



Hewlett Packard
Enterprise

BayNex

HPE-베이넥스 서버 비즈니스 파트너를 위한 솔루션 소개

2020년 10월 13일(화) 15:00 - 16:30

비정형데이터 저장을 위한 HPE Qumulo 스토리지 소개

한국HPE | 박재현 차장

데이터 진짜 중요한가요?

“데이터가 미래다” 외치는 국내·외 CEO



버지니아 로메티

IBM CEO

“앞으로는
데이터가
승자와 패자를
가를 것”

제프 베조스

아마존 CEO

“우리는 절대로
데이터를
내다 버리지
않는다”

정태영

현대카드·현대캐피탈 부회장

“현대카드는
이제 금융회사가
아니다.
'데이터 사이언스'
기업이다”

출처 : <https://brunch.co.kr/@bzup/370>



4차 산업혁명 시대의 핵심자산, 데이터!
데이터 활용능력이 곧 국가의 기업의 경쟁력입니다.

출처 : 과학기술정보통신부

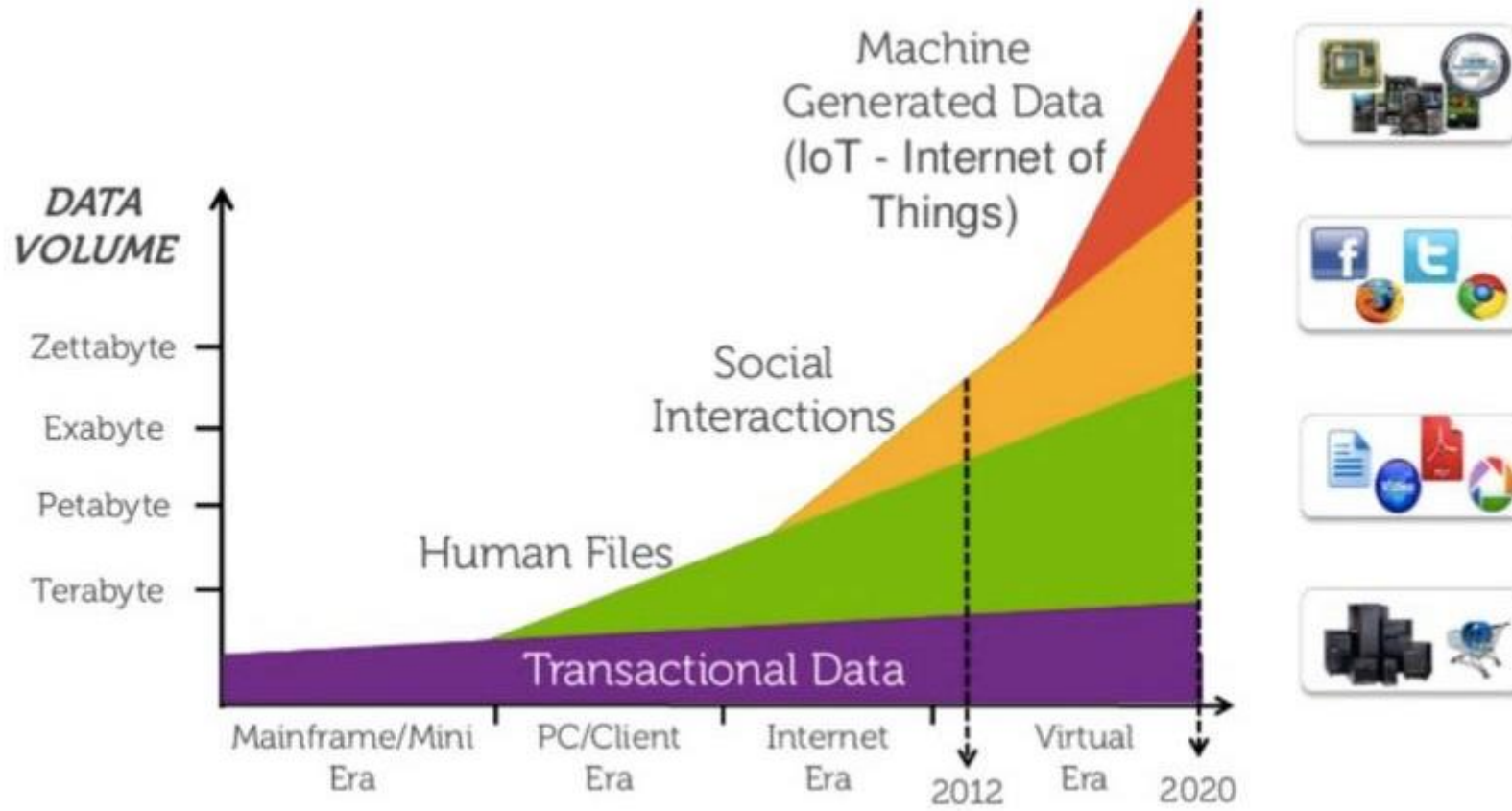
로고	기업	국가	시가총액 USD	시가총액 KRW
	애플 		19003 억달러	2259 조원
	사우디 아람코 		17572 억달러	2089 조원
	마이크로소프트 		16079 억달러	1912 조원
	아마존 		15865 억달러	1886 조원
	구글 		10175 억달러	1210 조원
	페이스북 		7647 억달러	909 조원
	알리바바 		7037 억달러	836 조원
	텐센트 		6521 억달러	775 조원
	버크셔 해서웨이 		5092 억달러	605 조원
	존슨앤존슨 		3912 억달러	465 조원

사용자 데이터, Machine 데이터, 기업 데이터를 수집

출처: 미스터캡 <http://www.mrktcap.com/>

HPE-베이넥스
서버 비즈니스 파트너를 위한 솔루션 소개

급증하는 데이터 유형은 무엇일까?



출처: <https://medium.com/@melodyucros/ladyboss-heres-why-you-should-study-big-data-721b04b8a0ca>

