GIGAOM

Enrico Signoretti, Arjan Timmerman, Max Mortillaro
Jun 29, 2021

GigaOm Radar for Enterprise Kubernetes Data Storage v1.0

Traditional Storage and Software-Defined Storage

Data Storage, Hosted Kubernetes

GigaOm Radar for Enterprise Kubernetes Data Storage

Traditional Storage and Software-Defined Storage

Table of Contents

1. Summary

The ways applications are developed and deployed are changing rapidly, pushing the storage industry to evolve quickly as well. Applications are moving to containerized environments, and containers—and thus Kubernetes—are becoming more prevalent. Businesses are rearchitecting their applications to fit this new model, and storage solutions that support containerized environments are in high demand.

Kubernetes adoption has accelerated to the point where it is now the dominant container orchestration tool. Enterprises are still transitioning, but the adoption of Kubernetes has grown exponentially. Because of the increasing interest in container-based application development, IT organizations began working on proofs of concept, which then moved to development and test platforms. During the pandemic, many of these companies fast-tracked their Kubernetes environments into production, building these environments on top of the existing infrastructure.

During that accelerated transition, those initial small Kubernetes infrastructures running fairly simple applications with limited data storage needs developed into environments that support the livelihood of the organization. Soon, more and more stateful applications began migrating to these platforms, requiring additional resources and performance.

At the same time, enterprises of all sizes began embracing hybrid cloud strategies that are becoming more complex and structured. We are moving quickly from a first adoption phase, when data and applications are distributed manually and statically in various on-premises and cloud environments, to a new paradigm in which data and application mobility is the key to flexibility and agility.

These days, enterprises want the freedom to decide where applications and data should run. The public cloud is known for its flexibility and agility, but on-premises infrastructures are still better in terms of efficiency, cost, and reliability. More and more companies do development and testing on the public cloud, with production on-premises, in the cloud, or both, depending on the business, regulatory, economic, and technical needs of the particular enterprise.

Kubernetes has become instrumental in executing this vision, but it needs the right integration with infrastructure layers, such as storage, to work. Persistent and reliable data storage, along with data management and security, are vital factors to consider when evaluating and implementing Kubernetes deployments in enterprise environments today. These factors expand the scope of the orchestrator to a broader set of applications and use cases across different types of on-premises and cloud infrastructures. The goal is to provide a common data storage layer that is abstracted from physical and cloud resources, with a standard set of functionalities, services, protection, security, and management. At the moment, many enterprise storage solutions are reasonably compatible with Kubernetes, but they don't work with the public cloud. This is a limitation you should always take into account when evaluating this technology, especially if you are planning a hybrid cloud environment.

HOW TO READ THIS REPORT

This GigaOm report is one of a series of documents that helps IT organizations assess competing solutions in the context of well-defined features and criteria. For a fuller understanding consider reviewing the following reports:

Key Criteria report: A detailed market sector analysis that assesses the impact that key product features and criteria have on top-line solution characteristics—such as scalability, performance, and TCO—that drive purchase decisions.

GigaOm Radar report: A forward-looking analysis that plots the relative value and progression of vendor solutions along multiple axes based on strategy and execution. The Radar report includes a breakdown of each vendor's offering in the sector.

Solution Profile: An in-depth vendor analysis that builds on the framework developed in the Key Criteria and Radar reports to assess a company's engagement within a technology sector. This analysis includes forward-looking guidance around both strategy and product.

2. Market Categories and Deployment Types

For a better understanding of the market and vendor positioning (see **Table 1**), we assess how well solutions for Kubernetes are positioned to serve specific market segments.

- **Small-to-medium enterprise:** In this category, we assess solutions on their ability to meet the needs of organizations ranging from small businesses to medium-sized companies. Also assessed are departmental use cases in large enterprises, where ease of use and deployment are more important than extensive management functionality, data mobility, and feature set.
- **Large enterprise:** Here, offerings are assessed on their ability to support large and business-critical projects. Optimal solutions in this category will have a strong focus on flexibility, performance, data services, and features to improve security and data protection. Scalability is another big differentiator, as is the ability to deploy the same service in different environments.
- **ISP/MSP:** In this category, solutions that are suitable for ISPs and MSPs are assessed. These should include multi-tenancy capabilities and the ability to throttle performance per tenant.

In addition, we recognize two deployment models for solutions in this report—enterprise storage systems with a Container Storage Interface (CSI) plug-in and software-defined solutions with optimizations for containers.

- **Software-defined storage**, though not native to the container world, has the advantage of decoupling storage policies from the storage media. SDS solutions are hardware-agnostic and so they can execute on a broad variety of hardware. They are also distributed, and in that regard are similar to Kubernetes-native solutions. Because SDS solutions are not bound by a defined physical storage architecture like their traditional storage counterparts, they offer more flexibility and can be installed in the cloud. Furthermore, it seems likely that SDS solutions will become containerized over time, with even deeper integration into the container ecosystem.
- **Traditional storage with CSI compatibility** is usually built on physical controller-based architectures. These solutions, whether providing block or file (or both) storage capabilities, have been extended to include container support via CSI plugins. Typically administered through a proprietary management console, the CSI plugins allow the container orchestrator to automate storage operations that would otherwise require manual configuration at the storage-array level. Though this is the least flexible of the Kubernetes storage alternatives, traditional storage has the advantage of being already present in many organizations and offering the quickest path to adoption—if not in features, at least in management and ease of use. A major challenge facing traditional storage, however, is the sheer number of backend operations taking place in container environments, which are very dynamic and involve many resource allocation/deallocation operations as containers are created and destroyed. Traditional systems designed for mostly static, stateful, and persistent provisioning may face performance and queuing issues and not be able to keep up with the large volume of operations.

Note that GigaOm is publishing a companion Radar report focused on Kubernetes-native storage that addresses solutions built specifically to support containers. Typically, components are containerized and run within an organization's container environment. These storage solutions are tightly coupled with the container orchestrator and are *container-aware*, so that when the orchestrator spins up or destroys a container, it also handles storage provisioning and de-provisioning. Storage operations in this scenario are automated and invisible to the user.

