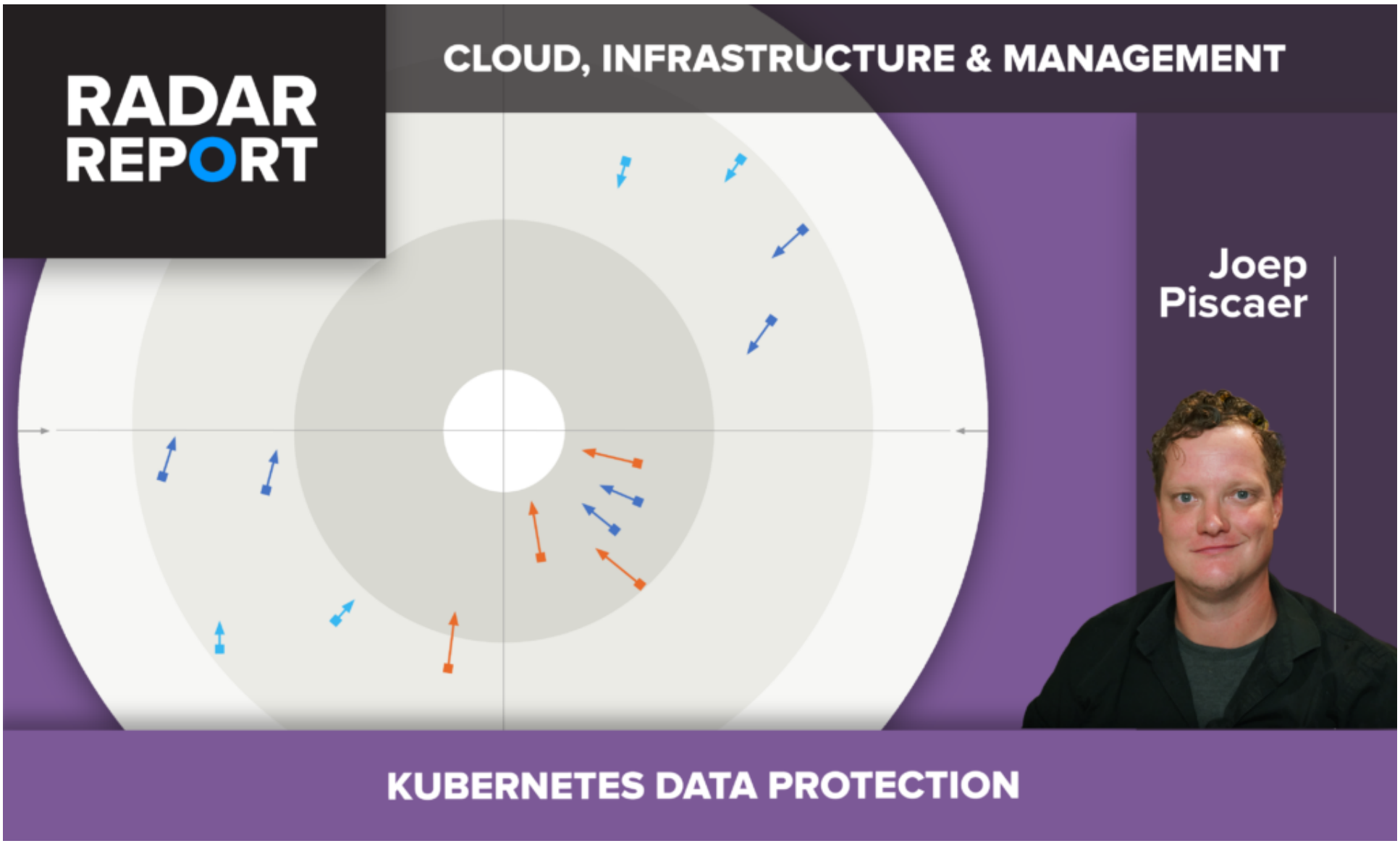

Joep Piscaer
Jan 26, 2023 -- Market Radar

GigaOm Radar for Kubernetes Data Protection v3.0

Table of Contents

1. Summary

Kubernetes is the industry standard for container orchestration, and it's being used by born-in-the-cloud startups and cloud-native enterprises alike. It's found in production on-premises, in the cloud, and at the edge for many different types of applications, including some that Kubernetes wasn't initially built for.

Kubernetes was never really meant for stateful applications, and by default, it lacks many data management and protection features. However, many organizations are building and running their stateful applications on top of Kubernetes, indicating there's a gap in functionality between what Kubernetes offers and what the (enterprise) market wants.

Unfortunately, existing data protection tools, mostly built for legacy technologies such as virtual machines (VMs), do not fit well into the container paradigm. However, vendors are adapting existing solutions or creating new products from scratch that are better aligned with the cloud-native and container worlds.

Many of these solutions include data protection and other data management features, such as data integrity and security, disaster recovery, and heterogeneous data migration capabilities. There's some overlap among data storage solutions, data protection solutions, and data management solutions in the cloud-native space, with each solution offering some adjacency in terms of features.

We have seen a particular focus on ransomware and other data integrity and security features in the last year, with vendors developing protective measures against different kinds of attacks, including ransomware, abuse of misconfigured cloud resources, and more. The companion Key Criteria report dives into the capabilities we expect to see in this space—namely, cloud-native data storage, protection, security, and migration.

The market for cloud-native data protection is growing rapidly, with both incumbent vendors and challengers in the market competing for completeness of features across the four key pillars. Differences can be observed between those targeting more traditional infrastructure alignment and those targeting fully cloud-native environments.

We continue to see the lines blurring between container-based workloads running on Kubernetes and other cloud services (like using cloud-native services for storage and databases), meaning data protection solutions increasingly must support these cloud services in addition to Kubernetes. The key differentiator observed in this year's report is the ability to back up application data as a whole across technologies, including VMs, containers, databases (running in a VM, container, or as a cloud service), and storage (across a variety of on-premises, container, and cloud technologies).

Last year's differentiators—mainly multiplatform, multicloud, multiteam, multiple environment (including edge), and self-service features—are now commonplace, with only a few exceptions.

This GigaOm Radar report highlights key Kubernetes data protection vendors and equips IT decision-makers with the information needed to select the best fit for their business and use case requirements. In the corresponding GigaOm report "Key Criteria for Evaluating Kubernetes Data Protection Solutions," we describe in more detail the key features and metrics that are used to evaluate vendors in this market.