데이터 특성의 변화

정형 데이터

데이터 유형

고객 신상 데이터
매출 데이터
재고 데이터
회계 데이터

크기

MB, GB, TB

속도

활용주기 : 수시간 ~
수주

비정형 데이터

동영상, 음악, 사진
메시지, 소셜미디어 등

PB, EB, ZB

분, 초 단위로 활용

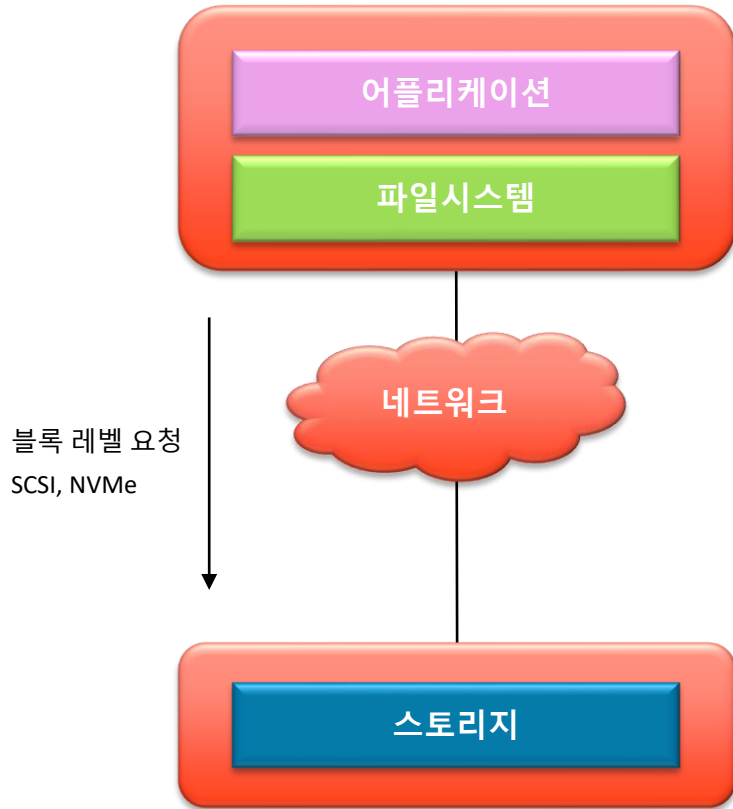
A promotional image for the movie Avatar. It features Neytiri on the left and Jake Sully on the right, both looking forward. The background is a lush, bioluminescent forest with a large moon in the sky. At the bottom, there are flames and a burning structure. The text '~ 10 PB' is overlaid in the center in a large, white, sans-serif font.

~ 10 PB

AVATAR

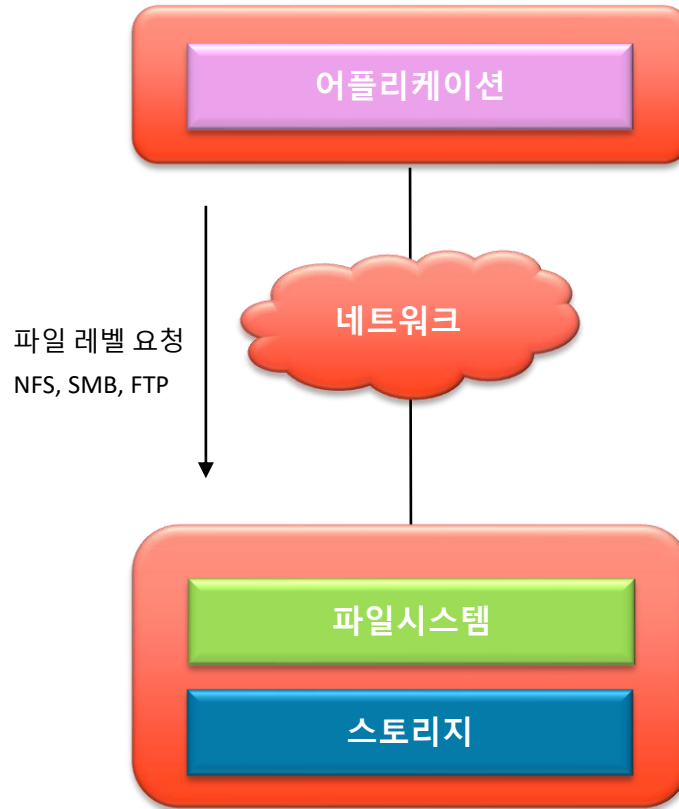
급증하는 비정형 데이터 어디에 저장해야 할까요?

블록



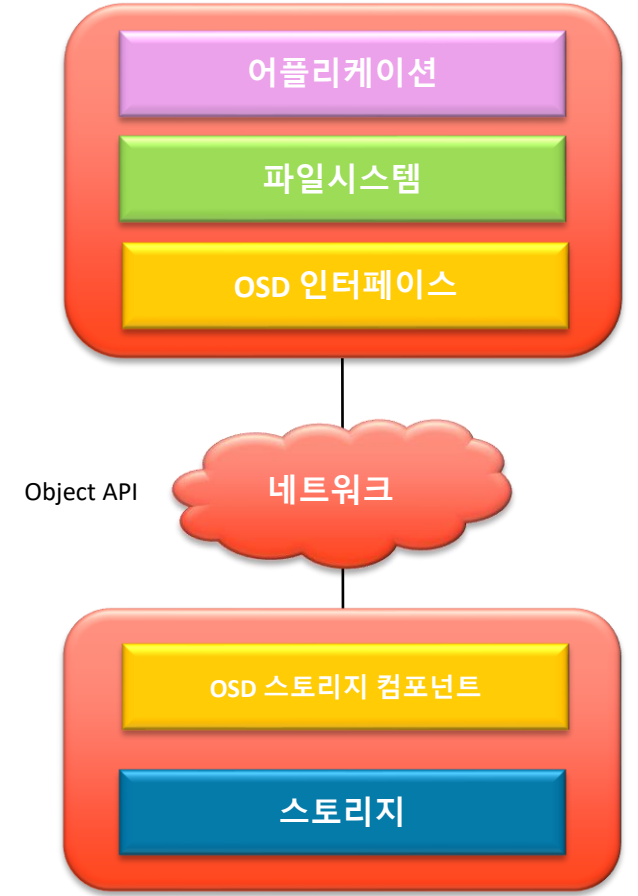
고성능, 빠른 응답 속도.
하지만 스토리지 용량 할당이 복잡.

파일

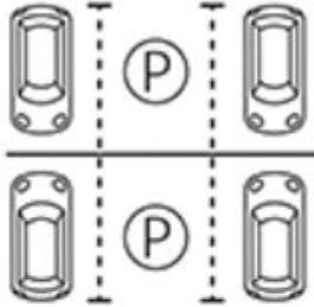


저렴하고 쉬운 구성이지만
파일 개수가 많아질 수록 성능 저하.

오브젝트

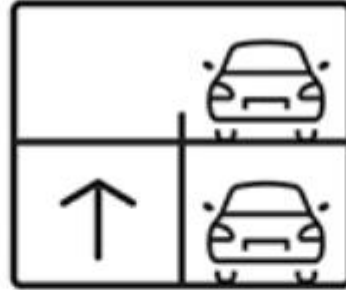


대용량, 데이터 변경작업시
처리가 느리고, 어디에 데이터가
저장이 되어 있는 파악이 어려움.



블록 스토리지

'주차장'에 비유 - 엄격하게 정의된 블록에 데이터가 저장됨 - 특정 '공간' 위치를 통한 액세스



파일 스토리지

'주차장 건물'에 비유 - 데이터가 계층형 레벨로 정리됨 - 경로를 역추적해서 액세스



오브젝트 스토리지

대리 주차'에 비유 - 스토리지 세부 정보에 대해 걱정할 필요 없음 - 데이터 저장 및 액세스가 용이함

출처: <https://medium.com>

Qumulo는?