Table 1. Vendor Positioning

	MARKET SEGMENT			DEPLOYMENT METHOD	
	Small/Med Enterprise	Large Enterprise	ISP/MSP	SDS	Traditional Storage + CSI
Commvault	+++	++	+++	+++	+++
DataCore	+++	+	–	+++	+
Dell Technologies	+++	+++	++	+++	+++
HPE	+++	++	++	+++	+++
IBM	++	+++	++	++	+++
Infinidat	+	+++	+++	+	+++
LINBIT	++	++	+	++	++
NetApp	+++	+++	+++	+++	+++
Pure Storage	+++	+++	+++	+++	+++
Storpool	+++	+	+	+++	++
VMware	+++	+++	+++	+++	+++
Weka	+	+++	+	+++	++

+++ Exceptional: Outstanding focus and execution
++ Capable: Good but with room for improvement
+ Limited: Lacking in execution and use cases
– Not applicable or absent

Source: GigaOm 2021

3. Key Criteria Comparison

Building on the findings from the GigaOm report, "Key Criteria for Evaluating Enterprise Kubernetes Storage," **Table 2** summarizes how each vendor included in this research performs in the areas we consider differentiating and critical in this sector. The objective is to give the reader a snapshot of the technical capabilities of different solutions and define the perimeter of the market landscape. **Table 3** then compares the vendors in terms of the evaluation metrics relevant in this sector.

Table 2. Key Criteria Comparison

	KEY CRITERIA					
	Advanced Data Services	Performance	Optimization	Multitenancy	Security	Monitoring, Alerting, Analytics
Commvault	+++	++	+++	+++	++	++
DataCore	+	++	++	++	+	+
Dell Technologies	++	++	+++	+++	++	++
HPE	++	+++	+++	++	++	++
IBM	++	+++	++	+	++	++
Infinidat	++	+++	+++	+++	++	++
LINBIT	+	+++	++	++	++	+
NetApp	++	+++	++	+++	+++	+++
Pure Storage	+++	+++	+++	+++	+++	+++
Storpool	+	++	+	+++	++	+
VMware	++	++	+++	+++	+++	+++
Weka	++	+++	+	++	+	+

+++ Exceptional: Outstanding focus and execution
++ Capable: Good but with room for improvement
+ Limited: Lacking in execution and use cases
– Not applicable or absent

Source: GigaOm 2021

Table 3. Evaluation Metrics Comparison

	EVALUATION METRICS				
	Architecture	Scalability	Flexibility	Efficiency	Manageability and Ease of Use
Commvault	++	++	++	+++	+++
DataCore	+	+	+	++	+
Dell Technologies	+++	+++	++	+++	+++
HPE	++	++	++	++	++
IBM	++	+++	++	++	++
Infinidat	+++	++	++	+++	++
LINBIT	++	++	++	++	+
NetApp	+++	+++	++	+++	+++
Pure Storage	+++	+++	+++	+++	+++
Storpool	++	++	++	++	–
VMware	++	++	+++	++	+++
Weka	++	+++	+	+	+

+++ Exceptional: Outstanding focus and execution
++ Capable: Good but with room for improvement
+ Limited: Lacking in execution and use cases
– Not applicable or absent

Source: GigaOm 2021

By combining the information provided in the tables above, the reader can develop a clear understanding of the technical solutions available in the market.

4. GigaOm Radar

This report synthesizes the analysis of key criteria and their impact on evaluation metrics to inform the GigaOm Radar graphic in **Figure 1**. The resulting chart is a forward-looking perspective on all the vendors in this report, based on their products' technical capabilities and feature sets.

