GIGAOM



FEATURE PLAY

PLATFORM PLAY

Challengers

Leaders

VMware
Red Hat
Arrikto
NetApp
Ionir
Diamanti
StorageOS
Suse
Portworx
Robin.io
MayaData

◄─■ Forward Mover

GIGA OM

Enrico Signoretti, Max Mortillaro
Jun 25, 2021

# GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0

Data Storage

# GigaOm Radar for Cloud-Native Kubernetes Data Storage

## Table of Contents

# 1. Summary

This report focuses on cloud-native persistent storage solutions for Kubernetes. These are novel architectures specifically designed to address the needs of cloud-native applications without compromising on performance or scalability. They are usually not engineered to co-exist with other workload types, such as virtualization.

The adoption of cloud-native development frameworks and application refactoring initiatives continues to fuel demand for Kubernetes persistent storage. Organizations understand that the benefits of cloud-native workloads in terms of performance, scalability, and portability are key enablers to achieve business goals.

During the COVID-19 pandemic, enterprises already running cloud-native workloads met the abrupt challenges brought forth by lockdown measures with relative ease: Application portability enables frictionless workload movement from the data center to the cloud, and even across clouds, providing greater flexibility and responsiveness to business requirements than using legacy technologies.

Since our last report, evaluated solutions have evolved at an incredibly rapid pace. Some vendors have been adding broad sets of enterprise-grade features to their products, while others have focused on supporting specific use cases.

Many users are adopting Kubernetes solutions after trying traditional enterprise storage first. The latter is usually considered a safe bet for first deployment but it can't cope with the sheer number of backend operations required by Kubernetes at scale. This, together with the complexity involved in managing multi-cloud environments with traditional storage, encourages users to look for smarter and more efficient alternatives.

When compared to other types of systems, Kubernetes-native storage offers an environment that is more DevOps-friendly, helping to build a hardware stack that can be controlled by the operations team while enabling developers to allocate and monitor resources quickly, in a self-service fashion, when necessary. This is a major boon for enterprise IT organizations looking for the smartest way to evolve their processes and align them with the latest business and technology requirements.

# 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.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

4

# 2. Market Categories and Deployment Types

For a better understanding of the market and vendor positioning (**Table 1**), we assess how well solutions for cloud-native Kubernetes data storage 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: SaaS (cloud-only) and hybrid and multi-cloud.

- **SaaS:** The solution is available only in the cloud, as a managed service. Often designed, deployed, and managed by the service provider or the storage vendor, it is available only from that specific provider. The big advantages of this type of solution are its simplicity and the integration with other services offered by the cloud service provider.

- **Hybrid and multi-cloud:** These solutions are meant to be installed both on-premises and in the cloud by users, allowing them to build hybrid or multi-cloud storage infrastructures. Integration with a single cloud provider might be limited compared to the other option, and more complex to deploy and manage. On the other hand, these solutions are more flexible, and the user usually has more control over the entire stack in terms of resource allocation and tuning.

Note that GigaOm is publishing another Radar report on Kubernetes storage, focused on general-purpose enterprise storage systems that include software-defined storage and physical controller-based architectures. Enterprise Kubernetes storage allows organizations to leverage existing deployed storage platforms to deliver persistent storage capabilities without having to architect new solutions. These solutions are mostly suited to mixed-workload environments or large data centers with a sizable investment in storage infrastructure.

*Table 1. Vendor Positioning*

| | MARKET SEGMENT | | | DEPLOYMENT MODEL | |
|---|---|---|---|---|---|
| | Small/Med Enterprise | Large Enterprise | ISP/MSP | SaaS | Hybrid |
| Arrikto | +++ | ++ | – | – | +++ |
| Diamanti | + | +++ | ++ | – | +++ |
| Ionir | +++ | ++ | – | – | +++ |
| MayaData | ++ | +++ | +++ | – | +++ |
| NetApp | ++ | +++ | + | +++ | +++ |
| Portworx | +++ | +++ | +++ | – | +++ |
| Red Hat | + | +++ | +++ | +++ | +++ |
| Robin | +++ | +++ | +++ | – | +++ |
| StorageOS | +++ | ++ | ++ | – | +++ |
| Suse | +++ | ++ | ++ | – | +++ |
| VMware | +++ | +++ | +++ | +++ | +++ |