- 블록의 장점 : 블록단위의 저장 메커니즘, 4K 단위 (bStore, pStore)
- 파일의 장점 : 여러 Client가 동시 접속
- 오브젝트의 장점 : Single Name Space, 대용량, 분산저장

3가지 스토리지의 장점을 결합한

Scale-out 파일 스토리지

The most important advantage of scale-out NAS over object storage is that all your data is available in a single tier. Applications are built to talk to file systems. You don't have to re-write things to talk S3, for example. **Modern file systems can do what object storage does, but with performance.**

- Computerweekly

Gartner Magic Quadrant

설립 목적 : 하이브리드 클라우드 파일 스토리지

Isilon Core 개발자 18명이 나와 설립

가트너 매직쿼드런트 (분산파일 시스템과 오브젝트 스토리지 분야) 2018년도에 리더그룹으로 2년만에 수직 상승

Gartner's peer insight 4.9점 1등
(D사 4.3점, N사 4.6점)

Gartner's peer insight 100% 제품 추천 제품
(D사 72%, N사 80%)



Source: Gartner (September 2019)

As of August 2019 © Gartner, Inc

Scale-out NAS 기존 NAS와는 무엇 이 다른가요?

현대적인 Scale-out File Storage는 무엇이 다를까요?

Legacy file Storage

Scale-out NAS (PB)

검증 되지 못한 조합의
x86 Architecture

업계에서 검증된 **No.1** HPE Big-data Server Platform : Apollo 4200

Tree walk (Directory 구조)
100% 배제 못함

Tree walk (Directory 구조) 100% 배제
- 압도적 Scale, 실시간 Storage Analytics

Scale에 맞지 않는 Protection

Block Base의 효율적이고 빠른 Rebuild 시간을 가지는 Data 보호방식
- 100% Erasure Coding

100% 사용하지 못하는 비효율 성

“성능 저하 없이” 수십 PB, 수 백억 개 단위의 file 저장 지원

느리고 어려운 구식 고객 지원 방식

현대적이고 **즉각적**, 혁신적인 고객 경험 - 24x7 즉각 지원 가능한 Channel

실시간 관리 및 모니터링 불가,
사용량 예측도 어려움

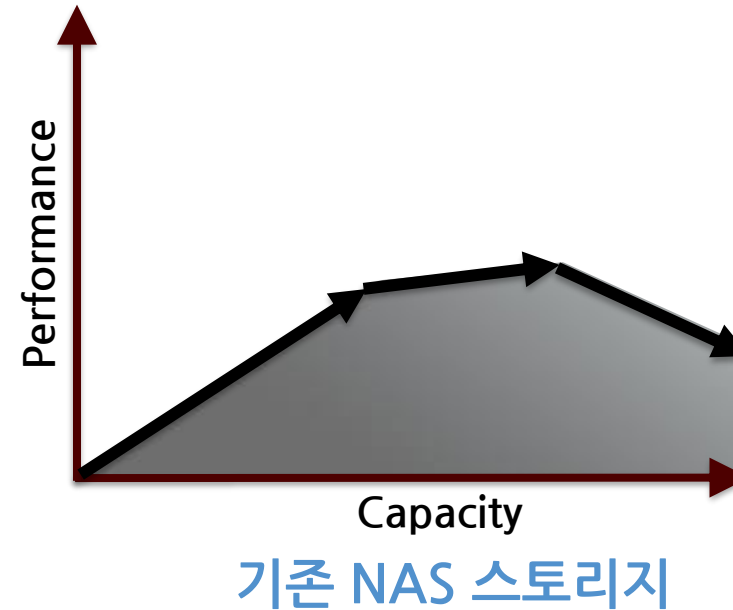
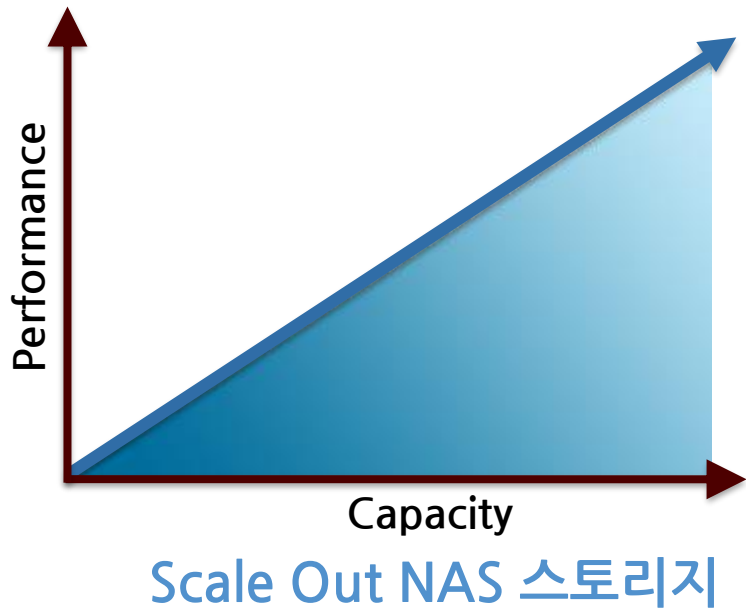
낭비가 적고 Data를 저장하는 것 외에 **관리, 예측이 가능**
- 혁신적인 실시간 Analytics