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, Deployment Types, and Architectures

For a better understanding of the market and vendor positioning, we assess how well Kubernetes data protection solutions are positioned to serve specific market segments and deployment types (**Table 1**). We also categorize solutions by their architectures and core features (**Table 2**)

We recognized three market segments for this report:

- **Small-to-medium businesses (SMB):** 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. In this category, ease of use and quick consumption models, like software as a service (SaaS), are important factors, as are simple pricing models such as subscription-based pricing. Additionally, SMB and individual contributors don't need the full weight of enterprise features for auditing and compliance.
- **Large enterprise:** Here, offerings are assessed on their ability to support large and business-critical projects. Optimal solutions in this category have a strong focus on flexibility, performance, data services, and features to improve security and data protection, as well as extensive data mobility features for various migration scenarios. Scalability is another big differentiator, as is the ability to deploy the same service in different environments. Lastly, enterprise features for compliance and security, like role-based access control (RBAC) and multifactor authentication (MFA), support for a wide range of identity providers, and audit logging, are taken into account.
- **Edge and other specialized use cases:** A new deployment type becoming popular is deployment in edge and IoT-like scenarios, including telco and retail deployments. In these cases, policy-based fleet management of data protection across many clusters is a key differentiator, as are dark deployments.

We also recognize two deployment models for solutions in this report:

- **SaaS (managed and hosted):** Available only in the cloud and as a managed service, this approach is usually based on a pay-as-you-go subscription model. Users do not need to manage the infrastructure or backup repositories, just backup policies and day-to-day operations. This deployment model includes first-party SaaS (run and managed directly by the data protection vendor) and third-party SaaS (run and managed by a managed hosting partner). Often, the SaaS deployment model still requires an agent or component to run on each protected cluster.
- **Self-hosted (on-premises or in-cloud):** These solutions are meant to be installed both on-premises and in your cloud environment. While they are more complex to deploy and manage, they are more flexible in terms of where and how they are deployed. This deployment type is more suitable for those with stricter requirements for operational control or who have specific deployment requirements. The architecture of this deployment type can vary among hub-and-spoke, fully per-cluster, self-contained deployments, and a number of other architectures.

Table 1. Vendor Positioning

MARKET SEGMENT			DEPLOYMENT MODEL	
SMB	Large Enterprise	Edge & Other Specialized Use Cases	SaaS (Managed & Hosted)	Self-Hosted (On-Premises or In-Cloud)

CloudCasa	+++	++	++	+++	++
Commvault	+++	+++	+++	++	++
Dell	++	+++	+++	+	++
Druva	++	+	+++	+++	–
HYCU	++	++	++	++	++
IBM	++	++	++	–	+++
Kasten	+++	+++	+++	+	+++
NetApp	++	+++	+++	+++	+++
Portworx	+++	+++	+++	+++	+++
Rakuten	++	++	+++	+	+++
Trilio	++	+++	++	–	+++
Veritas	++	+++	++	+	++
VMware	+	++	+++	++	++
Zerto	++	++	+++	+	+++

Source: GigaOm 2023

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

In addition, data protection solutions for Kubernetes can also be categorized according to their architecture and core features (**Table 2**):

- **Traditional data protection solutions that support Kubernetes**: These solutions have a mature feature set for VM-based environments, and have added support for Kubernetes.
- **Cloud-native storage with data protection capabilities:** These solutions offer data protection capabilities on top of a cloud-native data storage product. You can't use the former without adopting the latter.
- **Cloud-native data protection:** Data protection solutions specifically designed to work with Kubernetes.

Table 2. Architecture Comparison

	ARCHITECTURE		
	Traditional	Cloud-Native Storage	Cloud-Native Data Protection
CloudCasa	–	–	+++
Commvault	+++	–	+
Dell	+++	–	–
Druva	–	–	+++
HYCU	+++	–	–
IBM	++	++	–
Kasten	–	–	+++
NetApp	–	+++	+++
Portworx	–	+++	+++
Rakuten	–	+++	+
Trilio	–	–	+++
Veritas	+++	–	–
VMware	–	–	+++
Zerto	–	+++	+++
VMware	–	–	+++

Source: GigaOm 2023

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

*Note: **Tables 1** and **2** are provided as an aid to better understand each vendor's market positioning, deployment model, and architecture, but are not taken into account in the final scoring and positioning in the Radar chart below.*

3. Key Criteria Comparison

Building on the findings from the GigaOm report "Key Criteria for Evaluating Kubernetes Data Protection Solutions," **Table 3** summarizes how each vendor included in this research performs in the areas we consider differentiating and critical in this sector. **Table 4** follows this summary with insight into each product's evaluation metrics—the top-line characteristics that define the impact each will have on the organization.

Table 3. Key Criteria Comparison

	KEY CRITERIA					
	Interoperability	Environmental Awareness	Disaster Recovery & Business Continuity	Application & Data Migration	Cloud Data Services Backup	Data Integrity & Security
CloudCasa	+++	+++	++	+++	++	+++
Commvault	+++	+++	+++	+++	+++	+++
Dell	++	++	++	++	++	+++
Druva	+	++	+	+	++	+
HYCU	+	++	++	+++	++	++
IBM	+	++	++	++	–	+
Kasten	+++	+++	++	+++	+++	+++
NetApp	++	++	+++	++	+	+
Portworx	+++	++	+++	+++	+	++
Rakuten	+	++	+++	++	–	+
Trilio	+++	+++	+++	+++	+++	+++
Veritas	++	+	++	+	+	++
VMware	++	++	+	++	–	+
Zerto	+	+	+++	+++	+	+

Source: GigaOm 2023

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

Some of last year's key criteria have become commoditized and moved to table stakes, including multicloud and multidistribution support, as well as multicluster and self-service capabilities. Like last year, native integration with the Kubernetes application programming interfaces (APIs) and—at a minimum—container storage interface (CSI) integration for source storage support are considered table stakes. Finally, we consider operational security (encryption of management portal and backup data in flight and at rest, MFA, RBAC, and multitenancy) to be crucial.

