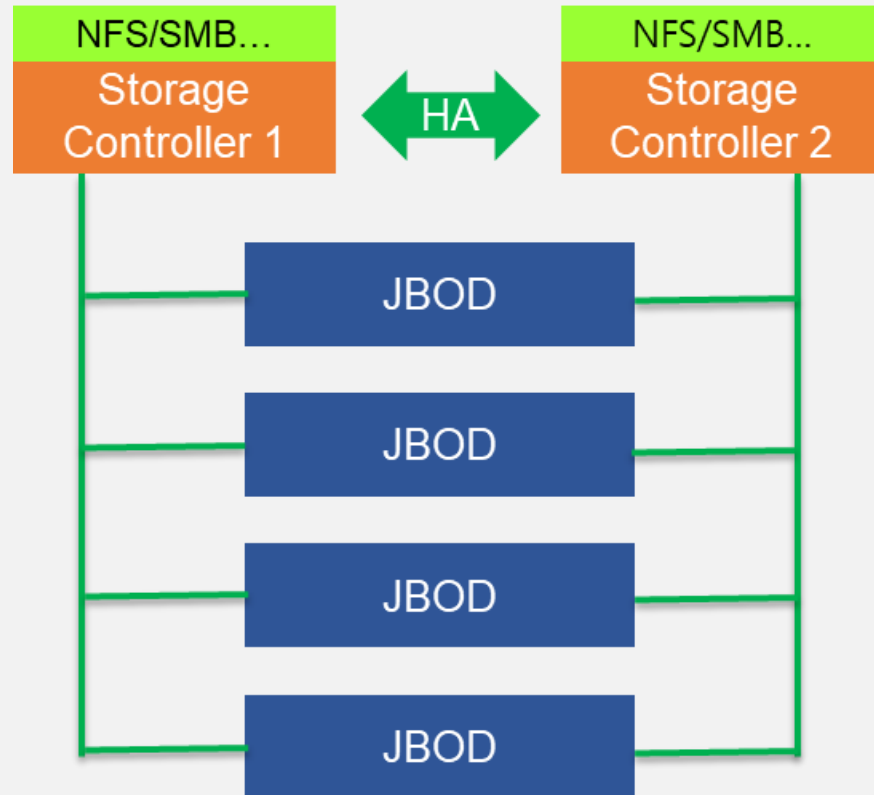
시스템이 수용가능한 Data 저장

저장할 Data에 최적화된 Storage

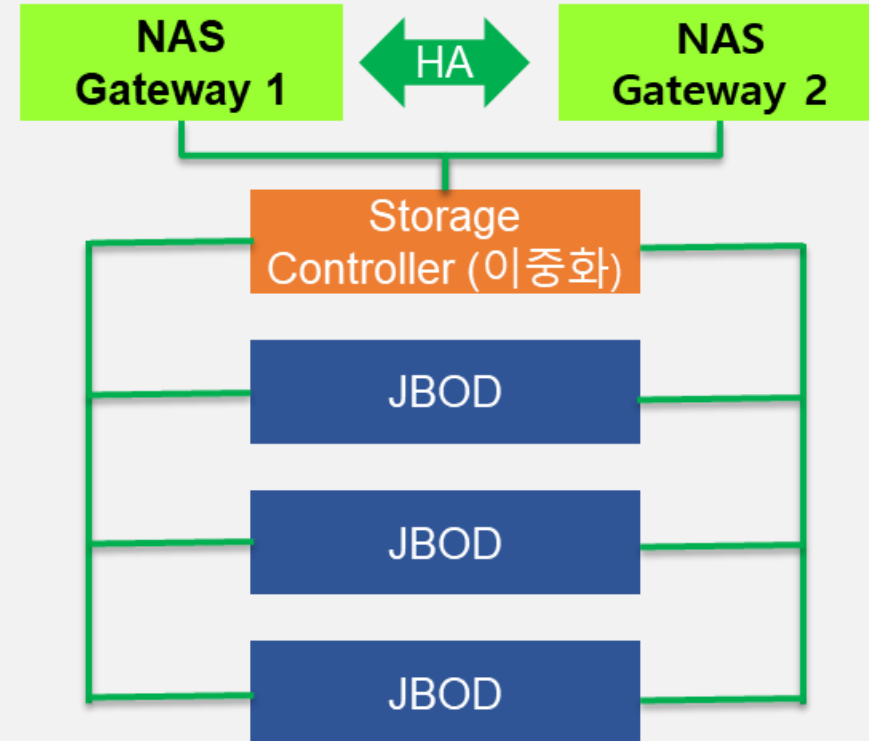
Traditional NAS architecture

IBM Spectrum NAS

구형 Architecture의 문제점 개선

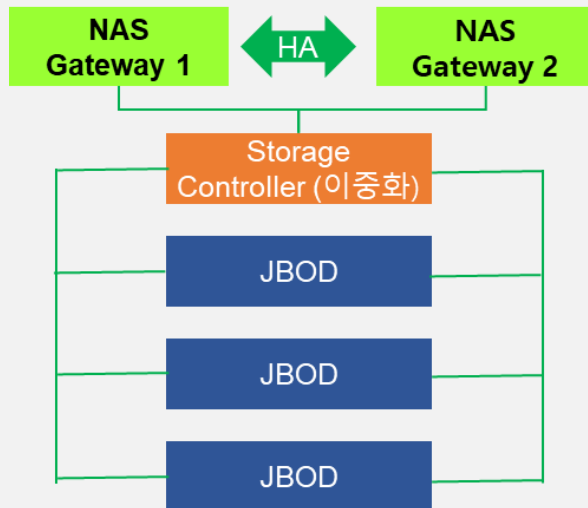
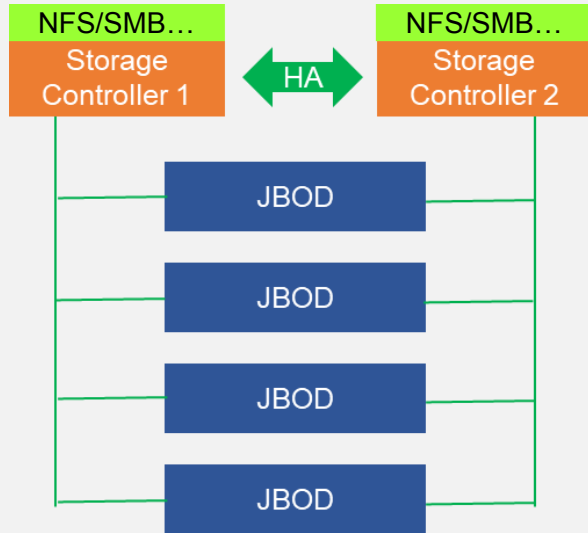


Storage Controller에서
직접 파일시스템 Sharing



SAN Storage + NAS
Gateway 조합

Traditional NAS의 문제점



- 대용량의 비정형 데이터 저장소 구축 시, 전통적인 Scale-up 방식의 Storage Architecture는 고객 요구환경을 만족시키지 못합니다.
 - ✓ Limited Scalability
 - ✓ Limited availability
 - ✓ Forklift Upgrades
 - ✓ Filer Sprawl
 - ✓ Reliability dependent on hardware
 - ✓ Limited self-healing