Feature 마다 구매 해야 하는
성가시고 비싼 License

All-inclusive License 정책 : 모든 기능이 이미 들어 있습니다

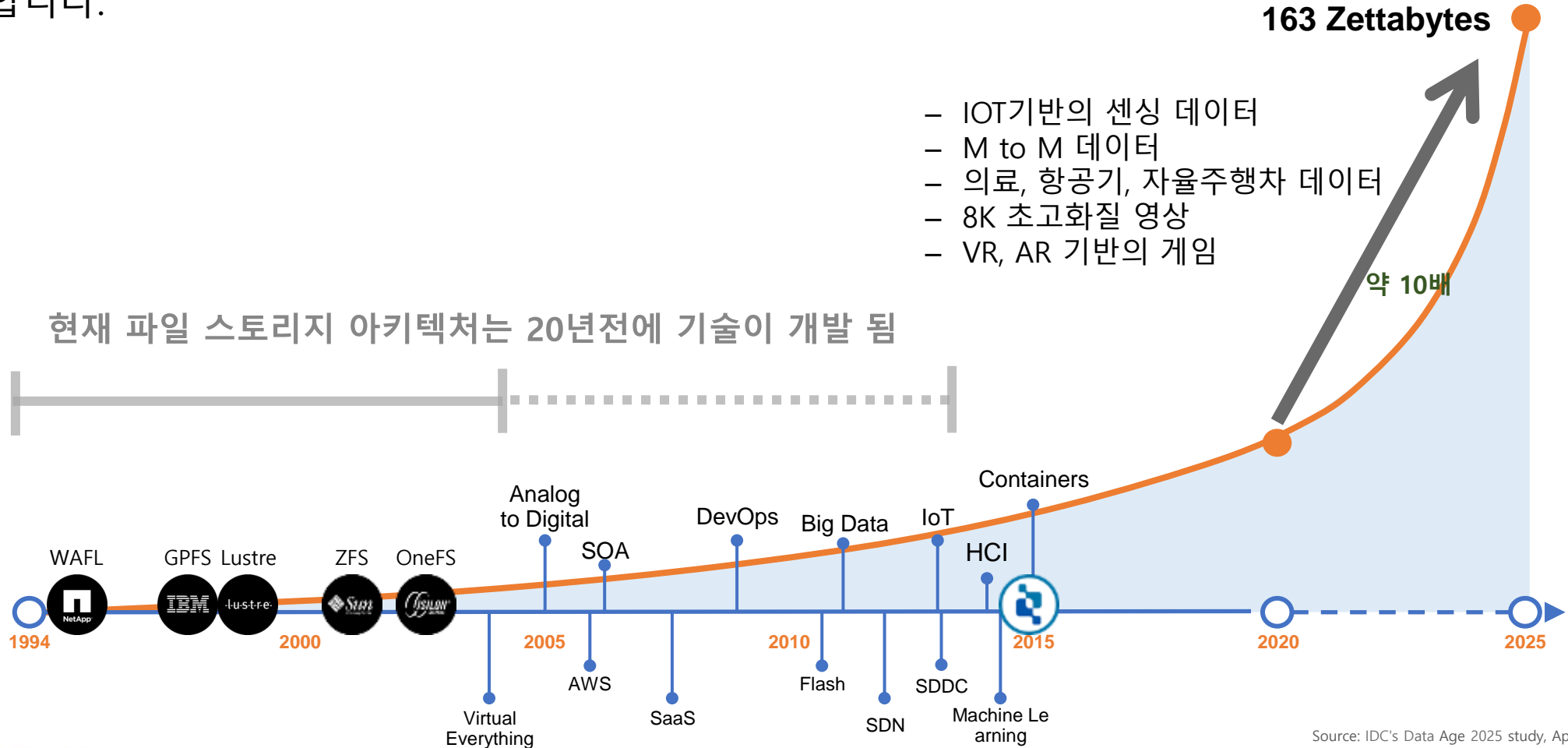
Scale Out NAS와 기존 NAS 비교

기존의 NAS는 이중화된 컨트롤러에서 다수의 디스크와 파일 시스템을 관리하여 용량과 성능의 한계가 있음
Scale Out NAS는 디스크가 장착된 컨트롤러를 추가하는 구조로 기존 NAS의 단점을 해결



PB급 스토리지를 지원하기 위해 개발된 아키텍처

큐물로는 PB급 이상의 데이터 저장 공간을 제공하기 위해서 탄생한 파일 시스템인 반면 기존 NAS는 모두 1990년대 혹은 2000년대 초반에 개발된 파일 시스템을 사용합니다. 따라서 5PB이상에서도 훌륭한 성능을 발휘합니다.



Source: IDC's Data Age 2025 study, April 2018

Scale-out NAS 도입시 중요한 건...

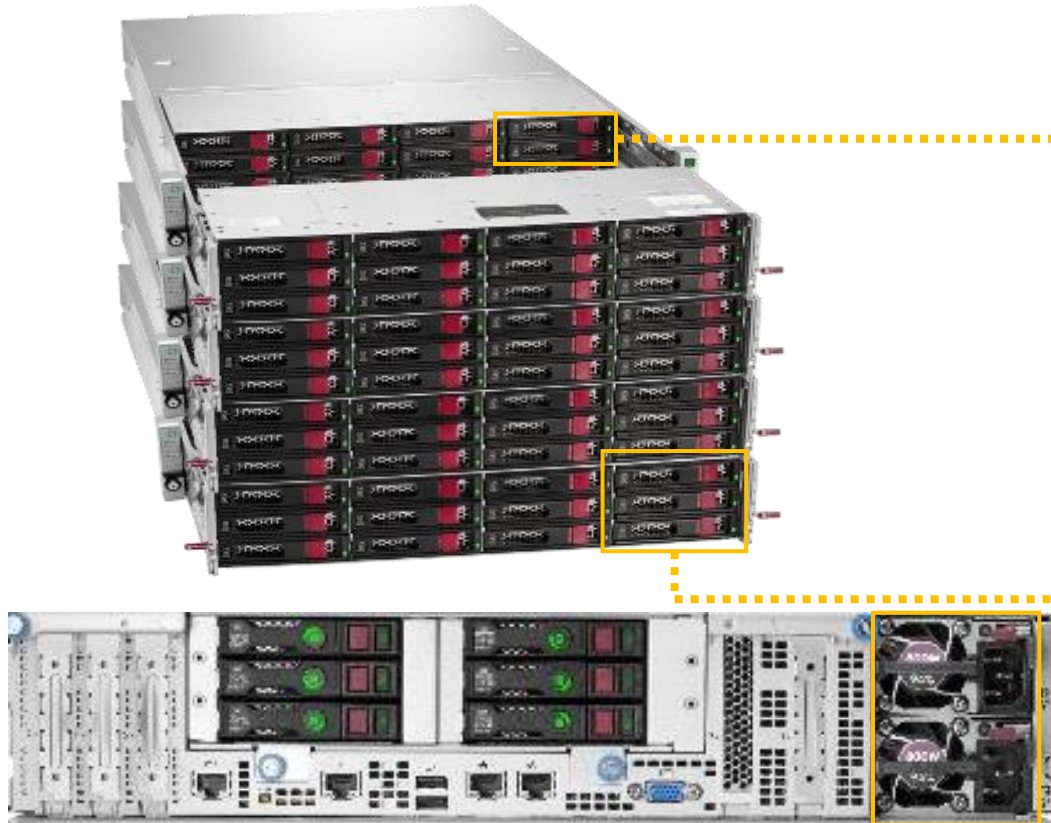
Scale-out NAS 도입 시 중요한 point 6가지

1. 이중화를 고려한 하드웨어 아키텍처
2. 개별 디스크를 교체 가능
3. PB급 스토리지를 지원하기 위해 개발된 아키텍처 : Btree
4. PB급 스토리지에서도 실시간 분석이 가능한 관리 툴
5. 데이터 효율성
6. 단일화된 라이선스 정책

1. 이중화를 고려한 하드웨어 아키텍처

HPE Qumulo 솔루션은 Scale-out NAS에 특화된 Apollo 4200 기반인 반면 타사는 구형 CPU를 사용하거나 Power Supply를 여러 노드가 공유하는 구형 하드웨어 아키텍처입니다.

Apollo4200 Gen10 Server



업계 1위 HPE ProLiant 서버 기반

증명된 최고의 Enterprise Server
최고의 안정성 / 고집적 효율성

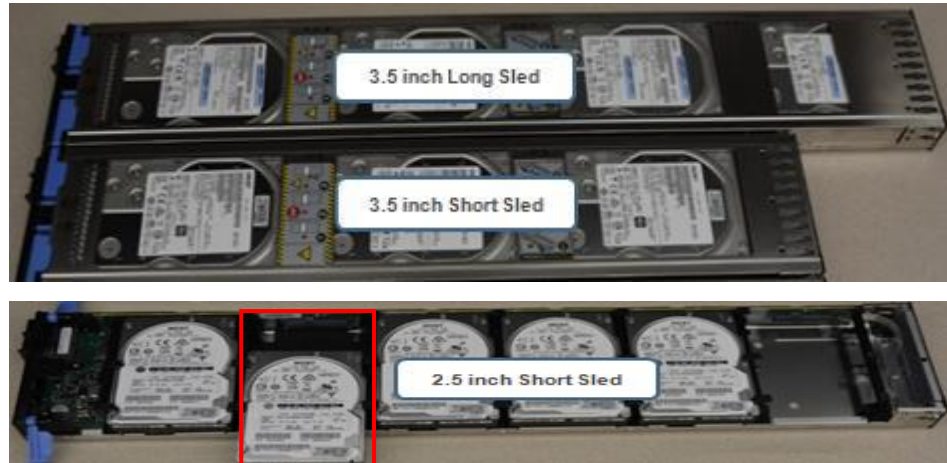
최신 Intel CPU 탑재

모든 파트의 이중화 구성

뛰어난 유지보수

2. 개별 디스크를 교체 가능

HPE Qumulo 솔루션은 모든 디스크를 개별 교체할 수 있지만 타사 일부 제품은 하나의 디스크 교체도 여러 개의 디스크가 Shut Down되어 다수의 운영되는 서비스가 영향을 받을 수 있습니다.