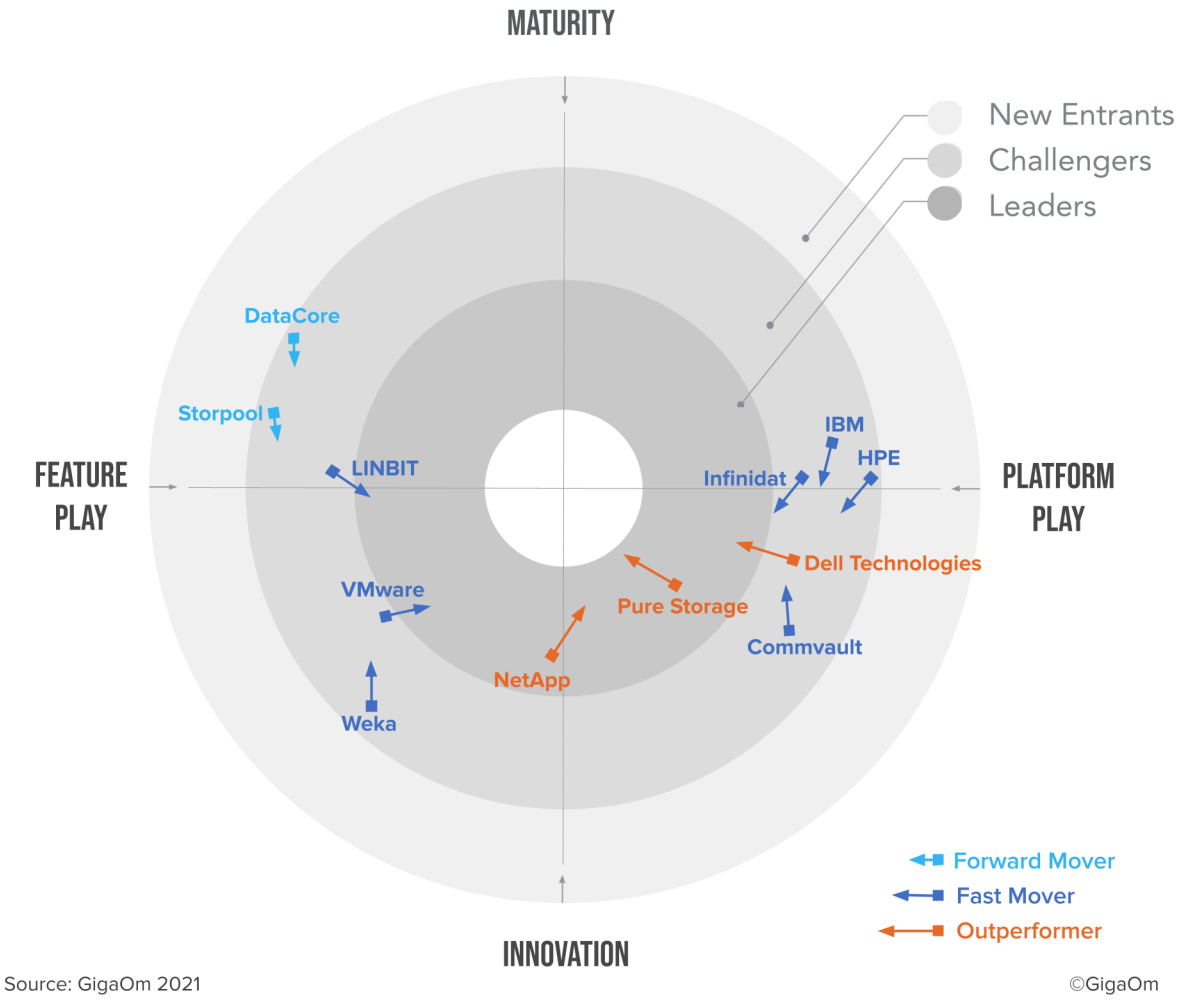

Figure 1. GigaOm Radar for Enterprise Kubernetes Storage

The GigaOm Radar plots vendor solutions across a series of concentric rings, with those set closer to center judged to be of higher overall value. The chart characterizes each vendor on two axes—Maturity versus Innovation, and Feature Play versus Platform Play—while displaying an arrow that projects each solution's evolution over the coming 12 to 18 months.

As you can see in **Figure 1**, Kubernetes data storage solutions for the enterprise are distributed across the entire radar, due to the broad spectrum of solutions and various levels of development and maturity.

The lower-right sector highlights the innovation leaders in enterprise Kubernetes storage. Thanks to its strategic acquisition of Portworx and the swift integration of Pure Storage Orchestrator in Portworx Essential, Pure Storage is now in a position to offer the most complete set of features for Kubernetes storage on its own as well as third-party storage arrays. Trailing Pure Storage, Dell Technologies is the only other vendor with complete support for Kubernetes across its entire range of storage arrays. In addition, Dell offers additional management features on top of its CSI plugin.

Crossing from feature play to platform play is NetApp, with its excellent Trident platform and its Astra software-as-a-service Kubernetes storage solution that can also manage on-premises workloads deployed on NetApp storage arrays. This solution is quickly maturing and offers an interesting take on advanced data services and application mobility.

The lower-left sector of the radar shows VMware, on its way from challenger to becoming a leader. Its Tanzu solution is very complete and offers substantial capabilities to organizations. The dependency on VMware vSAN is a concern, but with vendors leveraging its vVOLs technology, that dependency will lessen. Still, better integration with other storage platforms would be preferable.

Another challenger in this sector is Weka, which is moving from innovation toward maturity. Weka excels in high-performance primary storage, giving the company a strong presence in the world of HPC and AI. Weka partners with SUSE to deliver a Kubernetes storage stack based on Rancher, but with storage delivered by the WekaFS shared file system presented through a CSI plugin. Weka has also developed a partnership with Red Hat OpenShift and Run.ai to better support diverse use cases from its customer base. Weka scales well and is performance-oriented with low latency and strong throughput and IOPS capabilities. Its data protection feature, called Snap2Object Weka, allows snapshots to be taken in S3, and can enable immutable snapshots at the virtual filesystem level. Weka offers excellent capabilities for its market segment, but needs more and faster development for tighter integration into Kubernetes for things like CSI integrated snapshots.

On the right side of the radar, we find three vendors crossing or approaching the x-axis from Maturity to Innovation: Infinidat, IBM, and HPE. This kind of movement usually occurs when a vendor is in the midst of a profound transformation initiative. Infinidat offers a performant and robust platform with proven enterprise-grade capabilities including disaster recovery. Its solution is engineered to cope well under a variety of workloads and can help organizations segment container initiatives without negatively impacting production workloads. Trailing slightly is IBM, with a new solution for the OpenShift cloud platform called IBM Spectrum Fusion that targets mission-critical/AI workloads. The solution integrates well with the Red Hat/IBM ecosystems and includes a number of deployment packs to support various applications out of the box. Last to cross the dividing line is HPE. The company has a very compelling roadmap and a unified set of capabilities across its CSI plugins, and we expect it to execute on the roadmap in the coming 12 to 18 months.

Commvault is a fast moving vendor in the innovation area of the radar (lower right). Its Commvault Distributed Storage (CDS) solution provides a good set of Kubernetes capabilities through its CSI driver, including dynamic provisioning of PVs, lifecycle management, encryption, multi-tenancy, replication and more. Commvault CDS is improving fast and keeps adding new features at a high rate.

On the left side of the radar, crossing from Maturity to Innovation, is LINBIT, a software-defined storage solution built entirely on open source Linux components with a strong focus on replication capabilities, thanks to its DRBD component and its many replication modes.

Finally, two software-defined storage solutions can be found in the upper-left sector of the radar: StorPool and DataCore. Both offer entry-level Kubernetes storage capabilities while focusing their development efforts on other areas, as the smaller arrows indicate. DataCore provides basic capabilities with SANsymphony, but note that investments in MayaData and its OpenEBS cloud-native Kubernetes storage solution make major development efforts on SANsymphony redundant and uneconomic in light of MayaData's advanced capabilities. StorPool, on the other hand, has a robust platform with good enterprise-grade features, but those are not yet exposed at the CSI plugin level to be seamlessly integrated with automation tools.

INSIDE THE GIGAOM RADAR

The GigaOm Radar weighs each vendor's execution, roadmap, and ability to innovate to plot solutions along two axes, each set as opposing pairs. On the Y axis, **Maturity** recognizes solution stability, strength of ecosystem, and a conservative stance, while **Innovation** highlights technical innovation and a more aggressive approach. On the X axis, **Feature Play** connotes a narrow focus on niche or cutting-edge functionality, while **Platform Play** displays a broader platform focus and commitment to a comprehensive feature set.