+++ 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

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

6

# 3. Key Criteria Comparison

Building on the findings from the GigaOm report, "Key Criteria for Evaluating Cloud-Native Kubernetes Data Storage," **Table 2** summarizes how each vendor included in this research performs in the areas that 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 | Multi-tenancy | Security | Monitoring, Alerting, Analytics |
| Arrikto | +++ | ++ | ++ | +++ | +++ | ++ |
| Diamanti | +++ | +++ | +++ | +++ | +++ | ++ |
| Ionir | +++ | +++ | ++ | ++ | ++ | ++ |
| MayaData | ++ | +++ | ++ | ++ | ++ | ++ |
| NetApp | ++ | +++ | ++ | +++ | +++ | +++ |
| Portworx | +++ | +++ | ++ | +++ | +++ | +++ |
| Red Hat | ++ | +++ | +++ | ++ | +++ | ++ |
| Robin | +++ | +++ | ++ | +++ | +++ | ++ |
| StorageOS | ++ | +++ | ++ | ++ | +++ | ++ |
| Suse | ++ | +++ | ++ | ++ | ++ | ++ |
| VMware | ++ | ++ | +++ | +++ | +++ | ++ |

+++ 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

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

7

*Table 3. Evaluation Metrics Comparison*

| | EVALUATION METRICS | | | | |
|---|---|---|---|---|---|
| | Architecture | Scalability | Flexibility | Efficiency | Manageability and Ease of Use |
| Arrikto | +++ | ++ | +++ | ++ | +++ |
| Diamanti | ++ | +++ | ++ | +++ | +++ |
| Ionir | +++ | +++ | +++ | +++ | ++ |
| MayaData | +++ | +++ | ++ | + | + |
| NetApp | +++ | +++ | + | +++ | +++ |
| Portworx | +++ | +++ | +++ | ++ | +++ |
| Red Hat | ++ | +++ | ++ | +++ | +++ |
| Robin | +++ | +++ | +++ | ++ | ++ |
| StorageOS | +++ | +++ | ++ | ++ | ++ |
| Suse | +++ | ++ | ++ | + | ++ |
| VMware | ++ | ++ | +++ | ++ | +++ |

+++  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.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

8

# 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.
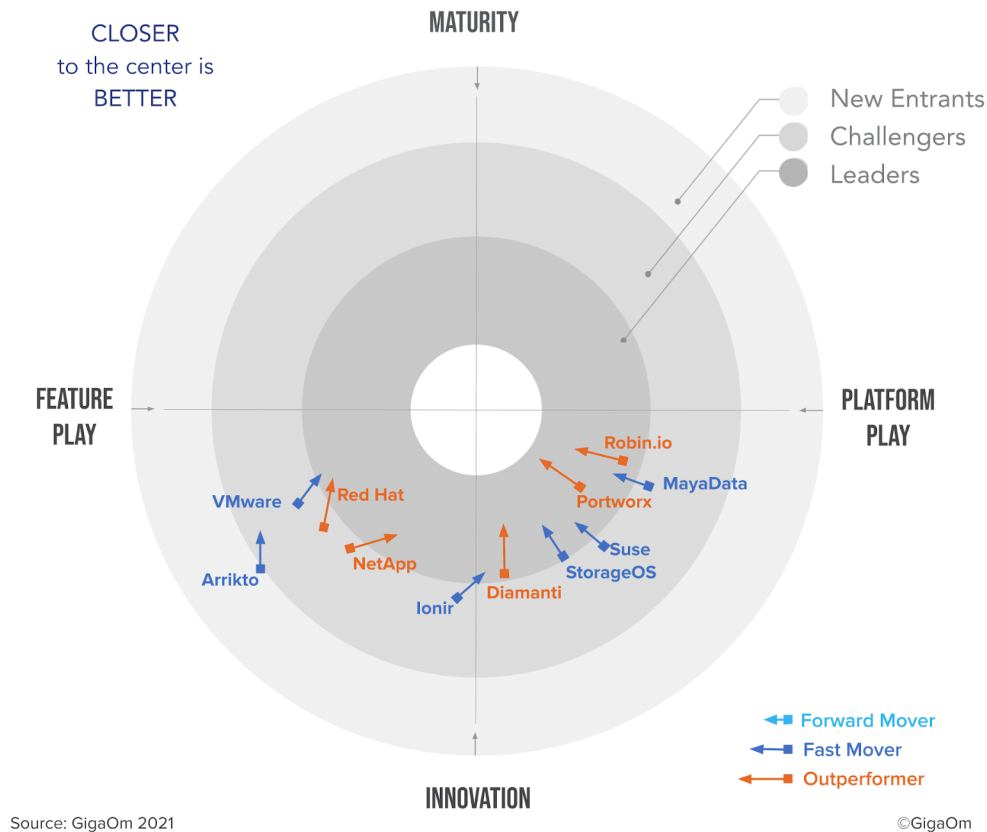


*Figure 1. GigaOm Radar for Cloud-Native Kubernetes Data Storage*

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 providing an arrow that projects each solution's evolution over the coming 12 to 18 months.