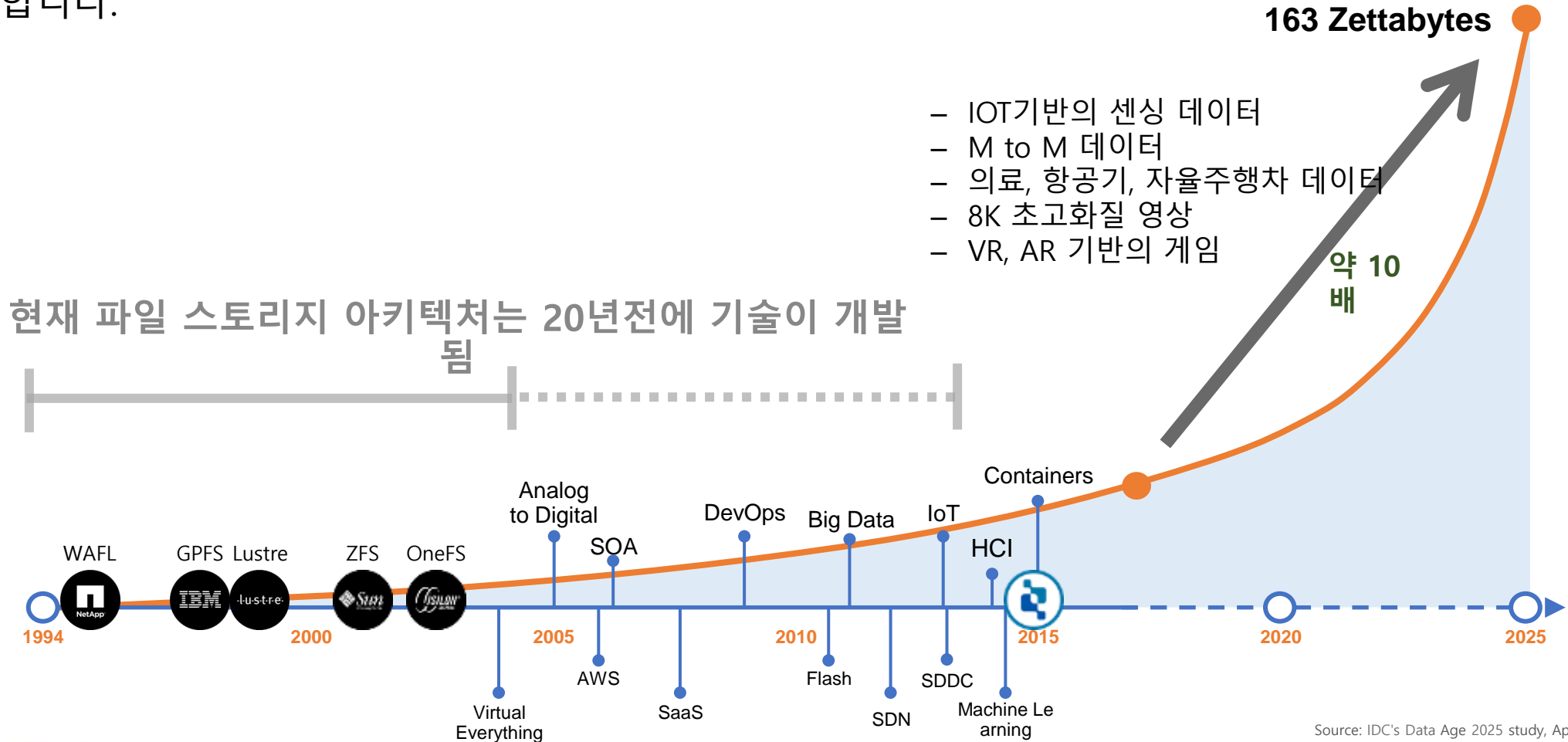


디스크

- Disk 한 개 교체 하려면 SLED 전체 offline
- 서비스 Risk가 높음

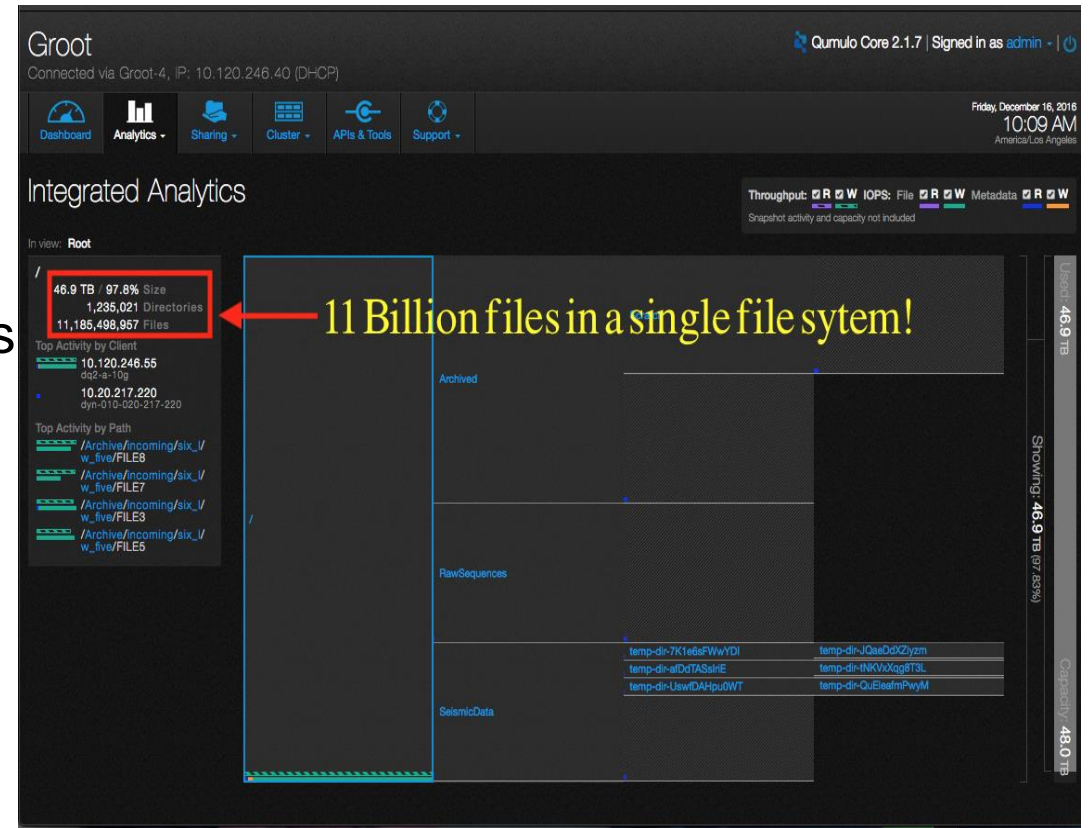
3. PB급 스토리지를 지원하기 위해 개발된 아키텍처

큐물로는 PB급 이상의 데이터 저장 공간을 제공하기 위해서 탄생한 파일 시스템인 반면 기존 NAS는 모두 1990년대 혹은 2000년대 초반에 개발된 파일 시스템을 사용합니다. 따라서 5PB이상에서도 훌륭한 성능을 발휘합니다.