Key criteria this year are features that enable data protection of entire applications that run a gamut of technologies, including VMs, containers, cloud databases, cloud storage, and more. We look at a vendor's ability to support application data protection fragmented across these four named technologies. We take special note of a vendor's auto-discovery capabilities on-premises and in the big three clouds—Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP).

Flexibility is a key differentiator in the market, and both source and target storage integration and support play a key role. We look for a vendor's ability to support a wide range of source storage (including integrations and optimizations for native snapshotting and changed block tracking) as well as a wide range of target storage (including S3-compatible on-premises and cloud storage, other cloud object stores, and various on-premises file and block options). This storage support also enables differentiating capabilities for data migration and data copy management for dev/test, cluster/cloud/region migration, and compliance use cases.

Last year's uptick in the need for data mobility and migration capabilities continues. As teams and organizations adopt Kubernetes and container-based applications, the need for operational control, and thus the ability to migrate stateful data between platforms, is growing because companies no longer have the luxury of dealing with greenfield projects only. The reality is that many companies are dealing now with migrating a complex landscape of applications to Kubernetes and are in need of advanced application migration and transformation tooling. This evolution of demand includes features used to create duplicates of application environments for application development and testing.

Similarly, disaster recovery features are enabled in different ways by different vendors. We look for synchronous and asynchronous replication technologies, as well as workflows that enable testing, failover, and failback use cases.

Data integrity and security features are crucial for data protection solutions. Vendors are developing protective measures against different kinds of attacks, including ransomware and abuse of misconfigured cloud resources. Immutability for backup targets (using S3's Object Lock functionality or similar) are commonplace, but support for air-gapped and offline media is intermittent, as is pattern detection, analysis, and remediation.

Table 4. Evaluation Metrics Comparison

	EVALUATION METRICS				
	Flexibility	Scalability	Performance	Usability	Security
CloudCasa	++	+++	++	+++	+++
Commvault	+++	+++	++	++	+++
Dell	+	++	+++	++	++
Druva	+	++	++	++	++
HYCU	+	++	++	++	++
IBM	++	++	++	++	+
Kasten	+++	+++	+++	+++	+++
NetApp	+	++	+++	+++	+
Portworx	+++	+++	+++	+++	++
Rakuten	+	++	+++	++	++
Trilio	+++	+++	+++	++	+++
Veritas	+	+++	++	++	+++
VMware	+	++	++	+	++
Zerto	+	++	+++	++	++

Source: GigaOm 2023

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

Cloud-native architectures are nothing if not continuously evolving. That means data protection solutions for Kubernetes must follow suit, especially with stateful data being fragmented across so many services and technologies. Flexibility is a key evaluation metric with which we gauge a solution's ability to flow with the tide. Not only are the table stakes of multiplatform, multicloud, multiteam, and multiple environment (including edge) key enablers of this flexibility, but the ability to back up data wherever it resides and wide support for storage sources and targets play a crucial role as well.

The most flexible solutions are those able to go beyond data protection and offer additional services that can simplify operations and speed up data management tasks for administrators and developers. These added services include migration workflows, data copy management, and disaster recovery workflows, which allow the data protection solution to fit a wide range of use cases, including new ones that involve new functionality added to the solution via new version releases.

We also evaluate solutions on scalability (vertical and horizontal), architecture (static or dynamic), resource usage and performance (autoscaling, native storage integrations, compression, and deduplication), security (both operational and for data integrity), and usability—the maturity and completeness of graphical user interfaces (GUIs) and dashboards, command-line interfaces (CLIs), software development kits (SDKs), and APIs.

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.

The GigaOm Radar plots vendor solutions across a series of concentric rings, with those set closer to the center judged to be of higher overall value. The chart characterizes each vendor on two axes—balancing 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.