As you can see in the Radar chart in **Figure 1**, cloud-native Kubernetes data storage solutions are very innovative and dynamic, which explains why there are no vendors at all in the upper part of the diagram, which usually denotes mature solutions.

In the lower right corner are six vendors with a platform approach. At the top of the pack are two with a strong enterprise approach to Kubernetes storage. Portworx, now owned by Pure Storage, leads this year's radar thanks to the completeness of its platform, strong enterprise-grade capabilities, and

excellent manageability. Robin.io, a platform with outstanding advanced data services and an application-centric approach to containers, follows closely.

Trailing slightly are Diamanti and StorageOS. Diamanti combines high resilience, performance, and simplicity of use with very good enterprise-grade features. StorageOS delivers high-performance Kubernetes storage for business-critical applications, thanks to its robust and scalable architecture.

Still in the lower right quadrant, two challengers are on a good course to join the leaders. Interestingly, both are open source projects. Longhorn, originally developed by Rancher Labs and recently acquired by SUSE, is now a CNCF sandbox project. This solution offers a flexible and performance-oriented architecture with growing advanced data services capabilities. The other one, OpenEBS, was originally developed by MayaData and gets contributions both from this company and a vibrant developer community. OpenEBS is a versatile solution with multiple deployment options and a focus on performance. Both solutions are built on solid and scalable architecture with great performance. However, enterprise-grade capabilities are not yet as developed as those of the other contenders in this section of the radar.

Ionir is a vendor with an intriguing trajectory: the company is crossing from feature play to platform play, and is on its way to becoming a leader. This solution has a strong focus on advanced data services and superior data management and data mobility capabilities due to its unique storage architecture and replication features.

The lower left corner of the radar shows five vendors, each of which are oriented more toward specific features, deployment models, or niche use cases.

VMware is a leader in this sector with VMware Tanzu. An excellently engineered solution, it offers the most convenient path to cloud-native Kubernetes storage for enterprises with a strong VMware footprint and large on-premises infrastructure, although Tanzu also runs on the cloud. Tanzu offers strong enterprise-grade features and excellent management capabilities that provide the fastest path to Day 2 operations for large enterprises.

Running alongside VMware, Red Hat OpenShift Data Foundation (ODF) is an enterprise-grade solution that boasts a rich set of capabilities. Unfortunately, ODF only supports OpenShift and therefore lacks the breadth of scope required to be positioned in the right (Platform Play) sector. Nevertheless, OpenShift users will find a compelling storage solution in ODF.

NetApp is interesting. Astra, its cloud-native as-a-service Kubernetes storage offering, shines in many areas. Built on Trident and Cloud Volumes, the solution is not only robust but offers great capabilities and an application-centric architecture. Astra is great for organizations seeking excellent performance, although the lack of support for non-NetApp underlying storage is currently a limiting factor.

Arrikto is a solution focused on MLOps and machine learning models that offers an innovative architecture and strong performance for Kubeflow. Among the highlights of the solution are advanced

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

10

snapshot capabilities that greatly reduce data staging times for machine learning training and tuning activities.

## 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

## Arrikto

Arrikto Rok and Rok Registry together form an innovative data management automation and collaboration solution for Kubernetes. This solution enables high-performing, instance-based NVMe drives to become persistent storage on the worker nodes, and is coupled with a transparent object storage back-end for snapshots, versioning, and a globally federated distributed collaboration platform. The company is becoming laser-focused on MLOps and accelerating ML models into production on Kubeflow.

**Strengths:** An innovative approach to data storage and management for Kubernetes that combines the performance of a local disk with the advanced data services enterprises expect, without inserting a new storage layer. The solution can be deployed on-premises and in the public cloud.

**Challenges:** Arrikto's approach is oriented toward ML workflows, making the solution suitable for companies deploying machine learning models at scale.