3. PB급 스토리지를 지원하기 위해 개발된 아키텍처

- 최대 SMB 디렉토리명 길이 : 260 characters
- 최대 NFS 디렉토리명 길이 : 32,760 characters
- 최대 파일명 길이 : 255 (protocol limited) characters
- 단일 디렉토리 내 최대 파일 개수 : 42억 5천개
- 단일 최대 파일 크기 : 9 exabytes
- 최대 저장 파일 개수 : 1800경



4. PB급 스토리지에서도 실시간 분석이 가능한 아키텍처

HPE Qumulo솔루션은 **Btree** 구조로 데이터 실시간 분석을 하실 수 있지만 기존 NAS제품은 **Tree walk**구조로 실시간 데이터 분석이 불가능합니다.

Btree (우편번호)



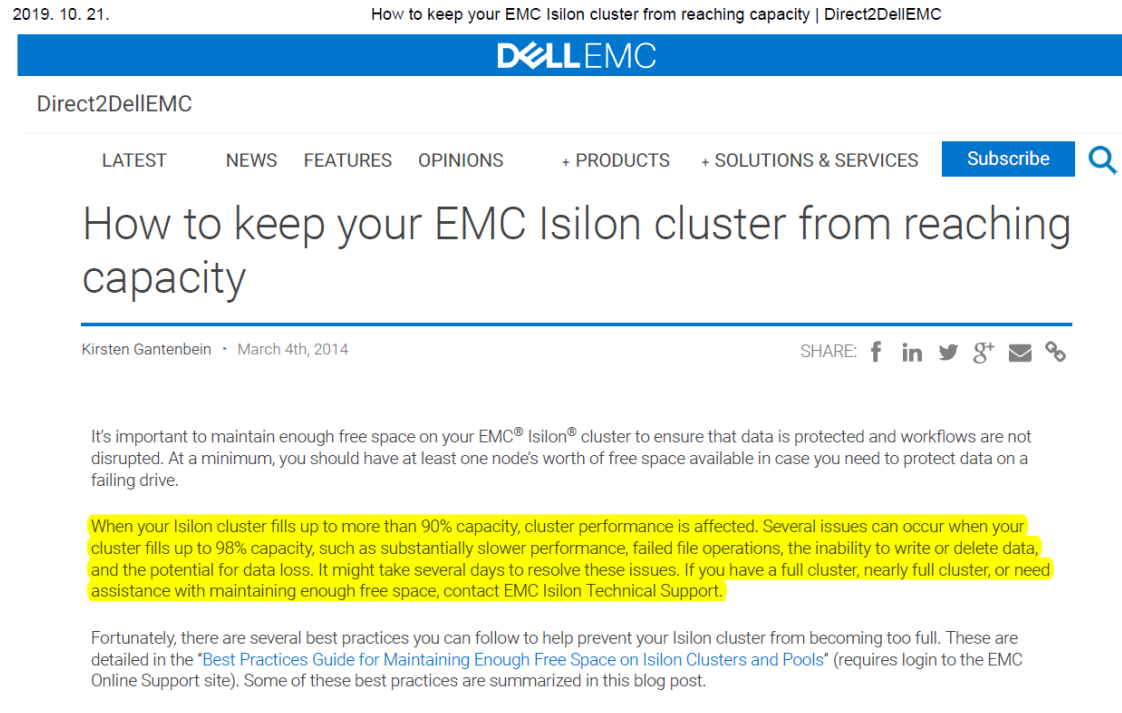
Tree walk

C: 내컴퓨터 > Users > park > OneDrive - Hewlett Packard Enterprise > Documents > 발표자료 > HPE Qumulo > HPE NAS시장 공략 전략

우리은하 > 태양계 > 지구 > 대한민국 > 서울시 > 영등포구 > 여의도동 > SK증권빌딩

5. 데이터 효율성

HPE Qumulo 솔루션은 Usable 용량의 100%까지 데이터 저장을 위해 사용할 수 있으나 기존 NAS 제품은 최대 85% 정도까지 사용이 가능하며, 90% 이상이 되면 언제 Full이 발생할지 몰라 위험해 질 수 있습니다.



경쟁사 공식 Blog 문서에서 발췌

전체 용량의 90% 이상을 사용하면 성능에 영향을 받음
(Field 에서는 85%를 권고)

98% 이상을 사용 하면 **성능 저하, File Operation Failed, Potential Data Loss** 까지 경고 함

해결을 하려면 몇일이 필요하며 DELL EMC 기술 지원이 별도로 필요함

이러한 상황을 피하는 방법은 모니터링 밖에 없음

전체 Storage 가용량이 예측 불가능하고 가변적이라 더욱 위험함

<https://blog.dell EMC.com/en-us/keep-emc-isilon-cluster-reaching-capacity/>

5. 데이터 효율성

같은 데이터를 저장했다고 다 같은 용량을 차지 하지 않습니다.

The image shows two side-by-side Windows File Properties dialog boxes. The left window is titled 'HPE Qumulo' and shows a file size of 9.33 TB (10,265,335,203,060 bytes) and a size on disk of 9.49 TB (10,438,318,596,096 bytes). The right window is titled '경쟁사' (Competitor) and shows the same file size but a much larger size on disk of 33.2 TB (36,605,997,875,200 bytes). The 'Size on disk' values are highlighted with red boxes. Below each window, the Korean words '고효율' (High Efficiency) and '저효율' (Low Efficiency) are written in large, stylized characters.

System	Files/Folders	Type	Location	Size	Size on disk	Efficiency
HPE Qumulo	29,515,930 Files, 119,018 Folders	All of type File folder	All in	9.33 TB (10,265,335,203,060 bytes)	9.49 TB (10,438,318,596,096 bytes)	고효율
경쟁사	29,515,928 Files, 119,018 Folders	All of type File folder	All in	9.33 TB (10,265,334,363,380 bytes)	33.2 TB (36,605,997,875,200 bytes)	저효율

