GigaOm Radar for Data Storage for Kubernetes

ENRICO SIGNORETTI

TOPIC: DATA STORAGE
GigaOm Radar for Data Storage for Kubernetes

TABLE OF CONTENTS

1  Summary
2  About the GigaOm Radar
3  Market Categories and Deployment Types
4  Key Criteria Comparison
5  GigaOm Radar
6  Vendor Roundup/Overview
7  Conclusion
8  About Enrico Signoretti
9  About GigaOm
10 Copyright
1. Summary

Kubernetes adoption is quickly accelerating, and enterprises are now in a transition phase. In the last few years, we have seen an increasing interest in container-based application development. As a result, IT organizations started to implement a proof of concepts and laboratories, which moved later to development and test platforms. In this period, the entire industry has matured, both in terms of the core technology (container formats and development tools) and orchestrators, with several companies trying to push their solutions (i.e., Docker Swarm, Mesos DC/OS, Google Kubernetes, and others).

Now that Kubernetes is the clear winner, the number of organizations moving to the production phase is finally growing as well. In most cases, Kubernetes infrastructures are still relatively small, and applications running on them are fairly simple, with limited data storage needs. On the other hand, more and more stateful applications are migrating to these platforms, requiring additional resources and performance. At the same time, enterprises of all sizes are embracing hybrid cloud strategies that are becoming more complex and structured. We are quickly moving from a first adoption phase where data and applications are distributed manually and statically in different on-premises and cloud environments to a new paradigm in which data and application mobility is the key for flexibility and agility.

Now, organizations want the freedom to choose where applications and data should run dynamically, depending on several business, technical, and financial factors. They choose the public cloud for its flexibility and agility, while on-premises infrastructures are still a better option from efficiency, cost, and reliability perspectives. In this scenario, it is highly likely that development and testing are made on the public cloud while production could be on-premises, in the cloud, or both, depending on the business, regulatory, economic, and technical needs of the particular enterprise.

Kubernetes is instrumental in executing this vision, but it needs the right integration with infrastructure layers, such as storage, to make it happen. Persistent and reliable data storage, alongside data management and security, are vital factors to consider when evaluating 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.
2. About the GigaOm Radar

HOW TO READ THIS REPORT

The Market Roundup/Landscape/radar profile is part of a series of documents aimed at giving the reader the tools to better understand a technology, evaluate it, and explore the market to find the best solutions for his organization.

In this context, and in order to get a complete view of the state of the solutions available in the market, the reader should consider the following documents:

**Key Criteria to Evaluate Data Storage for Kubernetes**, is an introduction to the technology, defines the necessary evaluation metrics, the key criteria to evaluate new solutions, and the impact of the latter on the former. It is dedicated to those end users that are approaching a new technology for the first time or want an update on the latest evolution.

**GigaOm Radar for Data Storage for Kubernetes**, offers a brief 360° view of the market including market and technical positioning of most notable vendors, a short introduction of their solutions and differentiations, including a high level graphic comparison of the vendors.

**Vendor Profiles for Key Criteria**, are easy-to-read deep dive documents that cover a single vendor regarding the solutions described in the other reports. They provide more details on the solution, how the vendor approached the key criteria, and the impact that its solution has on the evaluation metrics. This document helps end-users to get a quick but complete evaluation of a single vendor.
3. Market Categories and Deployment Types

For a better understanding of the market and vendor positioning (table 1), we categorized solutions for Kubernetes data storage by the target market segment (small-medium or large enterprise) and their architecture (enterprise storage systems with a container storage interface (CSI) plug-in, software-defined solutions with optimizations for containers, or cloud-native solution):

- **Small-medium enterprise**: In this category, we find solutions that meet the needs of very small businesses that can grow up to address those of medium-sized infrastructures. They can also be solutions adopted by large enterprises for departmental use cases without a very rich feature set, with limited data mobility and management capabilities, but are easy to use and deploy.

- **Large enterprise**: Usually adopted for larger and business-critical projects. Solutions in this category have a strong focus on data management and mobility features; they also provide additional features to improve security and data protection. Scalability is another big differentiator, including the ability to manage multiple clusters from the same user interface.

- **ISP/MSP**: Even though the number of solutions in the internet service provider/managed service provider (ISP/MSP) category is still pretty limited, they usually have the same characteristics as those of enterprises with an added focus on multi-tenancy and manageability.

- **Traditional storage arrays with CSI plug-ins**: This type of solution is the most common at the moment and, usually, is the first to be adopted by users. They are easy to deploy and allow the reuse of storage resources already in place, with a minimal initial investment. On the other hand, most of them are immature, limited in operational performance, data management, and scalability. Some of the systems in this category, not designed to cope with the number of back-end operations necessary to run a Kubernetes cluster, risk creating bottlenecks that can impact service level agreements (SLAs) of critical production environments.