## Diamanti

Diamanti's end-to-end solution targets enterprise-grade stateful applications and is based on two products: Spektra and Ultima.

Spektra is a container management plane that allows the management of Kubernetes clusters across clouds and locations (including core and edge), adding application and data mobility plus advanced data services, infrastructure observability and control. Ultima is the data plane part of the solution, which converges networking and storage and can run on-premises or in the cloud.

The Ultima data plane consists of a distributed storage platform that also provides L2/L3 networking capabilities, data protection features, container and VM support, and CNI/CSI plugins.

Ultima's capabilities in terms of storage and networking acceleration can be optionally augmented through the use of a hardware offload card called "Ultima Offload." Customers who opt for this solution will see significant performance improvements as well as better data efficiency through thin provisioning and data footprint optimizations, as well as reduced CPU usage on Kubernetes nodes. This approach is similar to that of cloud providers; for example, AWS and its Nitro PCIe cards

The solution comes with enterprise-grade features for both advanced data services and security. Data can be mirrored across availability zones; backup and restore as well as disaster recovery are supported across clusters and clouds; and volumes can be migrated across clouds. Diamanti supports role-based access control and multi-tenancy (with Spektra), allowing policy-based isolation between tenants and teams. Those features are also a plus for MSPs considering delivering Kubernetes-as-a-

service to their clients. Data-at-rest encryption is supported at the volume and disk levels, and an advanced, built-in key management system is also provided.

Diamanti has a feature-rich management platform that allows organizations to manage multiple clusters across various clouds. It embeds cluster and application lifecycle management capabilities to enable faster application deployments. The management platform also integrates granular observability capabilities, providing an overall view of the environment's health state and digging all the way down to the container level.

Diamanti has been gradually moving toward a software-only approach, enabling customers to run their platform on commodity x86 hardware through several OEM partnerships. The only component that still remains proprietary is the Ultima Offload card.

**Strengths:** High resilience, good performance, and ease of deployment and management remain important capabilities of this solution. An excellent set of enterprise-grade features make the solution ideal for larger organizations seeking robust data protection mechanisms. The initial expensive entry cost of the solution has been partially addressed by moving to a software-based deployment model.

**Challenges:** The architectural decision to offload storage and network operations to a dedicated hardware card limits hardware choice and can be a roadblock to broader adoption.

## Ionir

Ionir is a container-native, software-defined storage solution for Kubernetes with advanced data capabilities. The solution consists of an elastic and scalable distributed microservices architecture, which implements a CSI plugin that supports volume provisioning and snapshot management. Ionir uses NVMe over TCP, as well as the Intel SPDK framework to provide an efficient I/O path and avoid performance bottlenecks.

Ionir's metadata is based on a proprietary, patented database that records metadata on each write operation along with a name associated with the content of data and the time of the write. The timestamped record allows retrieval of the state of a volume from any point in time in the past at the granularity of one second; this, in effect, translates to continuous data protection.

Ionir leverages the unique metadata architecture to deliver advanced data management services, such as replication, migration, and disaster recovery. The solution allows persistent volumes to be copied or moved across clusters or even globally across clouds in 40 seconds or less, making it ideal for data migration or replication in time-sensitive environments.

Ionir also provides excellent data efficiency: Users can expect inline data deduplication augmented with compression, and thin provisioning. Erasure coding is not yet supported but is planned in a later release.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

13

From a security perspective, data-in-flight encryption is supported. Data-at-rest encryption is on the roadmap, leveraging the company's IP to avoid potential conflicts with data deduplication .

Ionir has an easy-to-use management interface that handles all of the supported activities, including snapshot clone operations and granular data restores. Monitoring is handled through Prometheus and Grafana, while ELK handles logging, tracing, and visualization of log events. Ionir allows customers to enable these tools through a simple one-click deployment process.

**Strengths:** Ionir is an excellent Kubernetes cloud-native storage solution with a strong focus on advanced data services, great efficiency, and superior data management and data mobility capabilities, making it particularly suited for enterprises with highly distributed cloud-native applications requiring data mobility and data protection.

**Challenges:** The solution currently has a few gaps in some areas, such as security (lack of at-rest encryption) and multi-tenancy (no QoS or throttling capabilities).

# MayaData—OpenEBS MayaStor

MayaData is a company that contributes to the development of OpenEBS, an open source Kubernetes cloud-native storage solution with one of the biggest active communities. MayaData provides professional support for this solution as well as additional paid features.