The closer to center a solution sits, the better its execution and value, with top performers occupying the inner Leaders circle. The centermost circle is almost always empty, reserved for highly mature and consolidated markets that lack space for further innovation.
The GigaOm Radar offers a forward-looking assessment, plotting the current and projected position of each solution over a 12- to 18-month window. Arrows indicate travel based on strategy and pace of innovation, with vendors designated as Forward Movers, Fast Movers, or Outperformers based on their rate of progression.

Note that the Radar excludes vendor market share as a metric. The focus is on forward-looking analysis that emphasizes the value of innovation and differentiation over incumbent market position.

5. Vendor Insights

Commvault

Commvault Distributed Storage (CDS) integrates with Kubernetes through a CSI plugin. CDS provides a good range of functionality covering the requirements that traditionally caused businesses to choose among the different storage options.

CDS has added enhanced CSI capabilities beyond basic integration through the Hedvig CSI driver 2.0, which is at this moment compliant with CSI v0.3.0 and v1.0.0 and Kubernetes 1.13 – 1.19. In addition to dynamic provisioning of PVs, CDS supports storage lifecycle management, including dynamic volume expansion, raw block volumes and on-demand and scheduled snapshots and clones. Enabling automated container data migration between CDS clusters across on-premises and multi-cloud environments is another useful addition.

CDS is intended to serve a broad range of primary and secondary storage use cases for its customers. As such, it supports on-demand and scheduled snapshots and clones for RTO/RPO reduction. CDS also enables automated container data migration between Commvault DS clusters across on-premises and multi-cloud environments via scheduled snapshots. The distributed storage architecture writes data to multiple locations (on-prem and cloud) simultaneously, eliminating the need to physically move data from one location to another.

Commvault offers a broad set of features on the CSI level, like snapshot capability, replication, and multi-tenancy, and on the security side, they offer encryption and KMIP capabilities, as well as RBAC, audit logs, and ransomware protection.

CDS has good monitoring options through its HTML 5 WebUI and is also integrated into the Commvault Command Center, which gives customers a holistic view of their Kubernetes environment. However, the company needs to prioritize visibility and to make a visualization investment based on customer adoption, such as Prometheus.

Strengths: CDS offers good Kubernetes CSI capabilities, and the company continues to develop its solution, adding better capabilities and support in a rapid but well-balanced way.

Challenges: CDS needs to support integration into other monitoring tools like Prometheus. Another feature that would put Commvault into a leadership position is orchestration.

Datacore

DataCore Software has an interesting approach to Kubernetes persistent storage. The company initially developed a CSI plugin for its software-defined storage SANsymphony platform to provide its existing customer base with Kubernetes storage capabilities, and the vFilO product is v1.13 compliant.

Recently, however, DataCore Software invested in MayaData to strengthen its Kubernetes storage capabilities. MayaData developed the MayaStor OpenEBS open source cloud-native Kubernetes persistent storage solution, and a team of DataCore developers is assisting MayaData to further develop OpenEBS. Kubernetes storage capabilities in the main DataCore products are expected to be developed in the future, and we might see a further integration between DataCore and OpenEBS.

Besides handling standard operations, the DataCore Kubernetes CSI plugin allows you to access point-in-time data through DataCore's continuous data protection feature. Low latency and high throughput are made possible using DataCore's in-memory caching and auto-tiering features. The solution is highly available due to the underlying storage architecture, which is interconnected either through Fibre Channel or multipath iSCSI. Both synchronous and asynchronous replication are supported, although no orchestration capabilities currently exist. Data footprint reduction can be achieved thanks to SANsymphony's support for in-line deduplication and compression, and the solution also supports multi-tenancy, such as QoS and capacity quotas.

DataCore's VVOL support is an enabler of Kubernetes technology through Tanzu and is something that helps DataCore/VMware users with a "Universal" VVOL adapter. This means that any storage managed by SANsymphony is thereby integrated with VMware Tanzu through VVOL.

Security capabilities consist of FIPS 140-2-compliant data-at-rest encryption; customers can either store security keys on SANsymphony servers or use an external key management server. Management can be performed through SANsymphony's own management user interface, but users can also feed metrics to Grafana through REST APIs.

Strengths: DataCore provides sufficient persistent storage capabilities to customers starting their Kubernetes journey. Performance and data efficiency capabilities are worth noting.

Challenges: Organizations seeking comprehensive Kubernetes storage capabilities may find themselves limited by the current feature set. Some roadmap updates might help DataCore Users, but a solid plan and description of what will be coming next is yet to be announced.

Dell Technologies

Dell Technologies' Container Storage Modules (CSM) lets customers use traditional storage arrays as multi-tenant, feature-rich enterprise storage for modern apps running on Kubernetes. This gives Dell customers a highly available Kubernetes infrastructure with replication, multi-tenancy, and resiliency support, for a good Kubernetes experience.

Dell supports cloud, on-premises, or hybrid storage deployment models; Dell EMC PowerScale and Dell EMC PowerMax are examples of cloud offerings. Dell notes that its on-premises cloud deployment model is still the one in use the most with customers, and with Dell PowerProtect Data Manager, which supports in-cloud backup protection (on AWS and Azure), the Kubernetes environment can be protected with snapshots and backup. And Dell is currently working on application-consistent snapshots.

Dell's entire CSM portfolio is compatible with the CSI v1.3 spec, and its drivers also support custom extensions for enterprise capabilities such as quota management, replication, observability, and more. Dell offers a rich Kubernetes storage portfolio, optimized for many types of workloads, with an extensive list of enterprise features. Every Dell CSM storage offering comes with the CSI driver, qualified with the latest K8s version.

Dell offers its EMC CSM to improve observability for monitoring, alerting, and analytics, as well as Grafana and Prometheus integration for selected Dell EMC arrays. The observability is meant to give K8s admins essential performance and capacity storage information that can be used to troubleshoot more complex application performance issues. Dell intends to make this available to its complete portfolio, but for now only some of the arrays can be monitored in this way.

