FUTUREWEI
Technologies

Primary Storage
Insight Report

V1.0	March 2022

FUTUREWEI® TECHNOLOGIES, INC.

Boston Research Center

Address: 111 Speen Street, Suite 114

Framingham, MA 01701

United States of America

Website: http://www.futurewei.com/

Contents

1 EXECUTIVE SUMMARY

Primary storage market will continue to grow at a favorable rate in the following years, although there are continuous competitions from server storage as well as public cloud. Slower growth rate (even negative growth rate) may occur in certain categories in primary storage products. It reflects adoptions of new data center strategies and new workload introductions such as AI/Machine Learning, big data and other workloads powered by 5G and other emerging technologies.

To continuously generate growth, fend off fierce competition from public clouds and server storage and provide more values to customers, primary storage vendors shall focus on several initiatives. Of all the initiatives, one of the key initiatives potentially could be data/application platform integration.

Primary storage market is becoming a red ocean market with increasingly competitions from cloud and among existing vendors. Profits and growth are slowing down. Storage vendors need to expand market space with new value creation. In today and future market, data and application has far more reaching impacts to customer's business than underlying infrastructure. Application and data platform has controlled data access point resulting in decoupling of infrastructure underlying and applications that use them. Building an integrated application platform by leveraging hardware and underlying infrastructure capabilities could be one of the key competences for primary storage products. Application platform shall leverage hardware and enterprise capabilities provided by primary storage to offer differentiation among other vendors with full stack optimization.

With popularity of data mesh methodology, more and more enterprises have decentralized their data, pushing data storage, models, and management into different business units. [1] Whoever controls data entry point will control entire customer data set and therefore, treat underlying infrastructure just as resources from edge, data center and cloud. How does primary storage vendor become major contributor to this effort instead of bystander? It will be the key for the next few years to define the role primary storage is going to play in the data center.

Vendors may often tout their hardware prowess to differentiate themselves from their competitors. Software or platform capabilities, so called "Soft Power", that could also distance vendor from their competitors. A start-up company may rely on "back technology" to differentiate itself from competitors and incumbents. Growing software capabilities and broader ecosystem supports to meet customer requirements with hardware differentiation as core competence can increase customer stickiness and expand business. They are the driving forces for business grow. It is especially true for sectors that are much mature, for example storage industry.

We expect green initiative and carbon emission conscious data center solution will be more important going forward. Countries are implementing more and more related regulations that have bigger impact on enterprises. Carbon emission

regulation conforming may become a barrier for enterprises to enter certain IT markets in the future.

Some primary storage vendors, NetApp® as leading example, have started to transition themselves into public cloud storage companies. It is a bold move. To hedge the risk, we are expecting those vendors will continue to invest their bread and butter – on prem primary storage market.

Below are listed some major points (some points could be related) deserving considerable attentions from primary storage vendors:

1. Fully embrace application as data and data/application platform integration in primary storage and take full control of data/application entrances.
2. Position primary storage product to provide modern data center infrastructure services including compute, storage, and other services.
3. Extensive ecosystem support by leveraging application and data platform integration. (Transform storage eco system problems into platform eco system problems)
4. Provide more application and data centric semantics to be integrated and managed by data/application platform.
5. Highly distributed architecture with modular design and managed with data/application platform and distributed OS.
6. Provide extensive support for PayU and STaaS including hardware agile deployment support and software supports
7. Extensive data sovereignty and security support to enable customer digital ownership
8. More energy efficient with carbon footprint accounting
9. Hardware specialization to accommodate dynamic workload and workload consolidation
10. Further advance AIOps and intelligent management to help customer reduce TCO

2 MARKET TREND

2.1 OVERVIEW

According to IDC, external enterprise storage system market is expected to grow in the following four years with a moderate rate typical for mature markets. [2]

2020–2025 Revenue ($B) with Growth (%)

Selected Segment Growth Rate

▲ All-flash arrays CAGR 8.4%
▲ Hybrid flash arrays CAGR 5.1%
▼ HDD only CAGR -5.8%

Total Market CAGR

4.7%
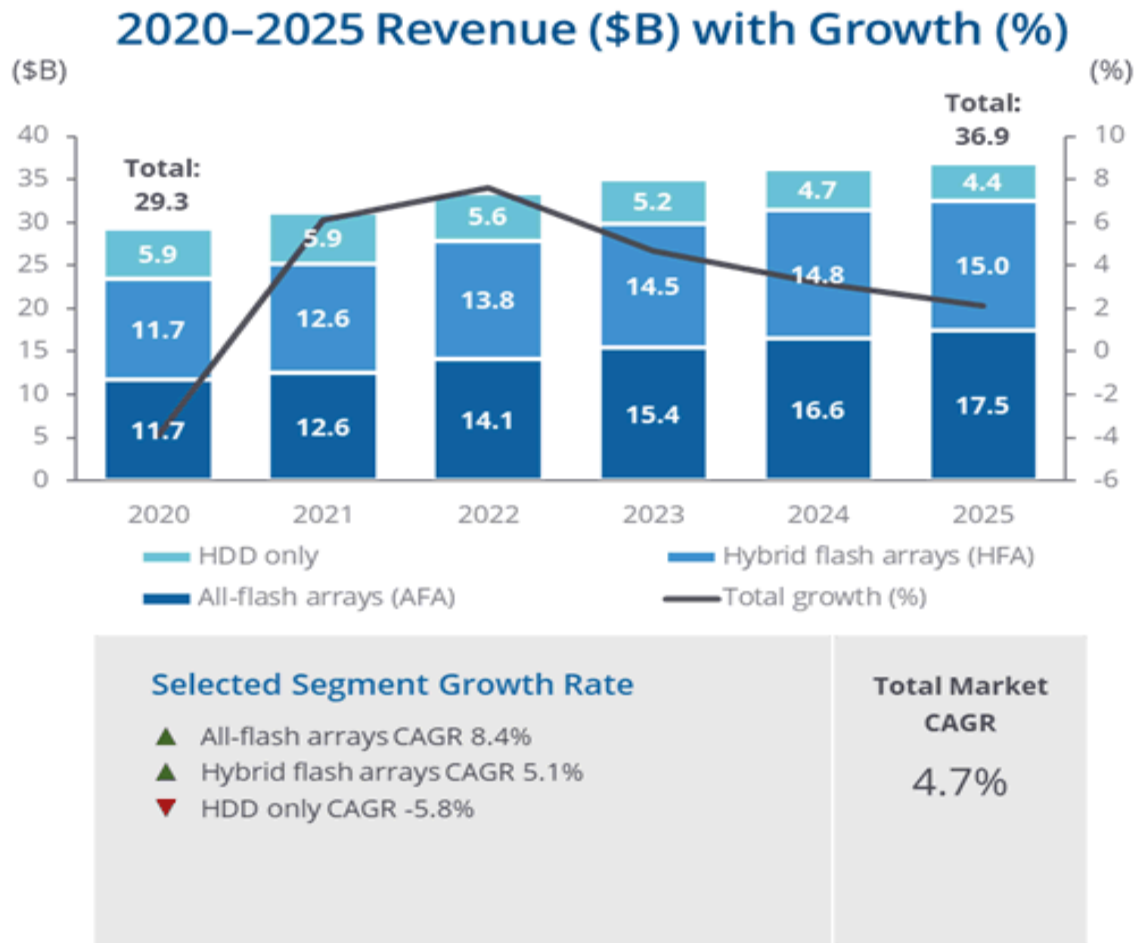
However, not all sectors or products are expected to grow equally.
Virtual SAN/HCI is expecting to grow a much faster rate 42.8% than others.