OpenEBS runs in containers and thus scales along with clusters. It has the ability to run on-premises and also on multiple cloud platforms. To demonstrate the performance capabilities and optimizations of the OpenEBS architecture, MayaData teamed up with Intel to test its performance with Intel Optane SSDs. The results were outstanding.

OpenEBS has multiple local persistent volume provisioners that customers can choose from based on the kind of workloads they need to support. One is OpenEBS LocalPV Hostpath, which allows applications to use StorageClass, PersistentVolumeClaims, and Persistent Volumes. Among these, Hostpath provides data protection integrations with Velero Backup and Restore and allows full integration with application code through YAML. Three other options exist—Device LocalPV (using local block devices), ZFS, and LVM backed storage.

The solution is highly modular, and organizations that want to leverage OpenEBS will have to identify whether suitable plugins are available to meet their requirements, such as to create application-consistent snapshots or implement data-at-rest encryption. In terms of advanced data services, good replication capabilities (across multiple availability zones) are worth noting, although only synchronous replication is currently available.

Optimization and efficiency both need improvement, as currently only the ZFS LocalPV deployment model provides compression and deduplication capabilities. While this is not a concern in smaller deployments, it can become a pain area for larger clusters.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

14

Because this is an open source solution with a strong focus on seasoned Kubernetes practitioners, it provides only basic monitoring integrations and a rudimentary graphical user interface. This may inhibit adoption among traditional enterprises.

Finally, it's worth noting that DataCore Software recently invested in MayaData and it also has a team of developers working for MayaData in OpenEBS development.

**Strengths:** A versatile and performance-oriented open source solution with professional support capabilities, multiple deployment choices, and an active contributor community.

**Challenges:** Capabilities are lacking in a number of areas. The use of plugins can cover some of these gaps, but this increases operational complexity. Increased focus on enterprise-grade features such as data protection, security, and manageability is necessary for the solution to keep pace with competitors.

# NetApp

NetApp Astra is a fully managed cloud-native Kubernetes storage and application-aware data management solution built on project Trident and NetApp Cloud Volumes as the persistent storage provider. Trident is NetApp's open source persistent storage provisioner and orchestrator for Kubernetes, while Cloud Volumes offers enterprise-grade, high-performance storage on public clouds. To develop Astra, NetApp took Trident's core technology and added an application-centric data management layer around it, using Cloud Volumes to store data, with both foundational blocks entirely transparent to users. Organizations can use Astra within and across public clouds as a fully managed service and, soon, on-premises on top of NetApp ONTAP storage arrays.

The solution supports automatic storage provisioning, application discovery, application-aware snapshots, and backups (scheduled and on-demand) for application data protection and business continuity, as well as cross-cluster/multi-cloud migration activities using application cloning. Backups can be locally stored in Cloud Volumes, or can take place out-of-band on different clouds, in which case customers have to handle their own cloud storage costs.

An interesting feature is the ability to detect certain popular applications (such as MySQL, Jenkins, and the like) and determine which components are in scope of the application. This allows Astra to not only undertake application-level operations (with all related containers) but also to take application-consistent backups and snapshots.

Astra comes with a clean and usable management interface that shows users all the information they need or the actions they can perform. Multi-tenancy is supported as well, with RBAC support and access granularity at the application level. At-rest encryption is supported when ONTAP is used as the storage provider as well as Cloud Volumes, with keys that are managed by the Cloud Volumes Service.

Besides a free tier, the solution is consumed under a simple "pay as you go" per application model. The solution will appeal to organizations that seek the simplicity of a fully managed solution and do not want to deal with frequent upgrades.

**Strengths:** Astra combines simplicity of use with a SaaS consumption model and advanced data services around an application-centric architecture.

**Challenges:** The subscription-based model can be unappealing to CapEx-oriented organizations. Currently, only storage from NetApp is supported.

## Pure Storage—Portworx

Acquired last year by Pure Storage, Portworx remains one of the most compelling and advanced solutions for enterprise Kubernetes storage. 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 modern, distributed, container-optimized cloud-native storage 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.

The solution offers broad deployment choices and supports bare metal and virtualized environments, including Pure Storage physical arrays, existing cloud block services, and cloud-based Kubernetes services, as well as those from other ecosystem partners, providing a consistent experience across infrastructures, platforms, and locations.