Although Dell is clearly offering a good solution for its customers, there are still issues that need attention, like modern application mobility, data security, and a better developer's experience, for its on-premises, hybrid, and cloud storage solutions.

Strengths: Dell has incorporated K8s into many of its storage solutions and is providing a solid and efficient Kubernetes solution. The effort and speed that Dell is putting into adding capabilities to its portfolio is also positive.

Challenges: Modern application mobility, data security, and an enriched developer's experience, as well CSM observability for all Dell EMC arrays, should be a priority for Dell.

HPE

HPE supports a mix of private, public, and hybrid installations for Kubernetes infrastructures, with on-prem support via its HPE Alletra and XP solutions and deployment of HPE Ezmeral Data Fabric as SDS on a public cloud. HPE currently has three different CSI drivers. One supports HPE Nimble Storage, HPE Primera, and HPE Alletra 6000 and 9000. The second one supports HPE XP, and the third supports HPE Ezmeral Data Fabric. Data protection is handled by the partner ecosystem and organizations can integrate their own data protection solution to the HPE Kubernetes storage of their choice.

Data optimization capabilities can be accessed by leveraging the underlying storage features, such as deduplication and compression. The solution supports multi-tenancy features such as QoS, by which users can limit both IOPs and capacity usage, and it also supports encryption.

While CSI gradually matures features and capabilities in the specification, HPE keeps a close watch on differentiating features in its primary storage family of products that could be suitable for implementing in CSI and Kubernetes.

The roadmap is quite interesting and the development pace of new features is good. At the moment, the absence of an orchestrator and additional features to manage large-scale environments can be a

limiting factor. It is not yet possible to get full visibility of the storage resources consumed by a Kubernetes application through standard analytics tools.

Strengths: HPE offers Kubernetes storage capabilities across all of its storage product range, making it simple for organizations to operationalize Kubernetes deployments regardless of their size or the nature of their infrastructures.

Challenges: The absence of a Kubernetes-centric management plane across the entire product range creates unnecessary operational overhead and increases complexity.

IBM

IBM plans to deliver enterprise Kubernetes storage capabilities in the Q3 2021 through IBM Spectrum Fusion, a solution designed for OpenShift, aimed at mission critical workloads and optimized for AI. A cloud-native version should also be available in Q3-Q4 2021.

IBM takes a two-layered approach to delivering persistent Kubernetes storage in the form of an integrated HCI appliance for containers: The infrastructure layer consists of a storage platform built on IBM Spectrum Scale, and the storage services are based on IBM Spectrum software. Users can start with six servers and scale up to 20, and they can also integrate NVIDIA A100 GPUs into the solution.

IBM Spectrum Fusion deploys a scalable containerized file system that handles local and remote backup and recovery. It includes application-aware disaster recovery capabilities plus support for data migration use cases, and offers data efficiency capabilities in the form of erasure coding support.

Security should include encryption capabilities and immutable snapshots as well as role-based access control. Unfortunately, specifications are not yet finalized.

The solution is managed using the IBM Spectrum Fusion HCI dashboard, which provides standard monitoring and alerting capabilities. Integrations are possible with IBM Cloud Satellite and OpenShift Advanced Cluster Management. IBM Spectrum Fusion also includes call-home support and troubleshooting capabilities.

An interesting feature of IBM Spectrum Fusion is the availability of application paks, which consist of ready-to-deploy application packages for popular applications, such as Cassandra, Kafka, MongoDB, and SAP HANA.

Strengths: IBM's offering is a Kubernetes storage solution specifically designed to easily deploy the Red Hat OpenShift container platform in hyperconverged mode. IBM Spectrum Fusion's value is enhanced when used in conjunction with other IBM products or solutions, including IBM Cloud.

Challenges: Security capabilities for the solution are not fully realized. The focus on AI workloads and the development centered on OpenShift limits broader adoption.

Infinidat

Infinidat offered a dynamic volume provisioner before there was a CSI standard, then released the InfiniBox CSI driver in 2020. In May 2021, the company released version 2.0 of its CSI driver, with new snapshot and iSCSI capabilities as the most notable features.

Infinidat's scalable snapshots feature has been built into the product since day one, and other core InfiniBox technologies, such as replication options including active/active synchronous and three-site solutions, are built around those snapshot capabilities. Various backup software platforms integrate with InfiniBox snapshots to provide application consistency and archiving options, including cloud integrations, and Infinidat supports snapshot scheduling with its own software as well.

InfiniBox built extensive performance monitoring into the InfiniBox GUI as well as into the InfiniVerse multi-system cloud service. All monitoring is also accessible through APIs, which can be used with open-source tools like Prometheus and Grafana and platforms such as SolarWinds, Veritas APTARE, VMware vRealize Operations, and more.

The Infinidat architecture is designed for multi-petabyte scale. The data plane gives customers bulk bandwidth and IOPS, and the control plane provides APIs that were carefully crafted to be non-blocking and scalable. The InfiniBox CSI driver leverages the same APIs as other user interfaces, and is scalable and efficient.

Infinidat generally focuses on large and relatively conservative enterprises that typically require high resiliency and long infrastructure lifecycles, as well as on mixing workloads on the same systems—and not just Kubernetes. These companies tend to prioritize multi-vendor integration and standards over bleeding-edge solutions that might end up being reimplemented as they reach standardization. There is an inherent tension in that conservative dynamic versus the very quick release cycles in the Kubernetes world, and Infinidat balances these considerations in its CSI activities.

Infinidat provides a solid CSI solution to its customers for a multitude of capabilities but relies heavily on its partners for others, such as application-consistent snapshots and in-depth monitoring, which forces customers to invest in other tooling to provide a streamlined Kubernetes environment to their users.

Strengths: Infinidat offers a solution that provides Kubernetes capabilities to a subset of the large enterprise market that is more conservative but still needs to provide a decent Kubernetes environment to its users.

Challenges: Building in overall Kubernetes support and fast adoption of new Kubernetes and CSI compliance will need to be a priority for Infinidat.

LINBIT

LINBIT SDS is a software-defined storage solution deeply rooted in the Linux ecosystem, with a strong emphasis on using Linux native open-source technologies. It comprises two components called LINSTOR and DRBD.