Worldwide External Enterprise Storage Systems Capacity Shipped by Topology, Installation, and Data Access, 2016-2025 (PB)

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2020–2025 CAGR (%)
Direct (DAS)	3,890	4,095	4,303	4,408	4,539	5,914	5,300	5,514	5,666	5,760	4.9
Mainframe networked (FICON)	1,199	881	1,042	1,248	1,126	1,134	1,024	943	975	1,072	-1.0
Open networked											
Ethernet file/object	16,307	17,534	21,183	21,460	24,170	29,844	33,234	40,152	46,302	52,417	16.7
SAN											
Fibre Channel	18,643	19,334	23,935	22,217	22,343	25,762	28,409	36,489	42,427	48,003	16.5
Virtual SAN	1,056	1,840	3,525	5,450	7,413	12,564	20,985	27,416	40,795	44,043	42.8
Ethernet block	8,995	10,752	10,899	12,103	12,734	14,421	14,702	20,284	25,045	30,309	18.9
Other	1,709	1,721	2,507	5,239	4,947	6,016	6,168	6,498	6,668	6,741	6.4
Subtotal SAN	30,402	33,647	40,867	45,009	47,436	58,764	70,265	90,687	114,935	129,095	22.2
Subtotal open networked	46,709	51,181	62,050	66,469	71,606	88,607	103,499	130,839	161,236	181,512	20.4
Total	51,798	56,157	67,395	72,126	77,271	95,656	109,823	137,297	167,877	188,345	19.5

*Virtual SAN includes storage capacity of systems sold as HCI appliances.

Source: IDC, 2021

In term of storage media type shipped by storage systems, we will see a trend of HDD sales are decelerating to a point that it is not significant market for big storage vendor anymore.

Worldwide External Enterprise Storage Systems Revenue by Storage Array Type, 2016-2025 ($B)

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2020–2025 CAGR (%)
All-flash arrays (AFA)	5.3	6.9	10.5	11.7	11.7	12.6	14.1	15.4	16.6	17.5	8.4
Hybrid flash arrays (HFA)	10.0	10.4	12.0	12.2	11.7	12.6	13.8	14.5	14.8	15.0	5.1
HDD only	9.8	8.3	7.4	6.7	5.9	5.9	5.6	5.2	4.7	4.4	-5.8
Total	25.2	25.6	29.8	30.5	29.3	31.1	33.5	35.0	36.1	36.9	4.7

Note: The value for HFA includes the full value of systems shipped with flash and HDD storage media.

Source: IDC, 2021

Worldwide external enterprise storage sales are still dominated by 3 big markets: United States, China, and Europe, with Chinese market having much fastest growth rate 12.3. Due to stagnant growth rate projects in US markets, vendors will and already have shift

their focus on European markets and other faster growth markets like Latin America and APJ market. We are going to see even fiercer competitions on those market fronts.

There are several main driving forces for revenue growth:
- All flash array with HCI
- Delivery model (consumption-based service model)

2.2   FAST GROWING HCI MARKET

HCI market is expected by most of market research firms to have best CAGR in the next few years. IDC projects 11.6%, Meticulous research gives 24.1 and Markets and Markets predict around 28%.

HCI products are also going through some architecture changes from CI architecture to HCI, DHCI and Composable HCI. [3]

Traditional vs. HCI vs. disaggregated HCI vs. composable infrastructures

	Traditional	HCI	dHCI	Composable
SCALING METHOD	Scale up or out— depends on vendor	Linear—all resources in one appliance	Compute and storage scale separately	Compute and storage scale separately
EASE OF MANAGEMENT	Low	High	High	High
WORKLOADS SUPPORTED	Virtual, containers, bare-metal	Virtual, containers	Virtual, containers	Virtual, containers, bare-metal
LEVEL OF AUTOMATION VIA API	Highly variable	Varies by vendor	Varies by vendor	High
APPLICATIONS SUPPORTED	Any	Mainstream, with some outliers	Mainstream, with some outliers	Any

SOURCE: SCOTT LOWE   ©2020 TECHTARGET. ALL RIGHTS RESERVED TechTarget

1. Management complexity (3 tier) TCO
2. dHCI independent storage and compute
3. GPU/DPU expensive need to repurpose and dynamic configuration, if commodity

Some of the major driving forces for those architecture changes include reducing management complexity and TCO, storage and compute decoupling and dynamic configuration, adaption to dynamic workloads (big data, AI/ML, cloud native applications) and hardware specialization.

2.3 Consumption based or PayU service model

By 2025, at least 50% of enterprises will leverage opex-based storage consumption models, compared with currently less 10% according to Gartner [4]. Facing strong competition from cloud providers, storage vendors are increasingly willing to offer ways to keep their customers operations on premises. They have to offer customer new programs on subscription bases. PayU or consumption-based service model enable customers paying only for the resources they use, moving from a Capex to an Opex model.

Customers could get cloud-like services in their data center and other on-premises locations with minimal upfront costs. However, services offered from vendors can vary significantly.
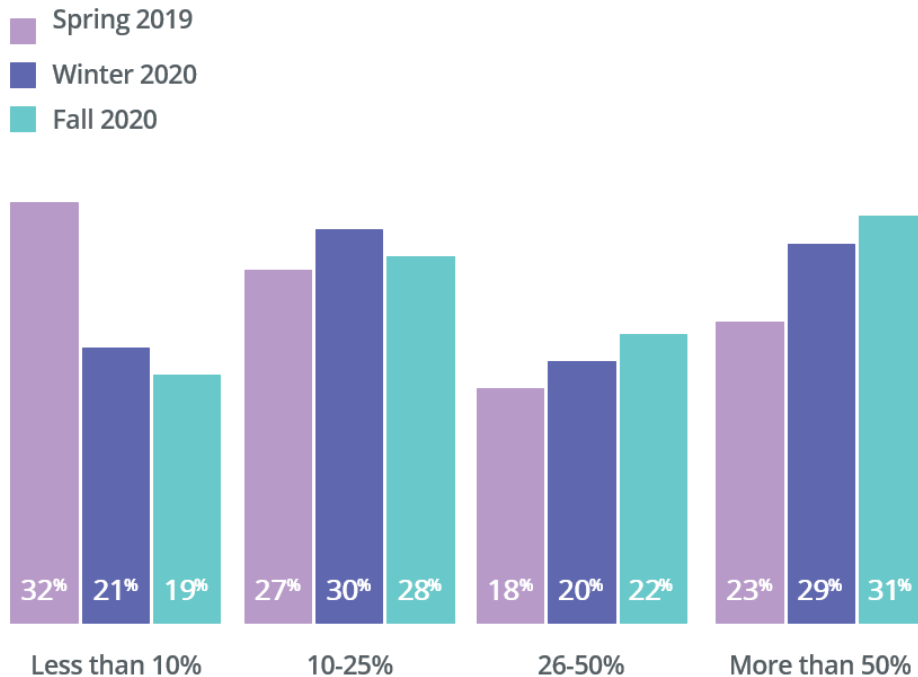
There are also challenges for vendors to offer such services:

1. Agile deployment and configuration to accommodate future workload and business logic changes.
   - Modular hardware design (Building blocks)
   - Finer granularity metering, monitor and predication tools
2. Separations and decoupling of application domain and infrastructure
   - Better integration with application and data platform
3. Flexible and reasonable pricing model
4. Customer finer granularity control of infrastructure
   - Rich and open APIs for data management platform
5. Data governance, risk assessment and compliance tools
6. Data migration and asset management strategy

## 2.4 Containerized Mission Critical Application Proliferation

In the next 5 years, there will be totally half billions cloud native applications created for enterprise, equal to number of applications created in last 40 years. Apparently, customers are containerizing their missional critical applications based on microservice architecture.

What percentage of your apps are containerized today?

- Spring 2019
- Winter 2020
- Fall 2020

	Less than 10%	10-25%	26-50%	More than 50%
Spring 2019	32%	27%	18%	23%
Winter 2020	21%	30%	20%	29%
Fall 2020	19%	28%	22%	31%

Are you using Kubernetes to orchestrate containers?

- Yes: 91%
- No: 9%

If you're using Kubernetes, are you using it in production?

- Yes: 75%
- No: 25%
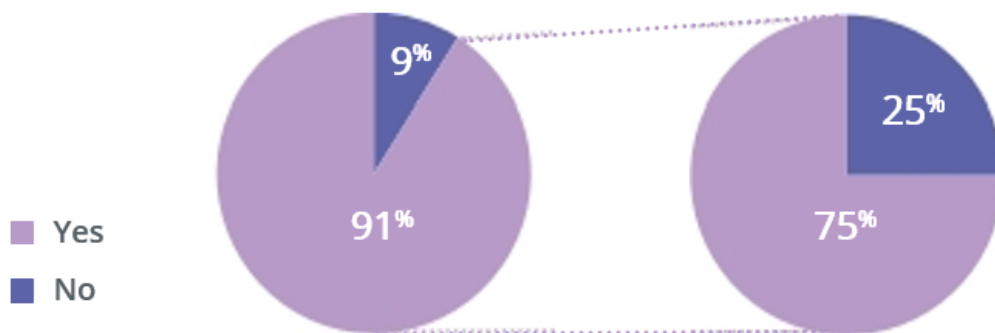
By 2022, there are about 70% of new applications have adopted microservice architectures. 35% of production application will be cloud native. By 2023, more than 90% enterprise digital economy will come from digital native IT environments.

Among those earlier adoptions, FSI and carriers are top customers driving those efforts.

Banks are embracing open baking concepts and increasingly leveraging edacity and agility offered from cloud native architecture. They have been building and transforming their online banking, payments, load services, wealth management and digital insights applications into cloud native architecture.

Carriers are also transforming their traditional business into container-based approach. For example, a major European carrier is building their entire IT infrastructure based on containerization infrastructure using VMware® and modular hardware design. Most of carriers also deploy more CNF to replace VNF in their core network services.

All those containerization efforts demand better integrations and supports from storage vendors beyond CSI. Applications requiring enterprise features will continue to demand those features after containerization. How to better leverage rich enterprise features and integrate with data and application platform is one of major factor for customers to choose storage products for their underlying infrastructures.

2.5 CLOUDOPS AND HYBRID CLOUD

Nowadays, discussion about hybrid cloud support seems like a cliché. However, a good viable hybrid cloud policy is vital to storage company growth. That is exactly how NetApp® does.

Based on NetApp® recent investor day conference, NetApp® thinks its public cloud annual recurring revenue will be $2 billion by 2026. Some analysts project NetApp® TAM has increase up to 96 billion by 2025 with CAGR.

It has set itself new revenue targets as a result. NetApp® 's hybrid cloud business TAM has a 4–6 per cent revenue CAGR with fiscal 2025 revenue targets of $6.6 to $7 billion. Public cloud revenues should be $1.3 billion to $1.4 billion with an expected 50 per cent public cloud revenue CAGR. [5]

NetApp Fiscal 2025 Segment Revenue Targets
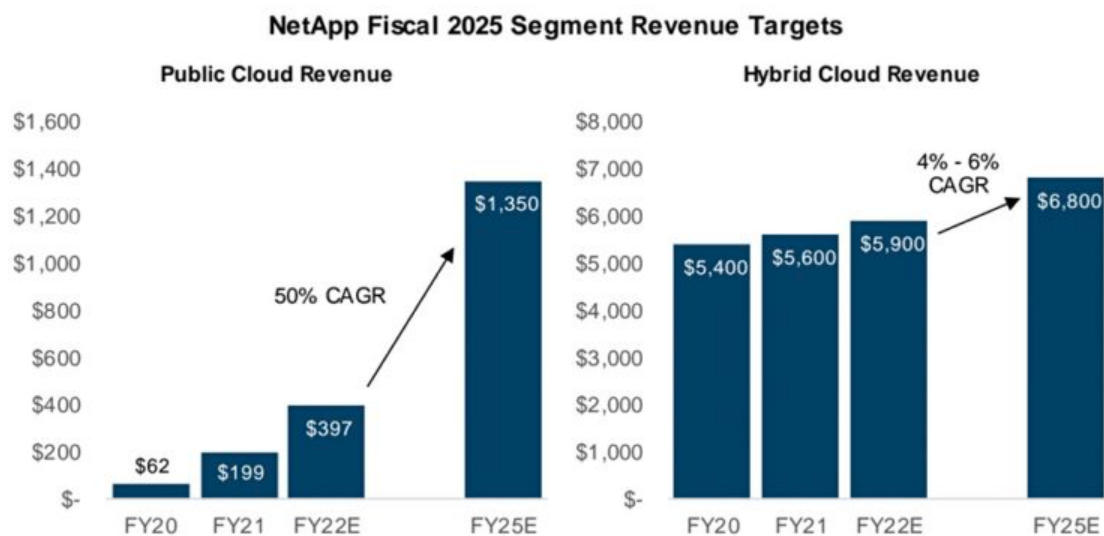
Note: All Values in Millions
Source: NetApp 2022 Investor Day and William Blair Equity Research

Software-defined storage initiatives extend on-premises storage workloads to operate seamlessly across the distributed hybrid cloud substrate with common management and

data services. SDS is central to disaggregated, scalable storage- compute architecture strategies.

By 2024, at least 20% of enterprises will leverage cloud storage management tools to integrate on-premises storage platforms directly with the public cloud for backup and disaster recovery (DR) use cases.