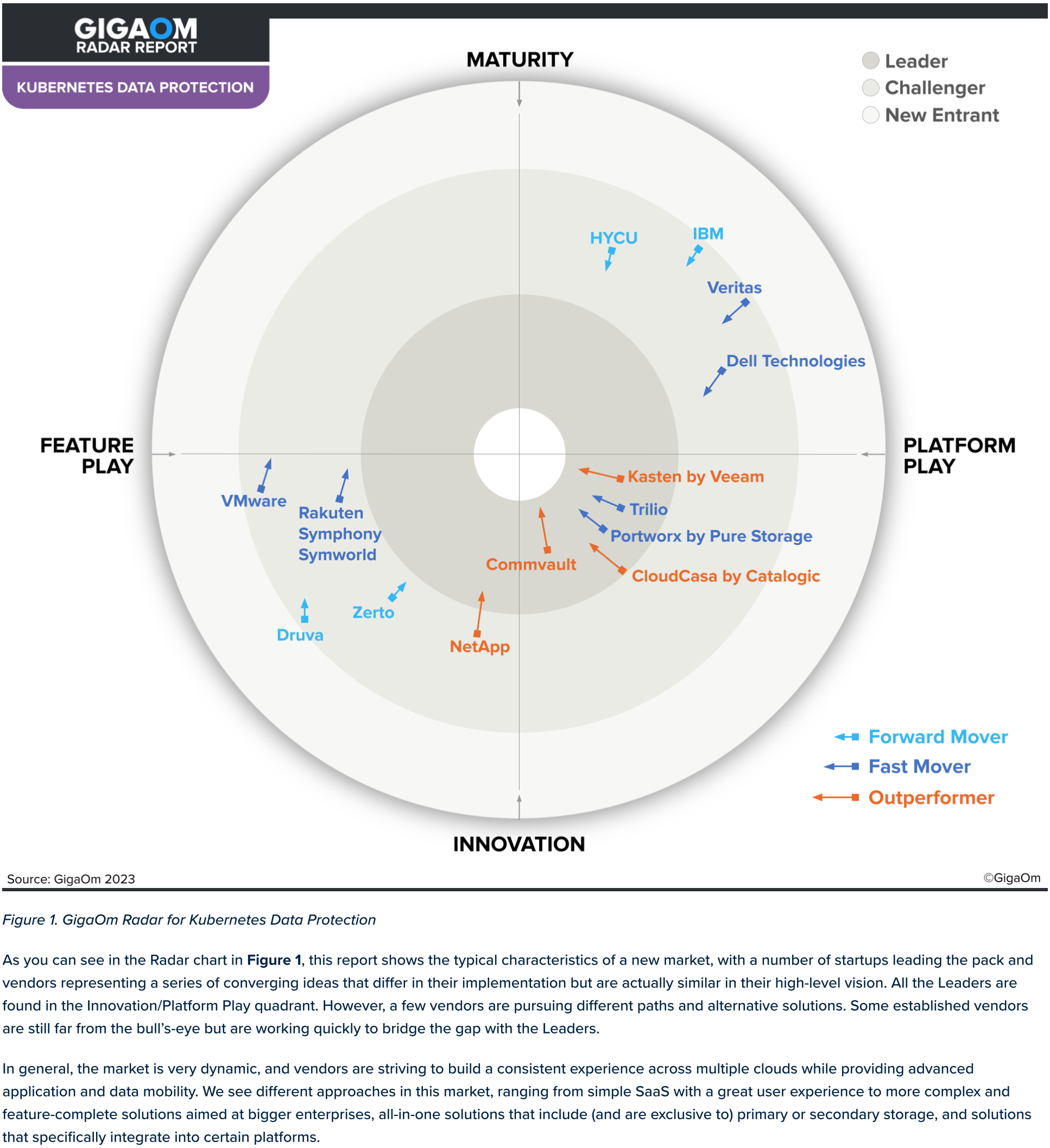

Figure 1. GigaOm Radar for Kubernetes Data Protection

As you can see in the Radar chart in **Figure 1**, this report shows the typical characteristics of a new market, with a number of startups leading the pack and vendors representing a series of converging ideas that differ in their implementation but are actually similar in their high-level vision. All the Leaders are found in the Innovation/Platform Play quadrant. However, a few vendors are pursuing different paths and alternative solutions. Some established vendors are still far from the bull's-eye but are working quickly to bridge the gap with the Leaders.

In general, the market is very dynamic, and vendors are striving to build a consistent experience across multiple clouds while providing advanced application and data mobility. We see different approaches in this market, ranging from simple SaaS with a great user experience to more complex and feature-complete solutions aimed at bigger enterprises, all-in-one solutions that include (and are exclusive to) primary or secondary storage, and solutions that specifically integrate into certain platforms.

In this context, a number of solutions are doing very well, including Kasten, Portworx, and Trilio. Some more established, traditional vendors like Commvault are also doing very well by combining solutions for SaaS applications, on-premises infrastructure (VM-based), containers, and databases efficiently. Some solutions, such as CloudCasa, tick all the right boxes for those buyers seeking a SaaS solution with a great user experience.

Additionally, we see some vendors combining data protection solutions with other functionality, like data storage. NetApp and Rakuten Symphony Symworld Cloud Native Storage (CNS) receive a virtual asterisk because their data protection features are great but exclusive to their respective storage platforms.

All in all, the data protection field is moving quickly and is accelerated further by Velero's growing importance in the Cloud-Native Computing Foundation (CNCF) ecosystem as the underpinning technology for many of the solutions discussed in this Radar.

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

CloudCasa by Catalogic

CloudCasa is a SaaS service that enables you to backup, restore, migrate, and secure Kubernetes-based applications.

Its pricing model is capacity-based. The free tier allows customers to back up Kubernetes resource metadata to CloudCasa and create unlimited local persistent volume (PV) snapshots with a retention period of 30 days as well as store 100 GB of backups in CloudCasa's backup target repositories. Paid tiers allow customers to use CloudCasa-provided backup targets from 100 GB and bring their own repositories (on-premises or cloud).

CloudCasa strongly focuses on the Kubernetes ecosystem, offering support for many Kubernetes-based platforms—including Azure AKS, AWS EKS, GKE, and DigitalOcean—but also has support for OpenShift, Rancher, Tanzu, and other platforms. In addition, CloudCasa supports CSI-compatible storage, and backing up cloud storage services in AWS, Azure (Disk and Files), and GCP, as well as network file system (NFS) PVs support. It also has native and deep integrations with the three cloud providers for backing up cloud databases (RDS and Aurora currently) and has support for cloud cluster and storage autodiscovery. However, it doesn't have support for protecting VMs running in KubeVirt.

CloudCasa supports bring-your-own backup repositories (which works with any S3-compatible target, including on-premises appliances) but also has a flexible selection of Azure and AWS object storage locations, including in different regions of the world.

CloudCasa has mature security features, including those for container, networking, configuration, and best practices benchmark scanning. These features will cover Kubernetes and can also scan cloud services, detecting misconfigurations and security vulnerabilities across identity and access management, key management, object stores, and more. Its SaaS interface meets AWS's strict security, efficiency, and reliability demands, which include MFA, suspicious IP throttling, fraud and brute force attack detection, and SOC2/ISO27001 compliance. It has support for data-at-rest and in-flight encryption, as well as ransomware protection features including immutability. It also supports private connectivity connecting on-premises backup targets or source clusters and CloudCasa's service. However, it doesn't support customer-provided encryption keys.

Application migration uses backup and restore workflows to migrate applications or data services between platforms. It supports heterogeneous restores between different storage classes and clusters, different regions, and different clouds or cloud accounts, enabling migration-to-and-from-the-cloud scenarios. When data is stored in a cloud backup target, it supports various disaster recovery scenarios, including on-site to disaster recovery site, on-site to cloud, and cloud-to-cloud. It has the ability to create the necessary clusters on the recovery side on the fly, simplifying recovery operations. CloudCasa does not include an optional storage layer and does not support synchronous replication, making the solution less suitable for low recovery point objective (RPO) and recovery time objective (RTO) disaster recovery scenarios.

Strengths: CloudCasa is a very complete data protection solution. As a SaaS solution, it's easy to deploy and manage. It has flexible backup repository options ranging from CloudCasa-provided to bring-your-own (on-premises). Its deep integration with AWS, Azure, and GCP make data protection of cloud resources (including databases) a breeze.

Challenges: While its security features are mature, there still is no support for customer-provided encryption keys. The lack of KubeVirt support is notable. While its migration capabilities are on-par with the competition, its disaster recovery capabilities are lacking replication.

Commvault

Commvault Backup & Recovery is a backup solution that supports more than Kubernetes workloads, making it suitable for hybrid applications that run across Kubernetes, VMs, and cloud services, consolidating backup operations on a single platform.