Portworx includes a comprehensive set of advanced data services, which is definitely one of the product hallmarks. PX-Backup handles data protection and supports application-consistent backups that are Kubernetes-complete; that is, not only is the data backed up, but also the entire application state, including all objects, application configuration data and dependencies. Granularity is provided, allowing organizations to back up either individual applications or thousands of applications and namespaces, and to define schedule policies as required. Restores can be performed locally or on any cloud.

PX-DR (an add-on module) expands those capabilities to provide best-in-class disaster recovery and data replication capabilities. It supports multi-site synchronous replication and zero RPO disaster recovery within a metro area, and multi-site asynchronous replication for cross-WAN connections. PX-Migrate handles multi-cloud and multi-cluster app migrations, as well as snapshots and application-consistent snapshots to the cloud.

PX-Secure constitutes the security layer of the Portworx solution, offering cluster-wide encryption; granular, container-based or storage class encryption (available when organizations bring their own

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

16

key management system); role-based access control, authorization and ownership mechanisms, as well as integration with Active Directory and LDAP through OIDC.

Finally, PX-Autopilot orchestrates automated space reclamation activities at the container volume level as well as on entire storage clusters through resizing activities, with the goal of keeping storage costs under control.

The solution is managed via PX-Central, a comprehensive management plane that handles multi-cluster management, CLI capabilities, proactive centralized monitoring, and cluster installation and setup functions. Integration with Pure Storage Pure1 allows this platform to consume telemetry data from Portworx and deliver app-centric analytics and, eventually, recommendations.

From an efficiency perspective, the solution handles compression for all snapshots, but true data reduction is achievable only when Portworx uses an underlying enterprise-grade platform with built-in data efficiency capabilities such as Pure Storage FlashArray.

Portworx can be deployed as a standalone solution, Pure Storage completely overhauled its container orchestration platform on physical storage arrays, and Pure Storage Orchestrator is now superseded by Portworx Essentials, which increases the maximum number of nodes and storage limits.

**Strengths:** A complete enterprise-grade solution with outstanding data management capabilities, unmatched deployment possibilities, and superior management features. Portworx remains the gold standard in cloud-native Kubernetes storage for the enterprise.

**Challenges:** Data efficiency capabilities are limited when the solution is not coupled with enterprise shared storage.

## SUSE (Rancher Labs)

Longhorn is an open source, cloud-native storage solution originally developed by Rancher Labs and recently acquired by SUSE. It is managed by the Cloud Native Computing Foundation as a sandbox project (although the designation can be confusing, since many organizations use Longhorn in production).

Longhorn provides highly available persistent storage for Kubernetes through a two-layer architecture consisting of a data plane and a control plane, where Kubernetes itself handles the orchestration. The data plane consists of distributed block storage that aggregates and pools the local disk capacity available on each of the nodes. The control plane, via the Longhorn manager, creates volumes by spinning up Longhorn engine instances on the node the volume is attached to, and then creates replicas on the nodes where these should be placed. The outcome is a distributed and resilient storage platform with high-performance characteristics. Although Longhorn prioritizes resiliency, performance is adequate and may see further improvements due to roadmap development activities.

This solution handles backups and snapshots using a copy-on-write block storage layer that allows point-in-time recovery. Those backups can be exported either to S3 or NFS for offsite storage. The same technology can be used for disaster recovery and replication use cases with an active-passive cluster topology, making multi-site disaster recovery possible. A feature called "disaster recovery volumes" also enables cross-region replication in the cloud, with defined RPOs and reduced RTOs.

The solution offers no particular data optimization improvements, although backups are compressed and based on changed block tracking. Some techniques are used on secondary storage to either reclaim unused space or apply some degree of deduplication on backup blocks within a single volume. There are no plans to implement data efficiency capabilities for in-cluster storage, due to a focus on high performance and high availability. Organizations are thus expected to leverage application-level data efficiency mechanisms.

On the security side, role-based access control is supported through Kubernetes, and integration with Rancher technology enables the use of Active Directory and other enterprise-grade authentication providers. In-flight and at-rest encryption for data volumes is currently not available but is on Longhorn's roadmap. Monitoring and alerting are handled through the standard Prometheus/Grafana integrations.

Organizations can deploy Longhorn as a standalone solution, or benefit from the strong integration points Longhorn has with the SUSE platform (such as one-click deployment options).

**Strengths:** By itself, Longhorn is a solution with a focus on high performance and high availability. This storage solution also offers new advanced data services and has some good features such as the ability to export backups to S3 or NFS.