We predict that due lacking of enterprise file capability from public cloud vendors(for example AWS EFS), enterprise file service could be a major driving force for primary storage vendor to enter public cloud service sector (NetApp® already did that with its data fabric service in public cloud and other startups like Qumulo, Nasuni also offer such services in public cloud), with Kubernetes as technology enabler.

2.6   AIOps

Driving down data center costs without sacrificing QoS is always one of the major goals for CIO for enterprises. More and more IT vendors including storage vendors have invested in AI/ML for IT operations and system automations. The goal is to provide more proactive and resilient platform with the potential to reduce overall support and maintenance costs and overall IT infrastructure management complexity.
We are expecting this trend to continue in the following years with AI/ML technologies widely adopted by enterprise customers.

Vendor will invest more on AIops with popularity of consumption-based service model to further drive down costs and improve their margins.

2.7   HARDWARE AND SOFTWARE ACCELERATION ADOPTION

With popularity of NVMe technologies and faster interconnects and media, we are going to see more enterprises adopt SCM, NVMe over Fabric in data center environments. However, it is still a question if customers are willing to invest more toward RDMA capable network (requiring DCB) than NVMe over TCP.

We are also seeing newcomers add GPU, highspeed interconnect enabling hardware(CXL) to their storage product. It mainly caters the demands of data-intensive applications including HPC and ML.

There are also a slew of companies working on boosting ML/HPC workload performance by leveraging RDMA like technologies for GPUs (GDS).

All those storage systems require ultra-high speed interconnect like RoCE/InfiniBand or even faster CXL etc.

Of course, SCM has been in the discussion of earlier adoption for quite a while. Now, with high speed memory semantic network emerging, it makes sense to have composable system with SCM as very front end pooled from various distributed nodes enabled by such high speed interconnect [6].

Server Element Resource Pools

CPU

GPU

FPGA

Optane

NVMe SSDs

Liqid
Command
Center v2.2

Composed Server

CPU

GPU

FPGA

Optane

OpenFlex

PCIe Gen 3, 4,
Ethernet,
InfiniBand

Gen-Z,
CXL
coming

Western
Digital
OpenFlex
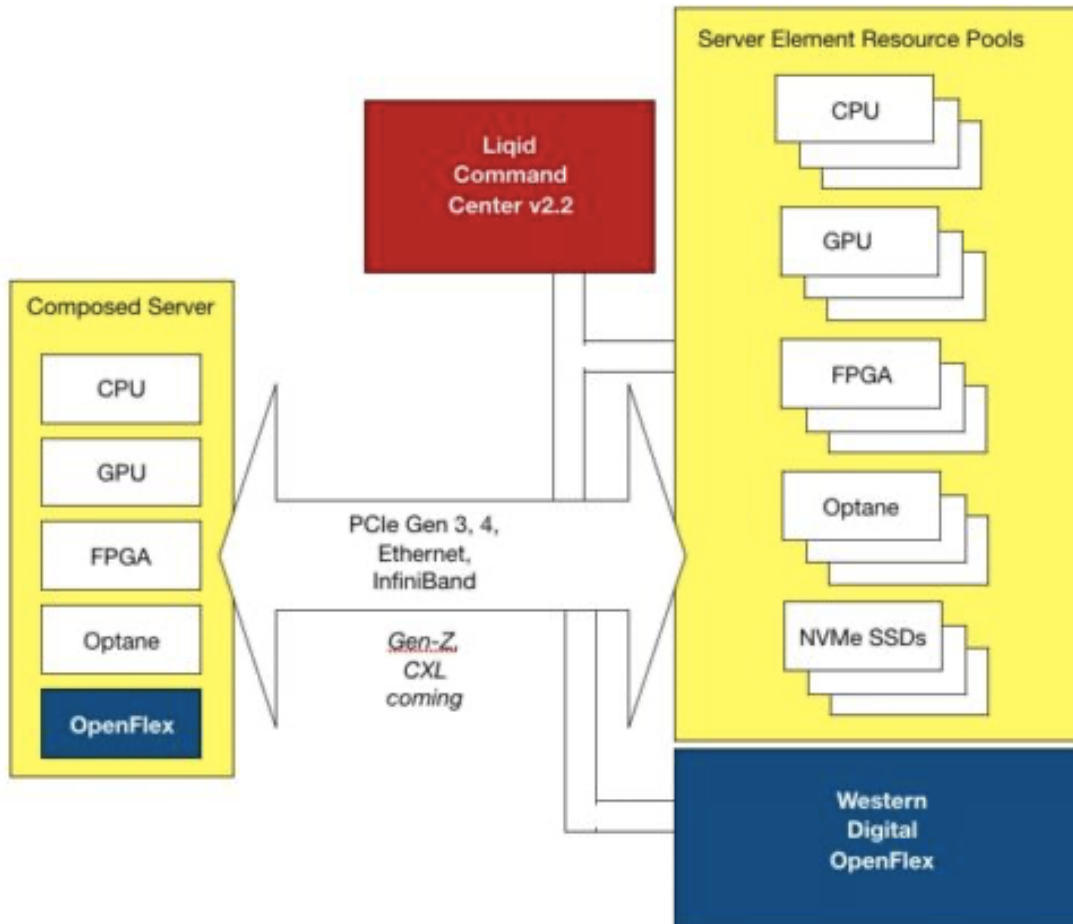
3 CURRENT PRIMARY STORAGE SURVEY

In this chapter, we briefly elaborate some of existing primary products in 4 aspects: Data services, media, host interfaces and AIops and management. We are not going to dive deeply into details. Some of existing primary storage pain points are also examined.
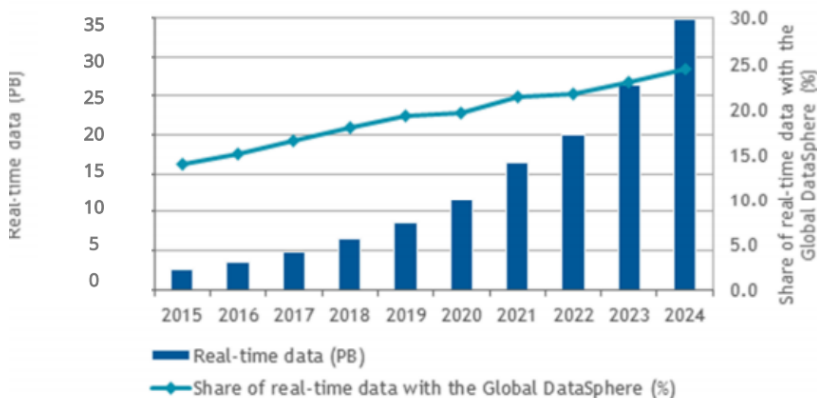
3.1 DATA SERVICES

Enterprise or primary storage has almost as complete enterprise data services as it can to this date. Those data services include Data protection, Data Replication, QoS, Multitenancy, etc. As enterprise applications are undergoing transition from VM based to containerized micro-service architecture, enterprise features are still demanded by those enterprise applications. **How to offer application-level granularity data services is one of the key developments for storage vendors.**

3.2 MEDIA

With SSD as main storage media for primary storage as a foregone conclusion, we are going to see less and less storage vendors to offer hybrid storage system in primary storage market (Not true in secondary storage market). The cost, production, and technology maturity of SCM are taken a hit due to current supply chain problems. Most of storage vendors experience various level of supply chain problems. For now, they most likely will focus on SSD shortage problem first. That may delay the adoptions on SCM and other faster storage technologies. **However, we are going to see some major vendors (some startup like memverge already did) to provide SCM as main storage in the next few years instead of caching**. It corresponds to changes of workloads in data centers.