LINSTOR acts as the solution's management system, with three functions: *controller* (manages the configuration and resources, which are stateless when running on Kubernetes); *satellite* (handles creation, modification, and deletion of storage resources on each node and runs as a stateless privileged container); and a *user interface* with its own command line and integrations via REST APIs with other management platforms. LINSTOR also handles the orchestration of the LINBIT solution and supports a multiple-processor architecture along with the x86 architecture, including ARM and others, making it suited for edge deployments as well.

The DRBD component handles distributed replication of storage and provides high-availability and data-replication capabilities. Multiple replication modes are supported: asynchronous, memory-synchronous, and synchronous replication. One-to-many replication schemes are also supported.

The solution is built on ZFS or LVM and supports in-line deduplication (with VDO). Organizations planning to use LINBIT SDS for multiple clusters can enable QoS via the configuration of *cgroups*.

In-flight encryption is used for communication between LINBIT SDS control-path components, and data-at-rest encryption is also supported. By configuring a single master passphrase, LINSTOR generates an individual (LUKS) passphrase for each volume. FIPS-140 certification is currently undergoing evaluation. The solution has no management interface of its own, with management available through integration with Prometheus and Grafana; LINBIT provides sample dashboards on its website. LINBIT SDS comes with a K8s operator that handles the complete rollout and life cycle of LINBIT SDS on top of a Kubernetes cluster. It can be integrated with Stork to better assist with scheduling decisions related to performance-based workload placement, and it comes with a HA-controller for quick failover of *Deployments* and *StatefulSets*.

Besides LINBIT SDS (a commercially supported product), a community edition called Piraeus Datastore is also available on Github. Piraeus Datastore was accepted in CNCF's Sandbox recently.

Strengths: An interesting architecture based on pure Linux components will appeal to open source-oriented organizations, LINBIT offers a lot of flexibility in replication schemes to cater to the needs of highly distributed applications. Multiple processor architectures are supported, making LINBIT suitable for alternative deployment options at the edge or on nascent architectures.

Challenges: The solution architecture can seem unnecessarily complex for smaller organizations or enterprises without much experience with the Linux ecosystem.

NetApp

NetApp Trident is an open-source, CSI-compliant dynamic storage orchestrator that enables customers to consume and manage their storage resources across NetApp storage platforms. It integrates with Kubernetes to dynamically provision persistent volume requests on demand. Additionally, Trident has a REST interface that can be used by any application to create and manage storage across the configured resources.

Moreover, Trident is the foundation for Astra (see our K8s radar for more information on this). NetApp Astra addresses issues like data protection, disaster recovery, portability, and migration use cases for Kubernetes workloads through the NetApp data management technology for snapshots, backups, replication, and cloning.

With Trident and Astra, NetApp offers a consistent security and tenancy model, which can be deployed on cloud-native as well as enterprise storage solutions. Trident supports CHAP for authentication, automatic export policy management to control access to NFS shares, and encryption at rest for ONTAP. Trident is also compliant with RBAC and other security norms. And with Astra, NetApp offers predefined roles to provide a set of permissions limiting certain users to performing only specific operations at the application level.

Netapp Trident and Astra include the monitoring tools needed to effectively monitor activity in Kubernetes clusters. Trident has a Prometheus exporter to consolidate infrastructure monitoring in a single platform, and NetApp also provides monitoring tools via its Cloud Insights product/service for the storage administrator.

NetApp's portfolio natively supports functionality like inline deduplication, compression, and compaction. Trident leverages storage efficiency functionality natively supported by NetApp's data storage portfolio and data movers. But integration with SnapMirror and SnapMirror Cloud technologies for moving data across on-premises and cloud systems is still on the roadmap for Astra.

Astra is a fully managed (SaaS) application-aware data management service that manages, protects, and moves Kubernetes workloads in both public clouds and on-premises. Astra is built on top of Trident and enables data protection, disaster recovery, and migration for Kubernetes workloads leveraging NetApp's data management technology for snapshots, backups, replication, and cloning.

Current Trident users will be able to extend their Kubernetes environment even further, when other capabilities are added. Astra will be soon available on-premises and in hybrid environments, as well as in multiple cloud environments. This will give customers application-aware data management functionality, including data governance, for Kubernetes deployments.

Strengths: NetApp provides strong Kubernetes capabilities in its storage solutions. With Trident and Astra, the customer gets the tools needed to offer a manageable, capable, and efficient Kubernetes storage solution.

Challenges: Integration between Trident and Astra is still under development, so customers are not yet able to leverage the complete NetApp feature set with Kubernetes, like SnapMirror for remote replication.

Pure Storage

The position of Pure Storage has significantly shifted since the last report. The key factor is the acquisition of Portworx last year. Previously, Kubernetes storage was provided through Pure Storage Orchestrator (PSO), which is now retired as a standalone product and is integrated within Portworx, in a version called Portworx Essentials.

Portworx Essentials is the free version of Portworx. Pure Storage adapted it to run smoothly on FlashArray and FlashBlade and has removed limits that normally affect free users of Portworx Essentials, such as a maximum of five nodes. Portworx is built on PX-Store, which aggregates and pools storage capacity, and a series of advanced data management components that are part of the Portworx Data Services platform.

PX-Store is a modern, distributed, container-optimized storage solution with elastic scaling, storage-aware class-of-service, multi-writer shared volumes, local snapshot capabilities, and multiple failover options (node-aware, rack-aware, availability zone-aware). Local synchronous replication for data center high availability is also supported.

Compared to the fully-fledged Portworx, there are features that are not supported in Portworx Essentials, such as PX-DR and PX-Backup. Stateful snapshot capabilities are available, as is the ability to perform Cloudsnap backups to cloud storage. Auto-scaling groups are available and support AWS, GCP, and Azure. Encryption is available at the cluster level when customers bring their own key management system.

The integration of Portworx Essentials on Pure Storage controller-based architectures significantly enhances data efficiency because users benefit from the data reduction capabilities offered by the storage arrays, which are superior to those offered by the standalone Portworx solution.