- **Software-defined storage with optimizations for containers**: The flexibility of this type of solution limits some of the deficiencies of traditional arrays while keeping the storage infrastructure compatible with traditional workloads and applications. The optimizations allow users to adopt Kubernetes gradually while granting good SLAs both to traditional and next-generation environments.

- **Cloud-native solutions**: These solutions are expressly designed to work with containers and Kubernetes across on-premises and public cloud environments, in a hybrid and multi-cloud fashion. Usually architected around a set of core features focused on data management and mobility, with specific data services tailored for containers, they take advantage of storage resources local to each single node in the cluster, cloud storage, or traditional enterprise shared storage such as network-attached storage (NAS) and storage area network (SAN) infrastructures.
## Table 1: Vendor Positioning

<table>
<thead>
<tr>
<th>Vendor</th>
<th>SMALL/MED ENTERPRISE</th>
<th>LARGE ENTERPRISE</th>
<th>ISP/MSP</th>
<th>TRADITIONAL STORAGE + CSI</th>
<th>SDS + OPTIMIZATIONS</th>
<th>CLOUD NATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAcore</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>DATARA</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>DELL EMC</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DIAMANTI</td>
<td>-</td>
<td>+++</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>HITACHI VANTARA</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IBM</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>INFINDAT</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>MAYA DATA</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>NETAPP</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PORTWORK</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>PURE STORAGE</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RED HAT</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>STORAGEOS</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+++</td>
</tr>
</tbody>
</table>

**Legend:**
- +++: Strong focus and perfect fit of the solution
- ++: The solution is good in this area, but there is still room for improvement
- +: The solution has limitations and a narrow set of use cases
- -: Not applicable or absent
4. Key Criteria Comparison

Following the general indications introduced with the Key Criteria for Evaluating Hybrid Cloud Data Protection Report, table 2 quickly summarizes how each vendor included in this research performs in the areas that we consider differentiating and critical for modern data protection. 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 2: Key Criteria & Evaluation Metrics Comparison

By combining the information provided by table 1 and table 2 the reader should be able to get a clear
idea of the market and technical solutions available in the field.
5. GigaOm Radar

All the Key Criteria and the critical feature impact analysis consolidate in the following graphic representation: The GigaOm Radar (figure 1). This vector-based graphics gives an overall perspective on all the vendors included in this research in terms of technical capabilities and features (table 2), execution on the vision, regardless of their market share or segment (table 1).

*Figure 1: GigaOm Radar for Data Storage for Kubernetes*
HOW TO READ THE GIGAOM RADAR

The GigaOm Radar focuses on technical capabilities, roadmap, execution, and the ability to innovate. In this graphic, each vendor representation of the market, there are three fundamental informational points:

- **The current position** in the chart gives information on the state of the solutions provided by the vendor.
- **The direction** indicates where the vendor is going with the product strategy and development.
- **The module** signifies how fast it is executing on vision and strategy.

The metrics on which the GigaOm Radar focuses:

- **Maturity:** Expresses the maturity and solidity of the product, user acceptance of the solution, and overall ecosystem sustainability.
- **Innovation:** The differentiation of the solution, innovative technical aspects, and overall vendor approach to the market.
- **Feature Play:** Gives indications about the focus of the solution in terms of single differentiating features and technical aspects of the product. It is usually including new vendors that are defining new product categories, niche players, and point solutions.
- **Platform Play:** Contrary to “Feature Play,” solutions positioned in this area provide a horizontal platform that can face a broader range of challenges, with a comprehensive feature set and an extensive ecosystem.

The GigaOm Radar organizes in three concentric areas and, the closer to the center, the better:

- **Leaders:** All the vendors that are competing for market leadership for the metrics described above, even if they are competing in different market segments.
- **Challengers:** All the vendors, with the potential to become a leader soon, niche, or traditional players with an established market, and other companies that have interesting solutions but still have to prove their real differentiation.
- **New entrants:** Usually companies with a limited feature set and too young to be included in inner areas, but with potential to move there soon.
- The central area of the GigaOm Radar is usually empty. It is reserved for extremely mature and consolidated markets with very few vendors remaining and solutions that are mature, comparable, and without much space for further innovation.

The market share metric is excluded from the GigaOm Radar to give a clearer view of the potential of the vendors. It is a forward-looking market landscape representation that leaves more room for innovation and differentiation instead of weighting established positions.
6. Vendor Roundup/Overview

DataCore

DataCore has a two-fold strategy about persistent storage for Kubernetes. One on hand, they are developing a CSI plug-in for its block and file storage products aimed at servicing its current customers. On the other hand, they are investing in MayaData, which is the sponsor behind OpenEBS, an open source solution designed for Kubernetes storage. This approach will help the company to give answers immediately while having access to intellectual property for the development of future solutions.