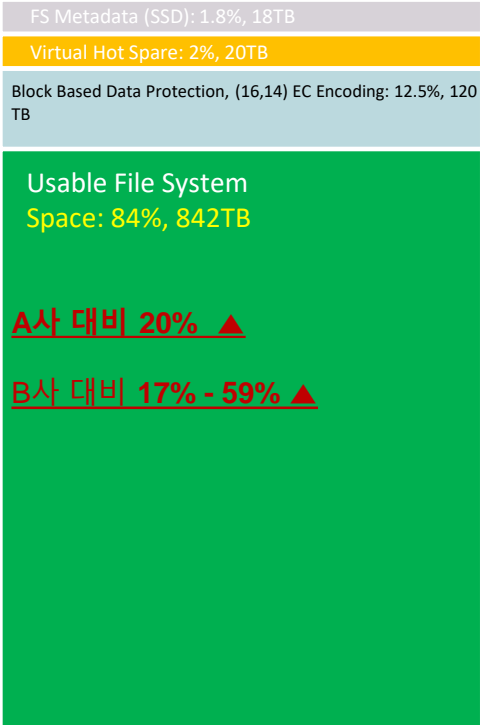
Size on disk:
9.49 TB

고효율

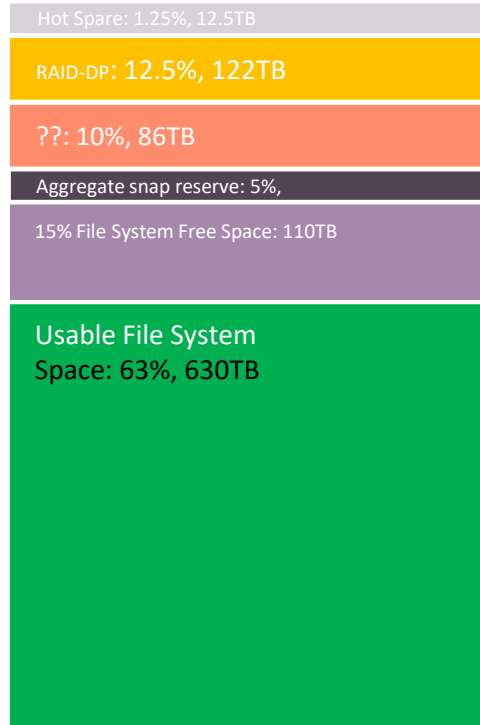
Size on disk:
33.2 TB

저효율

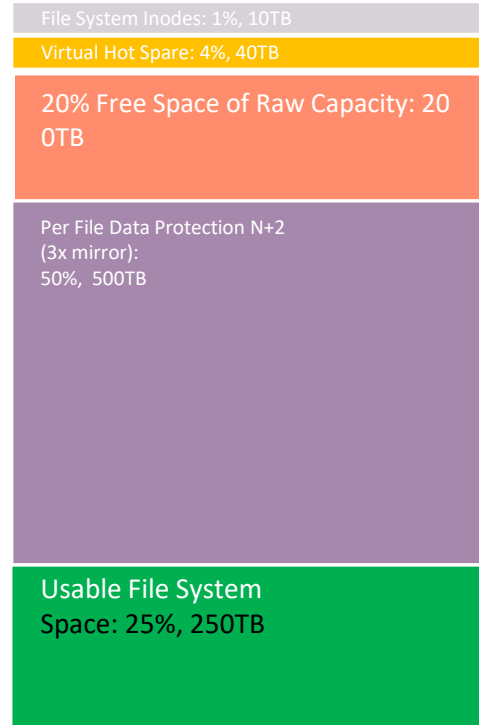
**Qumulo, 1PB raw
84% usable**



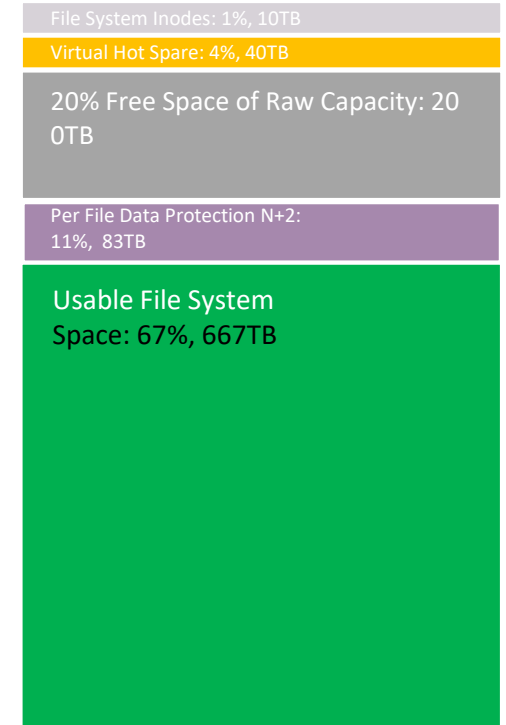
**A사, 1PB raw
63% usable**



**B사, 1PB raw
25% usable (2MB 이하 파일)**



**B사, 1PB raw
25% usable (2MB 이상 저장)**



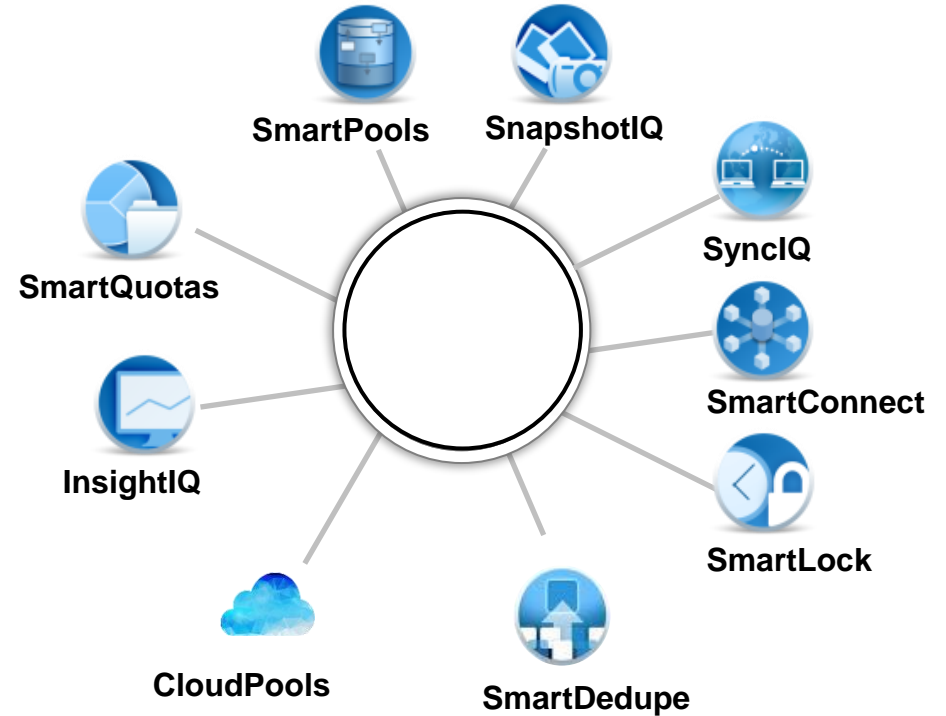
구 분	HPE Qumulo	A사	B사 (2MB 이하)	B사 (2MB 이상)
용량 효율성	84%	63%	25%	67%
데이터 보호 방식	Erasure Coding	RAID	3벌 복제**	Erasure Coding**
권고 최대 사용율*	100%	75%	80%	80%

6. 라이선스 정책

HPE Qumulo 솔루션은 All-inclusive 라이선스 정책인 반면, 기존 NAS 제품은 여러 기능을 쓰기 위해서 비용을 따로 받고 있습니다.