Portworx Essentials is managed through PX-Central, just like Portworx standalone. It integrates with Pure Storage Pure1, which consumes telemetry data from Portworx and delivers best-in-class app-centric analytics and, eventually, recommendations.

Although Portworx Essentials may feel limited compared to the standalone Portworx product, it allows organizations to seamlessly deploy cloud-native workloads on a proven Kubernetes storage solution, and as their needs grow, they can effortlessly migrate those workloads to the fully-fledged Portworx solution if they decide to adopt it.

Strengths: Portworx Essentials allows Pure Storage customers to get acquainted with the full Portworx experience immediately and without additional cost. The solution delivers a consistent

experience that customers can reuse if they decide to adopt the full Portworx product. The product offers excellent data efficiency and management with good monitoring capabilities.

Challenges: Limited data management and security features, although this is understandable due to the overlap with the potential full Portworx implementation.

Storpool

StorPool has developed a robust, high-performance software-defined storage solution that can serve mixed workloads from a single StorPool system. Beyond bare-metal and virtualized workloads, StorPool also supports container deployments and other workload types through its own CSI plugin.

This plugin leverages most of the StorPool mechanisms used for snapshots and clones, although snapshots are not supported currently via the CSI plugin itself, so snapshot operations must be done through API calls.

Various snapshot replication topologies are supported, and cross-site replication is also possible. Snapshots can also be used for data protection, with a low delta on data transfers due to incremental snapshots.

Data efficiency capabilities strongly depend on StorPool native snapshot mechanisms, a combination of thin provisioning, zero detection, and the use of TRIM/discard features. Clones are created by directly pointing at existing data without creating any copies, which can help reduce space usage. Currently, compression, erasure coding, and deduplication are not implemented. Multi-tenancy capabilities rely on a combination of templates and policy-based placement of data and IOPS and throughput limits, making the solution suitable for use by solutions providers. StorPool provides in-flight data encryption for snapshot replication, while at-rest data encryption is possible only when using self-encrypting drives.

Management is achieved through StorPool's own cloud-based management platform, which consists of a user-friendly interface that includes comprehensive analytics and proactive alerting capabilities coupled with actionable insights.

Strengths: StorPool is one of the most underrated software-defined storage solutions available on the market, with rich enterprise features. Its support for Kubernetes workloads is a great advantage for StorPool customers, especially because the solution's focus on high-performance workloads makes it ideal for business-critical, performance-sensitive containerized applications.

Challenges: Advanced data services have a lot of potential but are not yet natively supported by the CSI plugin, and most integrations have to happen through HTTP/REST API calls. Data efficiency capabilities are currently insufficient.

VMware

VMware Tanzu is built on top of vSAN and so it can be used either in standard on-premises VMware vSphere environments with vSAN, or as a part of VMware Cloud Foundation (VCF). VCF offers a full hybrid cloud experience and vSAN constitutes VCF's storage foundation.

When Tanzu is deployed on vSAN, it allows the consolidation of traditional virtualized workloads and cloud-native applications on the same layer, and is therefore best for organizations already using vSAN in production environments. This mode allows storage to be provided to cloud-native workloads from the same storage clusters without any architectural changes. VMware also offers an additional deployment option via the vSAN Data Persistence platform (DPp). The Data Persistence platform is a framework for modern stateful service providers to use to build Kubernetes plugins or operators on, and for their underlying vSphere infrastructure. Stateful services running on the DPp can be deployed on a vSAN datastore with the vSAN *host-local shared-nothing* architecture (SNA) policy or in a second mode called vSAN Direct.

The first option, SNA policy, allows the application to control placement and take over the duty of maintaining data availability. The technology makes it easy for the persistent service to co-locate its compute instance and a storage object on the same physical ESXi host. With the host-local placement, it is possible to perform such operations as replication at the service layer and not at the storage layer. The second option, vSAN Direct, consists of dedicated hardware with optimal storage efficiency and near bare-metal performance. vSAN Direct allows modern stateful services to leverage the availability, efficiency and security features built into the modern stateful service layer and to have direct access to the underlying direct-attached hardware.

Part of Tanzu's strength derives from vSAN's Storage Policy-Based Management (SPBM) capabilities. Various storage policies can be created, each with different resilience requirements, capabilities (such as encryption), QoS (IOPS throttling), and so on. SPBM can be further expanded by organizations using existing API integrations to automate container provisioning workflows. Individual software vendors can integrate their application's native data management, replication, and service capabilities (such as app-level replication, erasure coding, encryption, and so on) directly into vSAN DPp to shift some of the storage policies at the application level and avoid resource waste.

Management of the Tanzu environment is handled through Tanzu Mission Control (TMC), which allows multi-cluster Kubernetes management on-premises and across clouds and draws on VMware's proven experience with alerting, monitoring, and analytics. Data migration and replication are available with Velero or VMware HCX.

The solution offers great security capabilities with software-based, in-flight and at-rest data encryption, FIPS 140-2 cryptographic modules, support for third-party KMIP-compliant key managers, and the ability to enable datastore-level encryption with a single click. RBAC is natively supported through vSphere and VCF.

Strengths: Tanzu is ideally suited for organizations with a strong VMware focus as they already have

all the building blocks in place to adopt Tanzu quickly and effortlessly, enabling a fast movement towards Day 2 operations. The solution is comprehensive, offers two deployment models, and will completely integrate into the enterprise landscape.

Challenges: Although very well architected, Tanzu's dependency on other VMware products creates a platform overhead that is unnecessarily complex for organizations looking for a pure cloud-native deployment model.

Weka

Weka's WekaFS solution makes it one of the leaders in high-performance primary storage, particularly with its strong presence in the world of HPC and AI where containerized workloads are broadly used and vendors need to be able to offer Kubernetes storage capabilities to their customers.

To do so, Weka partners with SUSE to deliver a Kubernetes storage stack based on Rancher, but with storage delivered by the WekaFS shared file system presented through a CSI plugin. WekaFS is also present in the Rancher application catalog and can be deployed directly from the catalog. And Weka also partners with Red Hat Openshift as well as Run.ai to further help its unique client base to use Kubernetes through Weka to its fullest.