Real-Time Workloads Are On The Rise

Worldwide Real-Time Data and Share, 2015-2024

Worldwide, data is growing at a 26.0% CAGR, and in 2024 there will be 143 zettabytes of data created

By 2021, 60-70% of the Global 2000 will have at least one mission-critical real-time workload

Source: IDC's Global Datasphere, 2020

3.3 Host interfaces

With faster media (SCM, SSD, etc) and fast speed RDMA capable network technology are developing in data center and cloud, NVMe over Fabric is expected to become dominant front-end protocol going forward for block access.

Type	SAN with SCSI front end					SAN with NMVe front end			
	FC SAN	iSCSI/FCoE	iSER			NVMe over FC	NVMe over TCP	NVMe over RoCE	NVMe over Infiniband
Network Infrastructure	FC	Ethernet	RoCE	Infiniband		FC	Ethernet	RoCE	InfiniBand
Transport Protocol	FC	TCP/IP	RDMA			FC	TCP/IP	RDMA	
Speed	High	Low	Higher	Higher		Higher than FC SAN	Higher than iSCSI	Very High	Highest
Hardware Cost	High	Low, commodity hardware	Between FC SAN and iSCSI	Very High, limited choice		High	Low, commodity hardware	Between NVMe over TCP and NVMe over FC	Very High, limited choice
Scalability	High	Very High	Lower			High	Very High	Lower	
Software Ecosystem	Mature	Mature	Niche, didn't take off			New, 2017 OS/Hypervisor support is catching up. Lack of support on Windows Easy upgrade from FC SAN	Very new, 2019 lack of OS/Hypervisor support	New, 2017 OS/Hypervisor support is catching up. Lack of support on Windows	

NVMe-oF will truly unleash the performance potential of flash media in data center. Three elements need to be aligned. First, users will need an NVMe-capable storage network infrastructure, in place. Second, the OS/Hypervisor need to support NVMe-oF. Third, customers will need storage array systems that feature native NVMe.

However, there are some customers are concerning overall costs and complexity associated to deploy DCB capable data center network. They prefer NVMe over TCP to NVMe over RoCE or InfiniBand. Although clearly, NVMe over RDMA capable network is better than NVMe over TCP, it does not warrant an overwhelmingly advantage for customers to move to that direction except in certain use cases with high throughput low latency workloads. On the other hand, NVMe over TCP offers cheaper, compatible way of harnessing performance from faster media. **We will see more customer adoptions on NVMe over TCP in the next few years.**

With high-speed interconnect such as CXL, Gen-Z and RDMA network technologies, **we are seeing vendors also moving toward even higher hierarchy to support memory semantic storage tiers with faster speed nonvolatile memory so called big memory computing** [7].

There are several major driving forces for such sea change:
- Market evolution calling for real time data processing and data centric business models enabled by AI/ML

- Emerging PM technologies with memory virtualization with intelligent data placement
- Prevalence latency sensitive and real time analytics workload

3.4 AIOPS AND MANAGEMENT

AIOps storage platforms has been developed in recently years. Vendors have invested in advanced AIOps capabilities to support reducing TCO, minimizing operational complexity and enabling SLA-based, managed, storage as a service platform. AIOps platforms enhance a broad range of IT practices, including I&O, DevOps, SRE and service management. However, the more focused outcomes are within the I&O domain and include anomaly detection, diagnostic information, event correlation, and root cause analysis (RCA) to improve monitoring, service management and automation tasks. With AIOps, vendors have added proactive features to prevent faults from happening (fault predication) instead of post-mortem processing (automatic ticket issuing). Various ML/DL mechanism are employed to facilitate those efforts. The functionalities for AIOps include Ingestion, Topology, Correlation, Recognition, Remediation. [8]

With more and more enterprise customers are choosing subscription-based storage consumption models. It is critical for storage vendors to provide continuous insights across IO operations.

AIOps Platform Enabling Continuous Insights Across IT Operations Monitoring (ITOM)
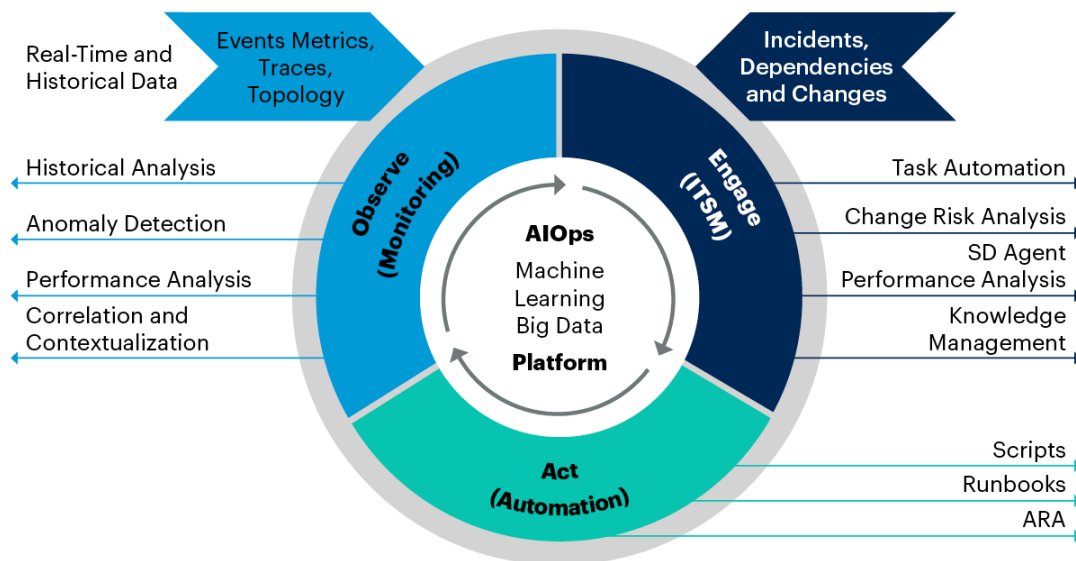
Source: Gartner
735577_C

Gartner.

Due to ample open source tools including Prometheus, Elastic Beats, Jaeger and Fluentd etc, most of storage vendors have provided certain level of supports for AIOps.

For storage vendors, how to integrate its own AIOps capabilities into a broader effort for IT AIOps is one of the key points to be developed in the future (Eco system support).

3.5 EXISTING PAIN POINTS

3.5.1 Server Virtualization

Virtualization technologies have been around a long time. It has driven overall data center compute efficiency tremendously. However, it does present security and performance problems which inherently with hardware virtualization.