Commvault's backup solutions include backup and recovery software and the Metallic SaaS service, and have very broad support for data sources, including databases like Amazon RDS, file and object stores like S3, Azure Blob and Files, Microsoft 365, and compute instances, including VMs, Amazon EC2 instances, and Kubernetes-based containers. Additionally, there is support for backing up CI/CD pipelines, including code repositories in Azure DevOps and GitHub. Its flexible deployment architecture and single interface across multiple deployments means the solution is multicluster and multicloud by default.

The solution is compatible with all CNCF-certified distributions (notably Tanzu and OpenShift) and cloud services (GKE, EKS, AKS, and OKE). It is integrated with CSI for snapshot-based backups. It can provide application consistent backups, thanks to the ability to use pre- and post-backup execution scripts to quiesce data on storage before taking the snapshots, minimizing risk of data loss. Many scripts for MySQL, Cassandra, MongoDB, and PostgreSQL are included. Furthermore, data and applications can be backed up and restored in an alternate Kubernetes environment, simplifying migrations and disaster recovery operations.

The Commvault solution is available across various cloud marketplaces. It works with a specialized access node (deployed as a VM) that interacts with the Kubernetes API server to discover and protect applications. Data movers are deployed on a cluster only during a backup or restore, with no other parts of the backup application running on-cluster.

The interface includes over 200 built-in reports and a custom reports builder, and shows both infrastructure requiring attention and at-a-glance recovery readiness status across clusters.

There are self-service capabilities for developers through the Command Center UI, thanks to integration with the Kubernetes RBAC system. Applications are discovered automatically using label selectors or entire namespaces, covering new applications with a blanket backup policy.

Security and ransomware controls hint at Commvault's long history in data protection. Its interface security is well taken care of with MFA support, data encryption (at rest and in flight), bring-your-own-keys, and support for various key management systems, air-gapped backups (optionally via Commvault's Metallic Recovery Reserve Storage Service), backup target immutability, and deep support for anomaly detection for both operations in the UI and deletions and changes at the storage target level for malware detection. Backup storage targets include HyperScale appliances, Metallic Recovery Reserve Storage Service, and object stores (including S3-compatible on-premises, NAS/SAN, and tape), while functionality encompasses support for deduplication and compression.

Additionally, Commvault's use of machine learning (ML) algorithms to optimize operational tasks is notable. For instance, these algorithms tune backup operations to keep backups within their set service-level agreement (SLA) automatically.

Disaster recovery has various options, including integration into Commvault's storage system for (a)synchronous replication and snapshot-based disaster recovery across cloud regions or on-premises data centers. The UI includes mature workflows for heterogeneous restores (like cross-clusters or cross-cloud), remapping of resources, and dev/test pre-seeding.

For edge use cases, Commvault has a physical appliance-based solution, HyperScale X, that supports Azure Stack, AWS Outpost, EKS Distro, and Google Anthos Bare Metal.

Commvault provides several purchasing options ranging from traditional perpetual licensing to subscriptions. Licenses are transferable between VMs and containers.

Commvault Metallic is aimed at organizations looking to move away from self-managed data management platforms and includes support for Kubernetes. It is integrated with HyperScale X, offering quicker on-premises restores using the physical appliance in disaster recovery scenarios and dark site deployments.

Strengths: Its broad support for VMs, containers, and (cloud) data services and databases in a single platform make it a great choice for hybrid and complex applications. Its security and ransomware controls are very extensive, making it suitable for larger enterprises.

Challenges: With flexibility comes complexity. Commvault's solutions require time and attention to implement correctly, and they are not as easy to implement or maintain as other solutions.

Dell Technologies

Dell PowerProtect is a container protection solution built on Velero, with appliance-based deduplication and compression using PowerProtect DD series appliances, which can be physical or virtual (running on top of cloud object storage).

The VM-based manager deploys the data movers automatically on protected Kubernetes clusters. It also deploys as a VM in public cloud environments to protect cloud-based Kubernetes environments. There is no SaaS version available currently, but future releases will include an operator-based deployment and will be available through various marketplaces. The Dell solution can connect to multiple on-premises or cloud-based Kubernetes clusters from a single manager.

PowerProtect uses Velero components to capture the Kubernetes-level objects while it leverages its own data movers (running in their own namespace). Backups run at the namespace level. Restores include remapping of storage classes.

It supports various cloud-based Kubernetes services, including AKS, EKS, and GKE. In the on-premises world, it supports VMware Tanzu Guest Clusters, Tanzu Kubernetes Grid Integrated Edition, Rancher, OpenShift, and Diamanti.

PowerProtect includes at-rest and in-flight encryption, as well as target immutability with retention locks. Optionally, it can be integrated with Dell's Cyber Recovery Solution for air gapping and advanced ML and anomaly detection to combat ransomware.

PowerProtect Data Manager comes with built-in support for MySQL, PostgreSQL, and non-sharded MongoDB and Cassandra. It also recently added first-generation support for Oracle. Currently, PowerProtect only supports backing up cloud-based databases like Amazon RDS via an additional product, Cloud Snapshot Manager. This functionality will be integrated in the future.

PowerProtect Data Manager includes workflows to support the migration of applications between clusters, including the underlying storage across disparate storage classes.

Strengths: PowerProtect protects more than just containers, making it a good solution for protecting Kubernetes-based applications when you're already using PowerProtect. It offers efficient replication of data between storage appliances across on-premises environments or cloud providers for disaster recovery purposes.

Challenges: PowerProtect requires backups to be stored on PowerProtect DD series appliances. While offering cost savings with more efficient compression and deduplication, this requirement limits the practical applicability of the solution to those already in the Dell ecosystem.

Druva

Druva has a solution designed to protect Kubernetes workloads in Amazon AWS, in both EKS (managed) and EC2-based (self-managed) clusters. The solution is extremely easy to set up and use; it is application-aware and is able to protect data stored in RDS databases as well. Druva gives AWS users a solution capable of protecting complex applications that take advantage of both containers and the AWS ecosystem. At the same time, Druva can protect common databases such as MySQL, PostgreSQL, and MongoDB with the mechanisms necessary to ensure end-to-end application consistency.

The solution is simple, well-designed, and scalable, with the core backup controller module installed as an operator inside the cluster. Each individual backup job is instantiated separately for improved scalability and parallelism.

Druva allows backing up and recovering data in different AWS regions for disaster recovery and application mobility, and integration with Kubernetes RBAC lets application owners use the CLI to perform day-to-day operations in a self-service fashion.