 - ✓ Limited flexibility in hardware configuration
 - ✓ High acquisition cost

비 정형 데이터 저장소 구축을 위한 새로운 전략

IBM Spectrum NAS

구형 Architecture의 문제점 개선

1. 비용을 최적화하고
2. 새로운 차원의 민첩성 및 유연성을 제공하고
3. 파일 데이터 증가에 원활하게 대처할 수 있는

**소프트웨어 정의 방식을
고려하십시오.**

● Solution 적용시...

- ✓ 파일 스토리지의 단순한 관리 및 할당
- ✓ 비즈니스 요구의 증가에 따라 용량, 성능 및 가능한 동시 접속 사용자 수를 자유롭게 증설 가능
- ✓ 저비용 상용 서버 (x86)를 사용하고 구성 및 제조사를 Mix 할 수 있는 유연성
- ✓ 업무 (어플리케이션)의 중요도에 따라 스토리지 성능, 데이터 보호 및 보관 정책을 정의/적용하는 계층적 데이터 관리 구축
- ✓ 표준 NFS, SMB 파일 Protocol에 대한 강력한 호환성 및 풍부한 데이터 관리 기능

- 유연한 성능 및 용량 확장성
- TCO 최적화
- 신뢰성, 정합성, 안전성 극대화
- Data 보호 및 가용성 극대화
- 운영 효율 및 관리 편의성
- 구형 Architecture의 문제점 개선