**Strengths:** The CSI plug-in has a simple and efficient design, aligns with the latest CSI specifications. Easy to adopt and without additional license fees.

**Weaknesses:** CSI specifications are still immature and can not take advantage of all the potential offered by the DataCore platform.

Datera

Datera Data Services Platform is a flexible, scale-out, software-defined storage solution that bridges the gap between legacy and modern infrastructures. It is integrated with Kubernetes CSI and aligns well with the latest API specifications.

**Strengths:** Storage Classes are implemented very well, enabling the DevOps team to take full advantage of the backend resources and flexibility of the solution.

**Weaknesses:** Advanced data services (data protection, remote replication, disaster recovery (DR)) are not yet implemented due to CSI limitations.

Dell EMC

CSI-Plug-ins are now available for all Dell EMC storage systems, and the company is actively working to improve them. The implementation is still basic, and there are not features aimed at simplifying the deployment and management of Kubernetes storage in large infrastructures.

**Strengths:** Simplicity, free plug-ins, and ease of adoption for Dell EMC customers allow users to take advantage of existing infrastructures and start quickly without further investments.

**Weaknesses:** The solution is still immature and does not allow planning for hybrid cloud infrastructures, no disaster recovery options, and no official integration yet with Prometheus or other Kubernetes monitoring platforms.
Diamanti

Diamanti offers an end-to-end solution that resembles hyper-converged infrastructures in the virtualization space. It comes with a fully supported standard Kubernetes distribution pre-installed, or the customer can choose Red Hat OpenShift instead.

**Strengths:** High resilience, performance, and ease of deployment and management are the most important characteristics of this platform, which also offers several features to overcome the limitations of CSI and Kubernetes networking. Thanks to the support of custom resource definitions (CRDs), Diamanti can seamlessly replicate data to the public cloud for migrations or DR.

**Weaknesses:** The solution is relatively expensive and focused on large enterprise deployments. Even though its basis is on commodity hardware and very efficient, Diamanti offloads many storage and network operations to the hardware, limiting hardware choice.

Hitachi Vantara

Hitachi Vantara offers a CSI plug-in for all its storage systems. The plug-in is available for all Hitachi Vantara customers and provides a quick path to adopt Kubernetes for Hitachi Vantara customers.

**Strengths:** Hitachi has forward a looking-vision around containers, and Kubernetes in particular, which is aligned with its current product portfolio regarding data analytics and IIoT while protecting existing infrastructure investments.

**Weaknesses:** Current CSI plug-in is still immature and lacks some features such as remote replication for DR. This is due to the company strategy, focused on adhering to CSI specifications and APIs without adding non-standard functionalities.

IBM

IBM offers CSI plug-ins for all its major storage platforms. The plug-in is open source and free to download and install.

**Strengths:** IBM customers can take advantage of the installed base to start quickly with their Kubernetes deployments.

**Weaknesses:** Plug-ins are immature, and IBM storage products do not allow for planning for hybrid cloud infrastructure deployments.
Infinidat

One of the main advantages offered by Infinidat is the possibility to consolidate a large number of applications, workloads, and data in very few systems. In this context, Infinidat provides a compelling solution for customers that want to consolidate Kubernetes applications alongside others.

**Strengths:** The new CSI plug-in aligns with the latest specifications. Additionally, Infinidat simplifies disaster recovery and data migrations across on-premises and major cloud providers thanks to its Neutrix Cloud service.

**Weaknesses:** Even though data volumes for Kubernetes can group for monitoring in the UI, there is not a specific integration with Prometheus yet.

Maya Data

Maya Data OpenEBS Enterprise bundles enterprise functionalities and support to the OpenEBS open source project, including a series of tools to improve data migration, mobility, visibility, and infrastructure hardening. Users and developers can adopt OpenEBS at no cost, while Maya Data offers a very flexible licensing.

**Strengths:** Maya Data offers a no lock-in approach for users that want to deploy an open-source storage solution for Kubernetes in a multi-cloud environment.

**Weaknesses:** The solution offers tools to control and migrate data between Kubernetes cluster but it has some limitations with its current implementation, limiting its potential for high demanding enterprise use cases. Future releases of the product will address this issue.

NetApp

NetApp is building Trident, a complete storage orchestration platform to address Kubernetes challenges. This allows simplifying data storage provisioning and management for DevOps teams while offering a consistent user experience across private and public cloud deployments.

**Strengths:** NetApp’s overall strategy is solid, and Trident is a software component that fits very well in NetApp’s Data Fabric vision. The solution shows a good feature set and an interesting roadmap, contributing to a good ROI on the NetApp solution.

**Weaknesses:** Trident is fully CSI compliant but still misses some advanced features that would simplify data migrations, remote replication, disaster recovery orchestration, and backup.
Portworx provides one of the most compelling solutions for enterprise storage dedicated to Kubernetes infrastructures. A cloud-native architecture combined with unique data services that enable enterprise organizations to deploy business and mission-critical applications without the limitations imposed by traditional solutions.