Additional cloud services and on-premises Kubernetes distributions will be added later, allowing companies to protect and move applications and data across multiple environments.

Strengths: This is an easy-to-use and well-integrated solution for the AWS ecosystem with the potential to evolve into a credible multicloud solution. Data protection for DRS can simplify application protection dramatically in some circumstances.

Challenges: The lack of multicloud support is a showstopper for most Kubernetes scenarios with applications deployed on-premises and in multiple clouds.

HYCU

HYCU's solution for Kubernetes centers around the SaaS-based Protégé. Storage integration is not CSI-based (nor are there any plans to support CSI in the short term), but HYCU leverages native cloud APIs (Google, Azure, and AWS) and existing storage providers for snapshot functionality (Nutanix, VMware). No other Kubernetes distributions or services are supported.

Kubernetes applications are auto-discovered via the YAML metadata. The policy-based backups are assigned using Kubernetes labels. Policies include options to do snapshots, full backups, long-term retention, and storage tiering from a single policy. Additionally, a copy can be stored outside the local region for disaster recovery purposes. Restores can be made to heterogeneous environments, including different clusters and different regions, with cloning functionality for test/dev. Workflows include mature remapping options for restores to disparate cluster configurations.

The web-based interface is simple and easy to use, but logins are cloud-provider specific; it reuses the identity provider active in the cloud where the solution is deployed. Each instance of HYCU is limited to use in a single cloud (depending on where it's deployed) and is not natively multicloud. A "manager of managers" interface is available upon request for multicloud capabilities. The UI includes self-service capabilities based on the Kubernetes primitives.

HYCU has some support for backing up cloud-native database and storage services, like Amazon RDS. Backup targets include S3-compatible storage services; the on-premises product also supports existing server message block and NFS targets. Immutability features are supported only on S3. Data encryption in flight and at rest is enabled and includes bring-your-own-keys as an option.

HYCU has basic support for data management and application migration and transformation but requires reuse of existing backup and restore workflows.

The pricing model is based on allocated source capacity, but options for VM-based and socket-based pricing also exist.

Strengths: Its integration into Nutanix and VMware-based environments make HYCU a strong option for those customers already running on these platforms. The backup policies and restore workflows are very mature, supporting migrations, transformations, and disaster recovery alike.

Challenges: The product is immature in supporting other platforms, and its lack of CSI support will remain a challenge for the foreseeable future. Multicluster support is minimal. The limited support for RDS will prove a disqualifier for some buyers.

IBM

IBM Spectrum Protect Plus (SPP) added the ability to protect Kubernetes workloads. This solution is designed to protect all modern environments, including virtualized infrastructures, and can be used by organizations of all sizes, including service providers.

SPP provides a native Kubernetes CLI and leverages Kubernetes orchestration capabilities to allocate resources necessary for backup jobs. The product's current focus is on Red Hat OpenShift. SPP takes advantage of Red Hat OpenShift APIs for data protection (OADP) based on Velero open-source technology, and IBM directly interfaces with Velero for other Kubernetes distributions, keeping the user experience consistent across distributions.

The SPP component for Kubernetes is deployed as a custom resource definition (CRD) operator. Administrators and developers can search the Kubernetes inventory easily and manually select the applications to protect. Comprehensive and automated discovery mechanisms will be added in future releases. Application consistency is not yet available, but IBM is working to implement mechanisms to ensure off-the-shelf application consistency for its IBM Cloud Paks, common databases, and custom applications.

IBM SPP boasts several licensing options, including perpetual licenses and subscriptions per terabyte of persistent storage under protection. The product is already available in the various cloud marketplaces. Ransomware protection is subpar compared to the competition, as is support for cloud databases and data storage services.

Strengths: Its strong focus on Red Hat OpenShift and IBM ecosystem (IBM CloudPaks) while providing protection for both traditional and modern workloads are positive attributes.

Challenges: IBM has a good roadmap for future product releases, but right now it presents several limitations for complex Kubernetes environments.

Kasten by Veeam

Kasten K10 is a cloud-native data management platform for Kubernetes. It includes backup, restore, disaster recovery, and migration functionality for container-based applications.

It's a self-hosted, auto-scaling solution, designed to run on each cluster it protects. It's installable via various cloud and Kubernetes marketplaces. Despite the lack of a first-party SaaS version, Kasten is offered as a service through local Veeam VCSP providers.

Kasten focuses on enterprise customers, and its key users are both cloud and IT administrators as well as developers for self-service backups and restores. It is extremely easy to use, with a high-quality GUI, consistent APIs, and a handy CLI. Kasten K10 is designed to cope with large-scale environments, but it can be deployed in small infrastructures as well.

K10 supports a wide range of Kubernetes flavors, including OpenShift, Rancher, Tanzu, EKS, AKS, and GKE, across on-premises and cloud. It has partnerships with various edge and retail deployment partners like K3s and AWS EKS Anywhere.

It has support for protecting VMs running in Red Hat OpenShift Virtualization or KubeVirt. Combined with the new integrations into Veeam Backup & Recovery (that bring Kasten into VBR's interface), it's well suited to protect complex applications spanning container, cloud-native, and VM-based deployments at large enterprise deployments.

K10 natively discovers data services (like MySQL, MongoDB, PostgreSQL, Amazon RDS, Kafka, and Cassandra) as part of applications, and automatically applies the right data management policies (for things like quiescing) using Kanister, an open-source data management framework initially developed by Kasten to create an industry standard for stateful data management. These blueprints are continuously updated and expanded, adding support for more database technologies and improving existing integrations. However, K10 does not support backups of AWS resources like virtual private cloud (VPC) or identity and access management (IAM) configuration.

K10 natively supports CSI, including snapshots and backups for traditional enterprise storage arrays and Amazon EBS and EFS, Azure Managed Disk, Google Persistent Disk, VMware, and container-attached storage solutions, but has deeper CSI integration for a number of CSI providers offering additional functionality and backup performance. It includes features like changed-block tracking, fast incrementals, and transfers between repository regions. When it can use a more optimal underlying storage integration, such as OpenStack, CEPH, vSphere, or a public cloud's API, it will offload tasks directly instead of using CSI. Space and network efficiency are assured, thanks to deduplication and compression techniques.