Container technologies are coming into help to alleviate performance issues associated with hardware virtualization technologies. Existing hardware vendors have been embracing container technologies for efficiently sharing their hardware resources with reduced performance penalties.

It may not be surprising with Kubernetes or other container platform as unified resource scheduler and manager; virtualization is shifting to container-based virtualization with hardware specialization.

3.5.2 AI/ML

"By the end of 2021, 60–70% of the Global 2000 will have at least one real-time, mission-critical, and AI-driven workload, and many of them will have several workloads. In many cases, these next generation applications have information technology (IT) infrastructure requirements that are not well met with legacy architectures, and IDC research indicates that roughly 70% of all enterprises will be modernizing their server, storage, and/or data protection infrastructure within the next two years as part of their digital transformation (DX) journey. Furthermore, over 91% of these enterprises deem successful infrastructure modernization projects as a critical success factor in their DX".

This statement from IDC generalizes some of pain points for data center to support dynamic workloads driven by AI/ML. Some of vendors like DDN, WekaIO support GDS to provide ample throughput with low latency for AI/ML workload.

With high-speed interconnect more prevalent, specialized GPU/DPU and storage intensive node connected with high-speed network is also a viable approach.

3.5.3 Security, Anti-Ransomware and Data Governance

Ransomware has increased 435% between 2019 and 2010 along. We have seen customers especially larger corporations are experiencing some major outage due to that. In the organizations that are targeted, time spent on migrating a ransomware event could

be potentially and disruption to business activities. The current approaches to ransomware center around detection and response. What roles do storage systems play in anti-ransomware domain need to be thoroughly discussed (Pre-execution detection, recovery, etc).

As all countries all focus on data sovereignty and security, we see a lot of regulations in EU, US and other countries adopted. How can storage system better support those eco system is also a hot topic.

3.5.4   PayU or STasS Business Model

As mentioned in the previous section, customers are embracing and asking for vendor to provide consumption-based business model. Although, it gives customers clear advantage for TCO reduction, it brings several challenges to storage vendors:

- consumption monitoring, metering and predication tools are more important for vendors to reduce cost of sales.

- Vendors shall implement AIOps to reduce their own O&M costs to improve their margin.

- Vendors shall implement strategies to power finer granularity agile deployment (Building blocks etc). Only 23 per cent in Europe have sufficiently capable automation/orchestration and other advanced management and application development and deployment capabilities to adjust their IT service delivery dynamically. Data sovereignty and Security

3.5.5   Green Initiative and Carbon footprint

Climate change and sustainability have become political factors that play a growing role in business management. There is more and more pressure for companies to limit their greenhouse gas emissions and take action to become more environmentally friendly.

Currently the global carbon footprint for data centers accounts for more than 2% of global carbon emissions (the equivalent of the world's entire airline industry), with the number expected to rise to 3.2% in 2025. By 2040, data storage is predicted to account for 14% of the world's carbon emissions. [9]

Carbon emissions from data centers and server rooms are a growing global concern with the rapidly increased demand for data processing and storage. With customers especially European customers require sustainability from data center providers. Storage vendors shall prepare to create products which are more energy efficient, easy for recycling, conscious of carbon footprint.

3.5.6   Edge Proliferation

Based on IDC, there is an average increase of 38% edge investment projected for the next 2 years.

With edge along with its data proliferating, network is become major bottleneck. This demands edge put more focus on data processing and minimize data sent back to data centers. At the same time, a unified data platform across edge, data center and cloud is preferred to facilitate data/application flow seamless from and to data center and cloud.

HCI is typical solution for this. However, how to adapt new workload such as AI/ML is one key aspect. **Vendor may need to build a competitive HCI solution with GKPU/DPU and unified data platform for edge storage and compute to have better growth beyond their primary markets.**

69% of orgs plan to increase edge investments in next 2 years. Average increase = **38%**
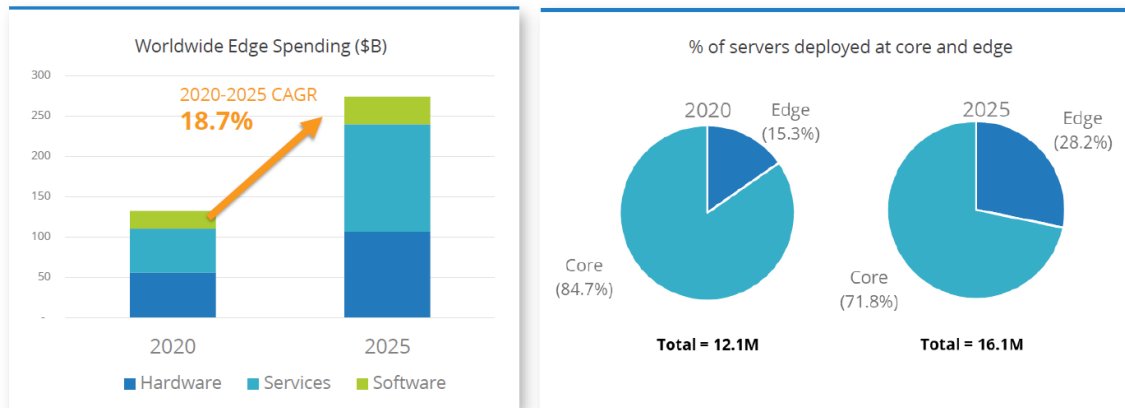
Source: IDC EdgeView, December 2021, n = 1500
IDC's Edge Spending Guide, 4Q21
Worldwide Edge Compute and External Storage Systems Infrastructure Market, 2021-2025: Market Trends and Outlook (IDC #US48190021, Sept 2021)

≡IDC

© IDC | 3

3.5.7   Eco System Support

With sheer amount of data predicted to grow exponentially, various new and existing applications are introduced into customer data centers. There is no single vendor that can build a complete data center solutions for its customers alone.

Qualified, certified, and tested by various third-party hardware and software may use tremendous resources. Although bigger companies may have resources and "tractions" to attract other vendors to qualify their solutions and products and conduct compatibility tests, this can't be applied to all storage vendors due to various reasons (geo-political, resource constrains, company relationships and customer appeals etc).

With more and more customers embracing data centric model, storage vendors may have chances to move their stack up to application level by integrating data management platform. Eco system supports for storage vendors may become bigger problems to solve by all participants in the data center.

3.5.8 Storage Cloudification

Cloud storage competition is real. Previously public cloud vendors can only provide consumer level storage services. As public cloud vendors are foraying into enterprise sectors, there are more and more enterprise level storage services offered by public cloud vendors.

One of public cloud vendors strength is application-level integrations. They are able to offer application or PaaS level services to customers, therefore, underlying storage service is invisible to customers. A lot of enterprises capabilities have been built on top of application or platform level. Underlying storage does not have to be "premium" storage system anymore.

Although it works very well so far, there are exceptions. For example, EFS is targeted to attract enterprise NAS customers to public cloud. AWS quickly found out EFS is lacking a lot of file services for enterprise customers. NetApp® seizes opportunities and offers their file services into public cloud. Initially, they deploy hardware systems into AWS and Azure, quickly they morphed into software only solution.