**Strengths:** Portworx Enterprise allows simplifying most of the operations that are currently limiting enterprise Kubernetes deployments for stateful applications, improving data protection and management processes with a positive impact on overall infrastructure TCO. It can work with local storage installed on cluster nodes or traditional enterprise shared storage.

**Weaknesses:** Most storage vendors offer free CSI plug-ins and backend resource orchestration, allowing enterprises to start with minimal or no investment.

Pure Storage demonstrated again its ability to propose solutions with the right combination of performance and usability that adopts a broad range of use cases. From this point of view, Pure Storage Orchestrator has the right characteristic to allow a smooth adoption of Kubernetes that leverages existing storage resources and minimizes initial investment.

**Strengths:** PSO is a good solution and has the potential to become even more, a key differentiator for Pure Storage when the end-user is evaluating storage solutions ready to support containers alongside virtualized and physical systems.

**Weaknesses:** The solution is not ready for all use cases, especially when backup and remote replication for disaster recovery are involved. Pure Storage has best practices in place to mitigate this limitation while working on the right implementation following the evolution of CSI specifications.

Red Hat

Red Hat OpenShift Container Storage (OCS), based on Ceph, has been designed for simplicity and ease of use. Integrated with OpenShift, all management operations are automated using Rook (a CNCF orchestration tool for storage) and automated, enabling DevOps teams with little or no storage knowledge to get a simplified hyperconvergence-like experience.

**Strengths:** Simplicity and ease of use of the solution, completely integrated with the Red Hat OpenShift Container Platform. The solution can be deployed on-premises as well as in the cloud. Simple support subscription model aligned with OCP support and licensing.

**Weaknesses:** Rigidity of the configuration and limited scalability can pose some risks for large
deployments, and the customer may be forced to replace OCS with a standard Ceph installation.

StorageOS

StorageOS is a cloud-native solution designed explicitly for Kubernetes. As such, the product is more flexible, easier to manage, and integrate with the Kubernetes platform than traditional storage systems. It can deploy both on-premises and in the cloud, providing the same functionalities in different environments.

Strengths: Good, lightweight, and efficient architectural design. StorageOS offers development licenses (up to 500GB of storage)

Weaknesses: StorageOS still misses data services features (e.g., snapshots and remote replication) and DR orchestration, limiting the possibility to use it in complex environments that need advanced data protection, data migrations, and fast cloning.
7. Conclusion

CSI specifications are still immature, and this reflects in the products proposed by most vendors. Many storage vendors have chosen a conservative approach that follows the development of CSI specifications. This approach means that the integration between the storage platform and Kubernetes is limited, offering scarce support for all those features that are usually 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 provide a basic CSI plug-in and expose limited functionality from the array. In this group, we find Dell EMC, Hitachi Vantara, and IBM. This severely limits the possibility of implementing Kubernetes for mission-critical environments, especially in hybrid cloud scenarios.

The second group of vendors (Red Hat, Infinidat, Datera, Pure Storage, NetApp, DataCore, StorageOS) opted for a more sophisticated approach that gives additional options to the end-user. Some enterprise features are still missing, but the overall strategy of these vendors is much more aggressive, with clear roadmaps and best practices to overcome the limitation imposed by current CSI specifications. Two of particular note in this group are NetApp and Pure Storage; their approach is more holistic and offers better flexibility in large infrastructure and hybrid cloud environments. Maya Data is worth a mention as well because of the open-source core (OpenEBS) and the innovative tools and services included in the enterprise support subscription.

The leading group is now composed of a couple of startups (Diamanti, Portworx), with products designed specifically for Kubernetes and able to overcome the limitations imposed by CSI. In this group, it is worth noting the different approaches. Diamanti is a hyper-converged solution aimed at building Kubernetes cloud-like experience for large enterprises, with high performance and ease of use as its primary characteristics. In contrast, Portworx is more focused on flexibility and consistent user experience across different environments, offering a very extensive feature set that extends data storage with data management functionalities aimed at building a consistent data services layer for Kubernetes that spans across different clouds.
8. About Enrico Signoretti

Enrico has 25+ years of industry experience in technical product strategy and management roles. He has advised mid-market and large enterprises across numerous industries and software companies ranging from small ISVs to large providers.

Enrico is an internationally renowned visionary author, blogger, and speaker on the topic of data storage. He has tracked the changes in the storage industry as a Gigaom Research Analyst, Independent Analyst and contributor to the Register.
9. 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.
10. Copyright

© Knowingly, Inc. 2020. "GigaOm Radar for Data Storage for Kubernetes" is a trademark of Knowingly, Inc. For permission to reproduce this report, please contact sales@gigaom.com.