Supported backup repositories include object stores, NFS shares, and existing Veeam backup and replication backup repositories. For any repository, K10 supports data-at-rest and in-flight encryption, as well as immutability for S3, MinIO, Cloudian, and others, preventing backups from being deleted or retention policies being relaxed in any cryptolocker attempts.

Backup data is always encrypted, both at rest and in flight. Encryption keys can be stored in the cluster or by using an external key management system like HashiCorp Vault or AWS KMS. It is also integrated with Kyverno and Open Policy Agent to ensure applications meet organizational data protection guidelines.

The Kasten interface supports managing multiple clusters from a single interface and supports RBAC and role and scope limitations for self-service access for non-administrative users to specific resources, like a single cluster or namespace. It uses Kubernetes Roles and ClusterRoles, replicating the already defined roles in the cluster. Kasten enables admins to standardize (global) policies to tenant clusters, allowing policies to be centrally managed but distributed across clusters. In addition, local cluster admins can define local policies.

The Application Transform Engine supports application data and metadata transformations (including KubeVirt VMs), migrations, and mapping, allowing use cases ranging from simple storage class mappings to cross-cluster, cross-region, and cross-availability zone (AZ), cross-distribution, and cross-cloud migrations. It also includes disaster recovery functionality to protect against cluster and availability zone failures, as well as storage system failures in on-premises scenarios.

Strengths: Kasten is a mature solution. Its RBAC features and centrally managed policy model are well aligned with large enterprise and self-service requirements. Its application-aware data management framework, Kanister, has Kasten- and community-provided blueprints for application-consistent backups.

Challenges: Its lack of first-party SaaS may be off-putting to some customers. As a very mature and complete solution, its pace of innovation is starting to suffer from first-mover disadvantages.

NetApp

Astra Control is NetApp's data protection, disaster recovery, and mobility product and is available as either a self-hosted/managed or SaaS solution. It is multicluster aware. It supports self-managed platforms like OpenShift Container Platform, Rancher Kubernetes Engine, and upstream Kubernetes, in addition to managed services in the public cloud (Google Kubernetes Service, Azure Kubernetes Service, and Amazon EKS).

However, Astra Control supports only a limited number of source storage providers, including those from NetApp itself (ONTAP, Azure NetApp Files, Cloud Volumes Service, Amazon FSx for NetApp ONTAP, and Cloud Volumes ONTAP), as well as cloud-native block storage services (Azure Disk, Google Persistent Disk, and Amazon Elastic Block Storage). The advantage of this tight coupling with NetApp's storage is the ability to offer asynchronous replication-based disaster recovery, snapshots, and more, allowing lower RTO and RPO. The downside is the requirement to have supported NetApp storage on both the source and the target clusters. Synchronous replication support is on the roadmap.

Astra Control supports object storage on Google, AWS, and Azure, and S3-compatible storage as a backup target, as well as using SnapMirror replication. It supports data-at-rest encryption per datastore and data-in-flight encryption. There is basic support for immutability on target object storage. Astra Control does not support cloud database services as backup sources without additional execution hooks, but it does support self-hosted databases (Cassandra, Elasticsearch, MariaDB, MySQL, MongoDB, PostgreSQL, Redis, and Kafka) through Verda, NetApp's open initiative for app-consistent snapshots.

Strengths: Astra Control is a basic but functional data protection solution for NetApp environments that leverages NetApp's trusted technologies for deduplication and inline compression. NetApp also supports the major cloud providers' block storage services. Opening up support to non-NetApp (on-premises) source storage is on the roadmap.

Challenges: Astra Control's capabilities over CSI are limited to the NetApp storage ecosystem. It does not natively support any cloud databases (requiring additional scripting), and its data integrity features are still developing.

Portworx by Pure Storage

Portworx Backup is a data protection solution for Kubernetes. It's compatible with any Kubernetes cluster on-premises and in the cloud, including OpenShift, Tanzu, EKS, AKS, and GKE, and works with any Kubernetes-compatible storage provider, including Portworx Enterprise, Pure's distributed storage solution for Kubernetes. Notably, it also supports backing up and restoring VMs running in KubeVirt, as well as any database that runs as a Kubernetes operator.

Recently, a SaaS version of Backup was introduced, making the solution fully managed for easier onboarding and lifecycle management. Additionally, Backup is available as a self-hosted Helm-based application in a Kubernetes cluster. It's also available in the AWS and IBM Cloud marketplaces.

It's a multicluster solution, able to protect any Kubernetes cluster across on-premises and cloud. Additionally, PX Central, a manager-of-managers, is able to manage multiple on-premises Backup instances centrally.

For each cluster it protects, it deploys Stork, which bridges the gap between the backup server and the administration cluster. Backups are policy-based, and applications are assigned to a policy via namespace and label selectors. A future release will also support SLA-based policies.

Additionally, Backup supports any CSI-compatible storage to perform and speed up application-consistent backups, Amazon EBS volumes, Google Persistent Disks, and Azure Managed Disks. Backup supports S3 (-compatible), Azure Blob, and Google Cloud Storage as backup targets. In recent versions, it added support to copy cloud provider snapshots (like EBS snapshots) into an S3 bucket, including storing that snapshot in a different region. Security is a critical aspect of the product, with data encrypted at rest and in transit. Portworx supports immutable backup targets for ransomware protection, and the roadmap shows additional data integrity functionality being worked on.

Backup supports pre and post hooks, and it has built-in support for Cassandra, ElasticSearch, Jenkins, MongoDB, MySQL, PostgreSQL, and RabbitMQ. It also supports backing up file shares (persistent volumes provisioned as file shares from a Pure Storage FlashBlade, Portworx proxy volumes, Amazon EFS or NFS exports).

It's multitenant aware, with mature RBAC functionality. It integrates with the Kubernetes RBAC controls and thereby adheres to the scope limitations and permissions set in Kubernetes so that users can interact only with their own namespaces or applications. It also has specific backup as a service (BaaS) roles of managing and delegating control for the Backup SaaS-tenant. It has recently added the ability for backup users to share backups (either specific point-in-time backups or all backups belonging to a namespace) with other users in the same tenant or instance.