The solution scales well and is performance-oriented with low latency, high throughput, and high IOPS capabilities. WekaFS is a hybrid solution that can be run on-premises or in the cloud.

WekaFS snapshots are supported only at the storage layer currently, as the capability is not yet available through the CSI plugin. The solution has a data protection capability named Snap2Object that allows snapshots to be taken in S3, and can enable immutable snapshots at the virtual filesystem level. Data mobility is enabled by creating a global namespace that includes both on-prem and cloud WekaFS filesystems, allowing workloads to be moved seamlessly without impacting operations. Weka is also available in AWS Marketplace and through Rancher, works with the AWS cloud.

Security features include in-flight and at-rest data encryption when customers use their own third-party key management system. Hashicorp Vault is also supported, as are role-based access control and LDAP/Active Directory services.

Multi-tenancy capabilities are tied strongly to the security implementation. WekaFS supports up to 64 logical tenants called "organizations," each with its own administrators, quotas, and separate data encryption keys.

For monitoring and alerting purposes, the solution provides fairly standard capabilities through Prometheus and Grafana.

Strengths: Outstanding performance on premises and in the cloud due to the WekaFS architecture, the solution also offers excellent security and multi-tenancy capabilities, along with some noteworthy

features such as Snap2Object and global namespace support.

Challenges: Abstraction of storage features through the CSI plugin is incomplete (lacking snapshot support, for example). Features aimed at improving application awareness, such as consistency groups, are not yet implemented.

6. Analyst's Take

CSI specifications have evolved a lot in recent years, and this is reflected in current Kubernetes storage solutions. Most storage vendors that chose a conservative approach last year are now developing their software rapidly to reflect evolving CSI specifications. This means that the integration between the storage platform and Kubernetes is moving at a brisk pace, and can now offer support for most of the features that are considered mandatory in an enterprise environment.

We identified three groups of vendors in this space, characterized by the level of sophistication of their approach and the features available on their platform. The most conservative ones—Datacore and Storpool—provide a basic CSI plug-in and expose limited functionality from the array. This severely limits the possibility of implementing Kubernetes for mission-critical environments, especially in hybrid cloud scenarios.

The second group of vendors consists of Commvault, Infinidat, IBM, HPE, VMware, and Weka. These companies take a more sophisticated approach that provides additional options. Though some enterprise features are still missing, the overall strategy of these vendors is much more aggressive, with clear roadmaps and best practices for overcoming the limitations imposed by current CSI specifications.

The leading group consists of three companies—Pure Storage, NetApp, and Dell Technologies—with products designed specifically with Kubernetes in mind that are able to overcome most of the limitations imposed by CSI, and even to add services on top of this functionality. Pure Storage, with its takeover of Portworx, emphasizes flexibility and a consistent user experience across different environments. It offers an extensive feature set that extends data storage, with data management functionality, to build consistent data services across different clouds.

NetApp evolved its Trident solution into Astra, which provides a promising and stable set of capabilities that let its customers build and maintain a compliant and fully featured Kubernetes infrastructure. Last, but certainly not least in this group, is Dell Technologies, which enables its customers to utilize traditional storage arrays as multi-tenant, feature-rich enterprise storage for modern apps running on Kubernetes. CSM gives Dell customers a highly available Kubernetes infrastructure with replication, multi-tenancy, and resiliency support, for a good Kubernetes experience.

That does not mean that these solutions are fully featured, because there are still challenges for all the companies that need to be addressed. With CSI and Kubernetes, as well as the storage arrays and software still in development, the market will keep moving at a rapid pace to provide the best possible environments for the customer and developers.

7 About Enrico Signoretti

Enrico has more than 25 years in technical product strategy and management roles. He has advised mid-market and large enterprises across numerous industries, and worked with a range of software companies from small ISVs to global providers.

Enrico is an internationally renowned expert on data storage—and a visionary, author, blogger, and speaker on the topic. He has tracked the evolution of the storage industry for years, as a Gigaom Research Analyst, an independent analyst, and as a contributor to the Register.

8 About Arjan Timmerman

Arjan Timmerman is an independent industry analyst and consultant with a focus on helping enterprises on their road to the cloud (multi/hybrid and on-prem), data management, storage, data protection, network, and security. Arjan has over 23 years of experience in the IT industry and worked for organizations across various verticals such as the Shared Service Center for the Dutch Government, ASML, NXP, Euroclear, and the European Patent Office to just name a few.

Growing up as an engineer and utilizing that knowledge, Arjan currently provides both technical and business architectural insight and management advice by creating High-Level and Low-Level Architecture advice and documentation. As a blogger and analyst at TECHunplugged.io blog, Gestalt IT, Amazic World, and other outlets, Arjan is also from time to time participating in podcasts, discussion panels, webinars, and videos. Starting at Storage Field Day 1 Arjan is a long-time Tech Field Day Alumni, former NLVMUG leader, and active member of multiple communities such as Tech Field Day and vExpert.

Arjan is a tech geek and even more important he loves to spend time with his wife Willy, his daughters Rhodé and Loïs and his son Thomas sharing precious memories on this amazing planet.

9 About Max Mortillaro

Max Mortillaro is an independent industry analyst with a focus on storage, multi-cloud & hybrid cloud, data management, and data protection.

Max carries over 20 years of experience in the IT industry, having worked for organizations across various verticals such as the French Ministry of Foreign Affairs, HSBC, Dimension Data, and Novartis to cite the most prominent ones. Max remains a technology practitioner at heart and currently provides technological advice and management support, driving the qualification and release to production of new IT infrastructure initiatives in the heavily regulated pharmaceutical sector.

Besides publishing content/research on the TECHunplugged.io blog, Gestalt IT, Amazic World, and other outlets, Max is also regularly participating in podcasts or discussion panels. He has been a long-time Tech Field Day Alumni, former VMUG leader, and active member of the IT infrastructure community. He has also continuously been running his own technology blog kamshin.com since 2008, where his passion for content creation started.

Max is an advocate for online security, privacy, encryption, and digital rights. When not working on projects or creating content, Max loves to spend time with his wife and two sons, either busy cooking delicious meals or trekking/mountain biking.

10. About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

11. Copyright