VS



Qumulo는 모든 기능에 대해 License 추가 비용이 들지 않는 All-inclusive license 입니다.

위 나열 된 모든 Feature들은 모두 별도 구매가 필요한 Usebased license 입니다.

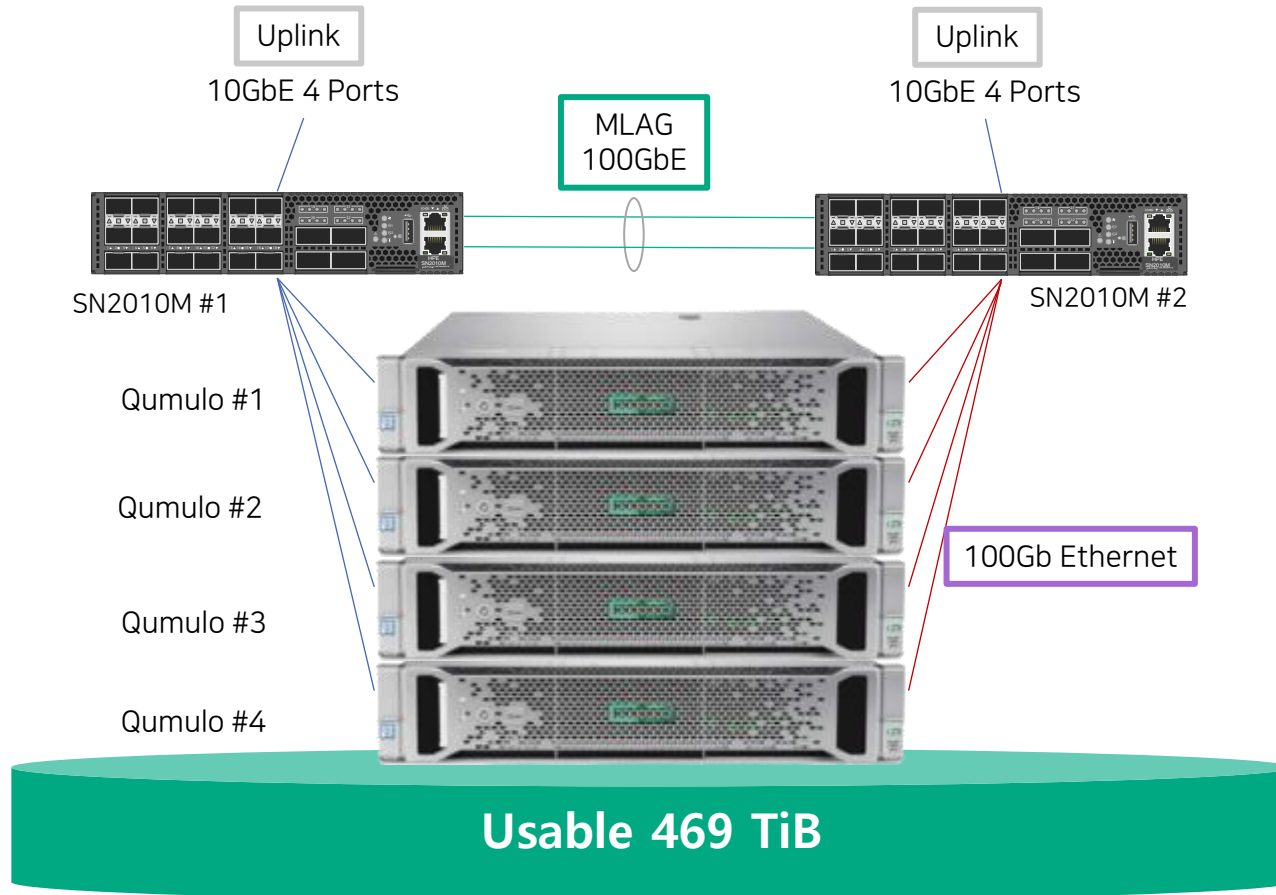
Scale-out NAS 도입시 중요한 건?

1. 이중화를 고려한 하드웨어 아키텍처
2. 개별 디스크를 교체 가능
3. PB급 스토리지를 지원하기 위해 개발된 아키텍처 : Btree
4. PB급 스토리지에서도 실시간 분석이 가능한 관리 툴
5. 데이터 효율성
6. 단일화된 라이선스 정책

HPE Qumulo 어떻게 구성

HPE Qumulo Gen10 192TB 구성 예시

- 192TB 모델의 구성도 예시입니다.



[대당 사양 : 1/2U]

18 x 10/25Gb ports

4 x 100Gb ports

* 옵션 사항으로 고객 스위치 연결 가능

[대당 사양 : 2U]

2 x Intel Xeon 2.2GHz 10-core

128GB Memory

6 x 1.92TB SSD

24 x 8TB SATA HDD



100Gb 2-port

Single Name Space

HPE Qumulo Worldwide best Scale Out 파일 스토리지

- 사양

All Flash NEW

구분	336TB	192TB	90TB	36TB	34TB	145TB
컨트롤러 CPU	1 x Inte Xeon 10core 2.2GHz	2 x Inte Xeon 10-core 2.2GHz	1 x Intel Xeon 10-core 2.2GHz	1 x Intel Xeon 10-core 2.2GHz	1 x AMD EPYC 24 cores 2.8Ghz	1 x AMD EPYC 24 cores 2.8Ghz
Cache Memory	128GB	128GB	64GB	64GB	128GB	128GB
SSD Cache	4 x 1.92TB SSD	6 x 0.96TB SSD	3 x 0.96TB SSD	3 x 480GB SSD	9 x 3.84TB NVMe + Boot Drive	19 x 7.68TB NVMe + Boot Drive
HDDs	24 x 14TB 7.2K HDD	24 x 8TB 7.2K HDD	9 x 10TB 7.2K HDD	9 x 4TB 7.2K HDD	N/A	N/A
Network	10/25Gb 2-port	100Gb 2-port	10/25Gb 2-port	10/25Gb 2-port	100Gb 2-port	100Gb 2-port
Management	iLO5 1Gb-T Management Port				iLO5 1Gb-T Management Port	
Power and Cooling	 Redundant PSU, 10 Fans			 Redundant PSU, 6 Fans		
Form Factor	2U rack mount				1U rack mount	

HPE Qumulo 어떤 업무에 활용할 수 있나요?

HPE Qumulo 적용 업무

이미지 가공/저장



지형 이미지
GIS(지리정보시스템),
위성정보, LIDAR

보안



고화질 CCTV,
바디 캠, 영상분석

고성능 컴퓨팅(HPC)



Simulation, Modelling,
HPC, Active datasets

Media &
Entertainment



VFX, Rendering, Editing,
Motion Animating,
Animation, Transcoding

AI / Machine Learning



Modelling, Simulations,
active archives,
Data Lake for AI

의료



PACS, Digital 병리학,
영상의학과(영상분석)

교육



Digital 도서관, HPC Labs,
GIS

순수 과학 &
의학 연구



수학계산, 현미경 Image,
유전체 분석

Enterprise File Storage



User Directory, Archives,
D2D backup,
ECM(전사 문서중앙화)

Commercial IoT



DFS, ADAS(자율주행/보조),
IoT Sensor Data(Splunk)