**Challenges:** No encryption capabilities are currently available, which can be a roadblock for regulated enterprises. Data efficiency features are absent in the primary data tier, which is where most of the data will be stored.

## Red Hat OpenShift Data Foundation

Part of a portfolio of solutions named Red Hat Data Services, Red Hat OpenShift Data Foundation (ODF) is a cloud-native storage solution based on Red Hat Ceph. The solution is scalable and resilient, and currently supports only Red Hat OpenShift, which is itself based on Kubernetes. For organizations considering consolidation on one technology stack, ODF provides frictionless operations at the storage layer.

ODF is versatile and supports block, file, and object storage. It can be deployed on-premises or in the cloud, and supports snapshots and clones. In the second half of 2021, Red Hat is expected to release new advanced data resilience solutions based on replication and disaster recovery mechanisms. For data protection, Red Hat's approach is to enable the ecosystem of third-party data protection vendors through its APIs.

ODF delivers strong performance without compromising on data optimization capabilities: Erasure coding, compression, and deduplication are currently supported. Multi-tenancy capabilities go beyond Kubernetes storage classes and include support for ResourceQuotas and LimitRanges, giving organizations control over resource usage and enabling them to overcome the hurdles of workload consolidation and the adverse impact from noisy neighbors.

The solution is excellent from a security perspective, with support for in-flight and at-rest data encryption (at the physical and volume levels). Key management is also supported by ODF. Monitoring and reporting capabilities are good, with integrations into the OpenShift console giving organizations all the basic performance and health metrics.

Finally, edge deployments are also supported when ODF is deployed in Compact Mode, starting at three nodes.

**Strengths:** A cloud-native storage solution with enterprise-grade features and an innovative approach to cloud deployments, ODF delivers solid value on multiple capabilities. Managed OpenShift services and storage options are available on multiple clouds now, enabling users to execute on their hybrid and multi-cloud strategy.

**Challenges:** Current support is limited to Red Hat OpenShift. Advanced data services remain a weak area. DR and advanced replication should be significantly improved in the next months. Even though RedHat OpenShift has good data-protection APIs, the approach to data protection relies on the solution ecosystem, which might deter small organizations looking for an integrated solution.

## Robin.io

Robin.io is an innovative, application-aware cloud-native Kubernetes solution with enterprise-grade capabilities. The solution can run anywhere, either on-premises (bare metal, virtual machines) or on all major public cloud providers.

The product, called Cloud Native Storage for Kubernetes (CNS), discovers and pools local disks of any type on Kubernetes cluster nodes, but can also pool storage capacity from cloud disks and SAN systems. Robin CNS delivers a highly available architecture with strictly consistent replication, auto resync on fall behind, and fast-failover capabilities. The solution enables bare-metal performance, live data rebalancing to avoid I/O bottlenecks, and the ability to enable QoS to throttle IOPs usage. QoS isn't limited only to storage, but also extends to CPU, memory, and network resources.

CNS shines with its advanced data services. Multiple replication modes are supported, with awareness at the node, rack, datacenter, and zone levels, providing organizations with sufficient granularity. To satisfy application-level deployment and performance requirements, advanced placement capabilities allow organizations to define fine-grained placement policies using affinity/anti-affinity rules. Robin's management interface includes an "application bundles" section that provides rapid deployment capabilities akin to an app store experience, while respecting best practices deployment topologies for those applications.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

19

The solution also supports snapshots and application-consistent, incremental forever backups. Replication capabilities can be used for data copy (application cloning), disaster recovery, and application mobility across clouds. Data compression is possible, and object storage is supported through integrations with MinIO.

Robin CNS supports encryption, although customers have to operate their own key management system. Monitoring and observability capabilities have been improved.

This solution is well suited to address edge computing use cases. Robin has a proven track record with the telco vertical, where edge deployments related to 5G infrastructure are one of the major use cases for containers.

Robin CNS is available in two tiers: a free tier, called Robin CNS Express (currently limited to five nodes and 5TB of data) and Robin CNS Enterprise, its mission-critical paid tier, which can be consumed either using a pay-as-you-go model or through an annual subscription.

**Strengths:** Robin.io delivers a comprehensive, feature-rich, enterprise-grade experience with an uncompromising adherence to cloud-native development and deployment principles. Advanced data services and application-awareness capabilities are among the highlights of this solution.