SDS is a key for hybrid cloud solution which customers are asking for. Traditional SDS vendors are not quite mature and lacking enterprise features (they are startups). Primary storage vendors shall take the opportunities to extend their footprint into public cloud by cloudifying their storage products. With cloud native application proliferation, it presents even more opportunities for primary storage vendors.

4   CURRENT VENDOR PRODUCT STRATEGIES
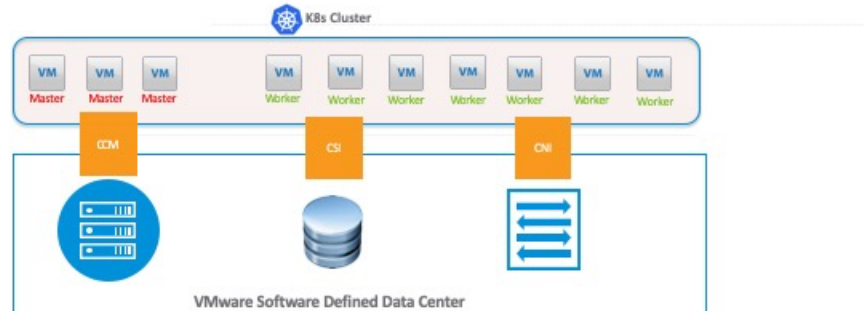
4.1   VMWARE®

VMware® is not a traditional storage vendor. However, with introduction of vSAN and popularity of SDS and HCI, market is seeing more and more storage solution offered from vSAN.

VMware® posts a unique challenge to most of storage vendors due to the fact that it dominates virtualization platform market. With VMware® enters container platform market with Tanzu platform, it presents a fearful domination to eschew primary storage vendor's market share. All virtualization workload can be consolidated into vSphere platform including traditional VM applications and containerized applications. [9]
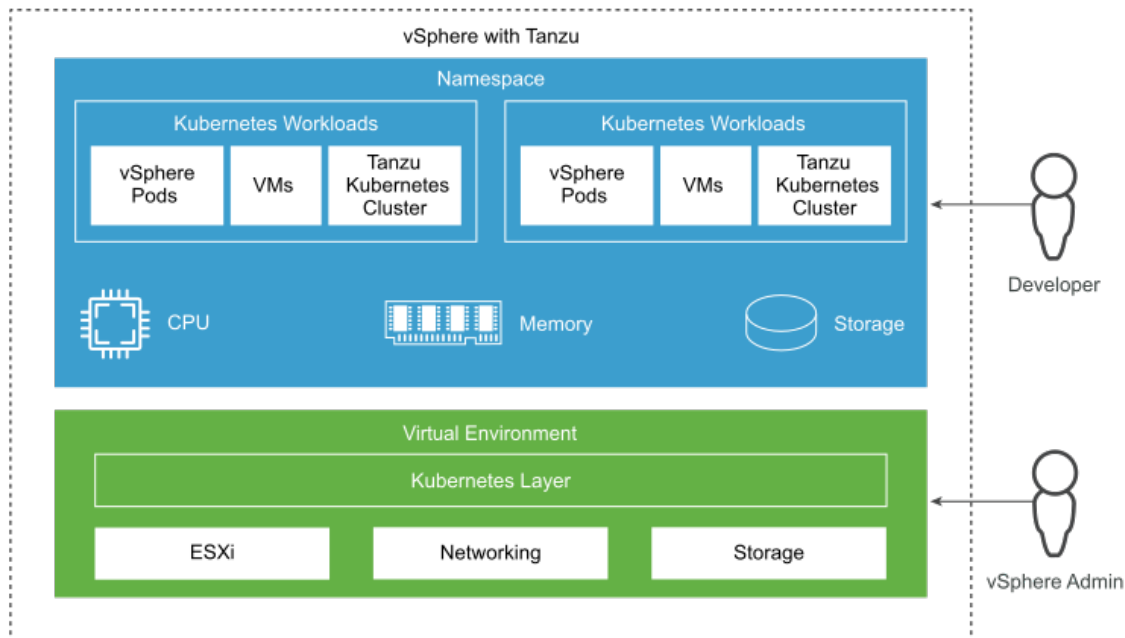
VMware® leverages its universal eco system supports (no vendor can claim no support to VMware®), it extends its already dominating position in vVol for virtualization market to container market via vSphere CSI support. **All other vendor storage will be just like another "smart" storage managed under vVol framework. Slowly, VMware® will eventually add more and more enterprise features(file support is underway) to replace those vendors systems.**

Kubernetes CCM/CSI/CNI

Traditional primary storage vendors shall prepare this and have a concrete strategy. It is not an easy task due to VMware® dominant virtualization platform position. However, with Kubernetes platform gaining popularity, it presents primary storage vendors an opportunity to integrate their product with such data platforms.

4.2   NETAPP®

NetApp® is fully committed to embrace the cloud. As mentioned in previous section, NetApp® CloudOps is becoming a faster growth revenue contributor in the company. NetApp® fully expects its business will grow even bigger share in its portfolio.

NetApp® also is foraying into future container platform data management market. Its Astro acquisition and later development will well position their products for hybrid clouds and future cloud native applications.

With a CloudOps portfolio helping customer move applications to the public cloud in an automated, manageable, and application-optimizing way, NetApp® is transition itself into a public cloud storage company.

It's a risky move. To hedge the risk, we believe NetApp® will keep investing their AFA storage systems going forward too.

4.3   PURE STORAGE®

For the first time, Pure Storage® has a net profit of $15 million for fourth fiscal quarter of 2022. It is due to their high growth in U.S. market with their FlashArray. AFA is still going to be its bread and butter for the foreseeable future since it is NOT a market leader in primary storage market. There are still a lot of catchup work to do.

With acquisition of PortWorx, Pure Storage® has entered SDS and storage cloudification arena. PortWorx can really offer customer data services by seamlessly deploying storage services across different clouds and data centers. As mentioned before, storage vendors need time to build enterprise services for their storage products, which typically are lacking in startup storage systems. Going forward, we will not be surprised that Pure Storage® is going to add more enterprise data services into PortWorx. This is the path similar like NetApp® .

4.4 DELL EMC

DELL EMC has a very broad and comprehensive portfolio. It is NOT expected it will consolidate their storage offering. It will still concentrate their efforts on its Power Series offering.

It does look like that Dell is investing more on APEX cloud and as-a-service offerings. With its multi-cloud focus, DELL is developing similar path as NetApp® with Project Alpine, which may deploy all its storage software to 3 main public clouds.

It's likely that all mainstream storage vendors and the up-and-comers will feel customer pressure to do the same as Dell and NetApp® – meaning Hitachi Vantara, HPE® , IBM® , Infinidat® , Pure, and VAST Data for example.

APEX StaaS offering provides consumption-based on prem solutions by leveraging all DELL storage products. APEX cloud service is built on VCF and VxRail to provide customer Hybrid Cloud, Edge solutions. It is also moving their storage infrastructures to the cloud but with a slower pace than NetApp® .