**Challenges:** Although the platform is very complete, further improvements in data efficiency (such as erasure coding and deduplication) as well as monitoring (predictive analytics, currently on the roadmap) would further strengthen Robin's position as a leader.

## StorageOS

StorageOS is a company focused on delivering cloud-native persistent block storage capabilities to Kubernetes through its product (also named StorageOS). The solution aims to address the storage needs of high-performance, mission-critical containerized applications.

The solution architecture consists of containers that are local to each Kubernetes cluster node. Each of these containers manages the local-attached storage present on the nodes they run on. The capacity across all these nodes is then aggregated and pooled through a StorageOS Pool layer, which is presented to all the pods in a given cluster.

The Kubernetes orchestrator can then communicate with StorageOS to provision or deprovision persistent volumes as needed for any of the containers executing on the nodes present in the cluster. The solution is highly available through cross-node volume replication and is built to deliver both scalability and performance, with particular attention to latency-sensitive workloads such as databases, message queuing, and high-frequency trading applications.

The solution currently supports synchronous replication and a feature called "delta sync" that replicates only the missing data in case of a cluster rebuild. For optimization and efficiency,

StorageOS uses data compression and an intelligent thin provisioning feature. Multiple storage classes are currently supported, but additional multi-tenancy features (such as QoS and affinity groups) are not yet available, though on the roadmap. Security is a notable area with role-based access control and namespaces, data-in-transit and data-at-rest encryption, as well as the ability to use unique encryption keys per volume.

Besides its own graphical user interface, the solution also integrates with Prometheus and Grafana, with a focus on IOPS, bandwidth, and free space. Finally, although StorageOS primarily targets business-critical applications, its architectural foundation is well suited to address edge use cases, thanks to its low overhead, and it should be a great fit once asynchronous replication capabilities are implemented.

**Strengths:** StorageOS is a solution with a lot of potential and currently a strong fit for performance-oriented cloud-native applications. The solution provides a robust and scalable architecture designed to meet demanding latency and throughput requirements, and is engineered to run on any platform, and with potential for edge use cases.

**Challenges:** Several key capabilities, such as snapshots, are not yet available on StorageOS, though they may be currently on the roadmap. A lack of advanced data efficiency mechanisms, such as deduplication and erasure coding, may become a challenge in the future.

# VMware

VMware Tanzu is built on top of vSAN and thus 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), a framework for modern stateful service providers to build Kubernetes plugins or operators on and for the 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,

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

21

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 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 quickly and effortlessly adopt Tanzu, enabling a fast movement toward 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.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

22

# 6. Analyst's Take

Cloud-native Kubernetes storage solutions are very dynamic. All of the vendors are innovating at a fast pace, but not all of them are focusing on the same priorities, and we can see in this Radar clear differentiation in three groups.

The first group consists of open source projects such as OpenEBS and Longhorn. These tend to focus on addressing the two primary challenges of such solutions: stability and performance. This translates into a strong effort to build resilient and scalable solutions with ongoing performance optimizations. For these solutions, advanced data services (data protection, replication, and so on), as well as data efficiency capabilities, are secondary.

The second group covers commercial platforms with a stronger emphasis on enterprise-grade features, although not all solutions are covering features equally. In fact, despite a couple of leaders with a complete platform and feature set, most of the solutions assessed in this report are focused on solving specific use cases. Thus, these gaps remain acceptable while cloud-native storage is still in its early stages and organizations are still exploring Kubernetes. Once enterprises become more comfortable with cloud-native workloads, such gaps will become disqualifying factors. It is therefore important that vendors seek ways to address these gaps while customers are still accommodating.

The third group consists of niche solutions, notably VMware, NetApp, and Arrikto. These are well engineered and usually offer excellent capabilities, but the deployment model, the underlying architecture, or the use-case focus prevent them from realizing broader adoption.

One concerning area seems to affect all of the vendors in this cloud-native Kubernetes storage report: data efficiency. Although cluster deployment still seems to be done in a fragmented manner, the growing use of Kubernetes as the foundation to build and run applications will inevitably push storage requirements to grow as well. This growth will lead organizations to consolidate Kubernetes storage, and consolidating Kubernetes storage at scale without enterprise-grade data efficiency mechanisms risks serious negative impacts in terms of storage costs and wasted space.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

23

# 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.

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

24

# 8 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.

# 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

GigaOm Radar for Cloud-Native Kubernetes Data Storage v2.0
This PDF prepared for: Adrian Escarcega (adrian.escarcega@gigaom.com)

27