5 FUTURE PRODUCTS

5.1 POTENTIAL FUTURE PRIMARY STORAGE

Predication is always a tricky part of IT. The infamous quotation for 640K memory has lead Bill Gates repeatably denying such claim. But we will do it anyway.

Software capabilities and broader ecosystem support are vital to storage vendor products especially in primary storage sector with hardware differentiation as core competence for better customer experiences.

Primary storage shall be able to provide agile hybrid cloud deployment for readiness of deploy anywhere including data center, co-lo and public cloud. It also need support on delivering subscription business model to reduce customer TCO with its agile deployment and AIOps.

Fully embrace application as data and data/application platform integration in primary storage and take full control of data/application entrances and provide more application centric semantics to integrate with data/application platform.

Primary storage shall have specialized hardware form to accommodate and consolidate dynamic workload such as AI/ML with highly distributed architecture with modular design and managed with data/application platform and distributed OS. With unified resource management and scheduler, application platform can leverage various specialized hardware for compute and storage needs from application.

Extensive data sovereignty and security support can shall be added into system to enable customer digital ownership.

NetApp® is leading the way for primary storage vendor to transform into cloud service provider. It is a bold movement; the result remains to be seen. We expect that most major primary storage vendors will continue to innovate in on prem storage markets.

5.2 REFERENCE ARCHITECTURE

Currently, storage vendors are working on enterprise data management platforms such as Astro, Portworx, K10 etc. It more likes the following diagram:

Existing Application Full Stack Architecture
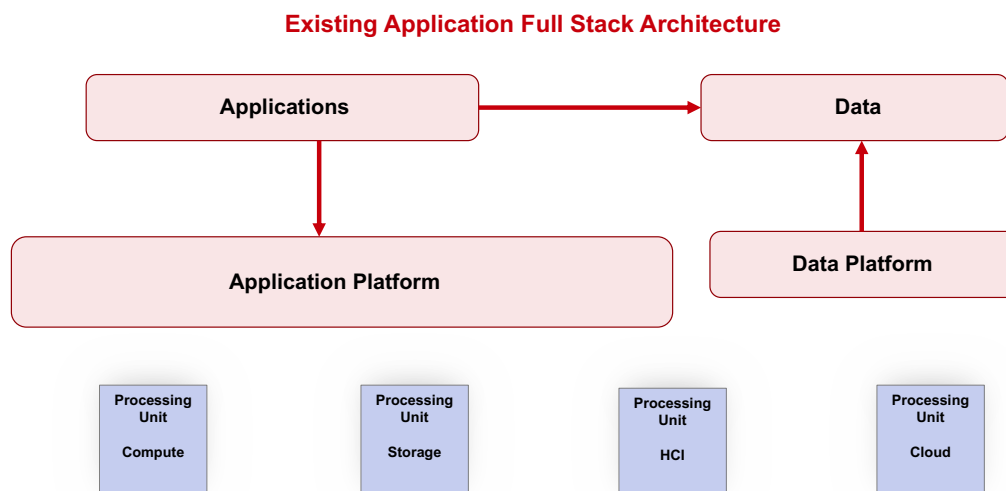
It is a very good first step towards full application/data platform development. However, it still treats data and application as separate entity. New generation of data platform shall treat data and application as one single entity such as data product. Future storage product shall be an integration with this new application/data platform as shown in the following diagram.

Application platform shall have the capabilities of following:
1. Universally (across applications) schedule, manage and expose resources underly infrastructure.
2. Translate application SLA to infrastructure and eventually storage SLA
3. Manage and safeguard data resources with data centric approach
4. Provide storage and compute services (Staas, IaaS) by leverage underlying infrastructure capabilities.

Primary Storage Reference Architecture
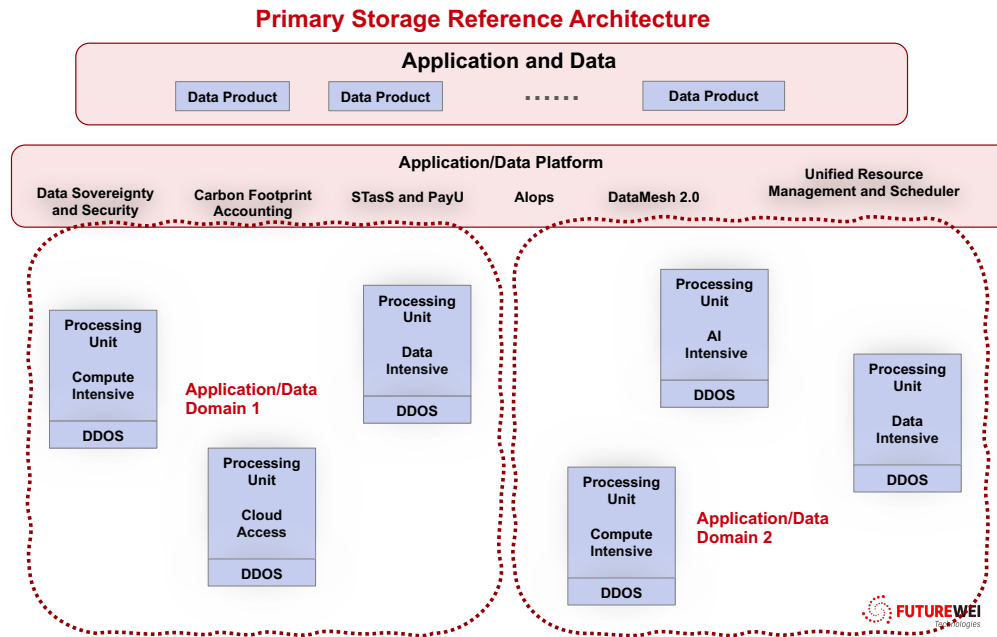
6 CONCLUSIONS

In this white paper, we have examined some recent primary storage trends. We also illustrate 10 key initiatives that may potentially have impact on future primary storage product. We also make recommendations on overall picture of future primary storage product. The goal of this whitepaper is more on promoting discussions than providing solutions.

7 BIBLIOGRAPHY

[1] Z. Dehghani, Data Mesh Delivering Data-Driven Value at Scale.

[2] "IDC Worldwide and U.S. External Enteprise Storage Systems Forecast".

[3] "HCI vs. traditional IT vs. dHCI vs. composable infrastructure".

[4] "Critical Capabilities for Primary Storage".

[5] "NetApp® is increasing targets thanks to CloudOps biz," [Online]. Available: https://blocksandfiles.com/2022/03/23/cloudops-business-looks-so-good-for-NetApp® -its-increasing-its-targets/.

[6] [Online]. Available: https://blocksandfiles.com/2022/03/25/liqid-composes-memverge-dram-and-optane-memory-pools/.

[7] "Digital Transformation Driving New "Big Memory" Requirements".

[8] Gartner, "Markete Guide for AIOps Platforms".

[9] VMware® . [Online]. Available: https://tanzu.VMware® .com/content/blog/a-deep-dive-into-the-kubernetes-vsphere-csi-driver-with-tkgi-and-tkg.