Portworx has mature data migration and transformation capabilities based on Stork, enabling application and data migrations between clusters, regardless of whether these are on-premises or in the cloud. These features do, however, require Portworx Enterprise to be installed on both the source and target clusters in addition to Backup. In situations without Portworx Enterprise, it does support backup-and-restore workflows to migrate data to a different heterogeneous cluster. It has dedicated workflows for cross-cloud, cross-region, and cross-cluster migrations.

Since the previous report and the introduction of the fully managed BaaS, it has introduced a fully functional "free forever" tier, consisting of 1 TB of application data. Customers transition automatically to a paid tier after that initial terabyte, which is based on application data protected per month. For the self-hosted version, various licensing options exist, including per cluster-node-hour or perpetual licensing.

PX-DR is an add-on solution to Portworx Enterprise, specifically designed for disaster recovery, leaning on synchronous, near-sync, and asynchronous data replication. However, Backup does not include these features and requires the use of Portworx Enterprise, its data storage product.

Strengths: This is a top performer in the Kubernetes data protection market, with flexibility in deployment (self-hosted or fully-managed SaaS). A free tier to get started quickly is another strong point, as are mature data migration features for those looking beyond data protection to data storage, data management, and advanced disaster recovery.

Challenges: For more advanced disaster recovery features, like zero RPO disaster recovery and near-sync disaster recovery across cloud regions and availability zones, Portworx Backup requires the storage capabilities of Portworx Enterprise. Support for cloud data services like Amazon RDS and additional data integrity features are planned for future releases.

Rakuten

Rakuten Symphony Symworld CNS is a storage solution aimed at persistent, stateful container applications. Included in the product are various data protection features, including synchronous and asynchronous storage replication as well as snapshot and backup functionality.

CNS can be installed on major cloud providers and distributions, including OpenShift, Rancher, EKS, AKS, and GKE, and it is available via various application marketplaces, including Google and OpenShift. CNS has friendly, consumption-based, per-node-hour pricing, with discounts for annual subscriptions, similar to those for public cloud offerings.

Some features, including the (a)synchronous replication, require the storage solution to be installed on both the source and the target clusters, reducing the relevance of the data protection features for those not looking to replace their storage solution with CNS. Similarly, CNS does not protect any data not stored on its storage volumes, including RDS databases, which limits the applicability of its cloud data protection features.

A key benefit of the storage layer, however, is CNS's full control over, and visibility of, the Kubernetes objects and underlying storage. The data protection features incorporate Kubernetes metadata and configuration natively to protect and replicate all Kubernetes objects and constructs. The storage layer is CSI-compliant, making it easy to quiesce storage volumes during backups and snapshots, and CNS includes scripts for popular data services like MongoDB, Cassandra, MSSQL, MySQL, Oracle, DB2, and PostgreSQL.

CNS supports snapshots (local and remote), backups to a separate repository (local or remote), synchronous replication between storage volumes (using a stretched-cluster setup across two availability zones), and multisite asynchronous replication (with replication across multicloud and availability zones supported).

One area where CNS shines is migration scenarios between disparate environments, including migrations from on-premises to cloud, or cloud-to-cloud. The storage layer's replication features are a clear benefit for migrations, making it easy to migrate an entire application, including its storage, between on-premises and public cloud. The same is true for cloning applications for dev/test scenarios. These scenarios do require CNS to be installed on the target cluster. Similarly, backups can be restored only to CNS storage volumes.

The GUI includes a monitoring dashboard and the tools necessary for fast troubleshooting, offering mature Kubernetes-based RBAC for self-service and multitenancy with multicluster features.

All backups are encrypted in flight and at rest at the application level, with support for external key management. Backup targets include S3-compatible object storage platforms as well as public cloud storage services such as Google Cloud Storage and Microsoft Azure Blob, and on-premises storage solutions via NFS.

CNS does not include many ransomware and security features. While immutable backups on WORM-compliant object storage are supported, the product does not include security scanning features.

Combined with other CNS products, like Symworld Orchestrator and Symworld Cloud Native Platform, CNS is an interesting all-in-one choice for storage and data protection for edge use cases, like air-gapped (or dark) clusters for 5G-in-a-box use cases.

Strengths: CNS is an end-to-end solution with a friendly pricing model and is well-integrated with Kubernetes. Multicluster and self-service capabilities fit well into large enterprises that consider the data protection features integrated into a storage solution acceptable. The (a)synchronous replication features add low RTO and RPO capabilities for those applications that need it. The storage layer adds flexibility to low-downtime migration scenarios. The easy-to-use GUI and CLI are very helpful for users who have limited experience with Kubernetes.

Challenges: Data protection is not a standalone product but part of the CNS storage solution, limiting its use of data protection features to applications running on CNS storage. Some of the features require CNS to run on the target cluster, thereby increasing cost and complexity, even though the pricing is competitive. Restores are limited to CNS-based volumes, and there are limited security features.

Trilio

TrilioVault for Kubernetes is a data protection and resilience platform for Kubernetes. Deployed as a Kubernetes operator or Helm chart (and available in various marketplaces), the solution allows you to back up data and applications on every supported platform and restore them locally or elsewhere for development, migration, or disaster recovery. Trilio can protect applications discovered via labels, namespaces, or Helm charts, or as operators seamlessly. It boasts native support for Helm charts, which includes backing up its history and context, solving specific restore issues with Helm charts being overwritten by the default chart at restore time. Deployments can be air-gapped for dark or edge deployments.

TrilioVault is very well integrated into Red Hat OpenShift environments with support for Advanced Cluster Management (ACM) to apply and manage data protection policies across fleets. It has support for OpenShift managed services with AWS (ROSA) and Azure (ARO) and OpenShift Virtualization (KubeVirt). It supports all certified Kubernetes distributions and cloud managed services as well, including Google GKE, Amazon EKS, Azure AKS, and Digital Ocean. Additionally, it supports SUSE Rancher, VMware Tanzu, and more. Further, the solution is able to backup MongoDB, PostgreSQL, InfluxDB, MySQL, Redis, etcd, Cassandra, and AWS RDS-based databases.

Multicluster support is achieved by linking together multiple per-cluster deployments in the UI. TrilioVault is built as an auto-scaling application, temporarily allocating the necessary resources for every backup job, ensuring scalability and performance for environments of all sizes. It spins up additional data mover pods as running backup jobs are started.

Backup targets can be S3-compatible object stores, Azure Blob, or NFS shared volumes, and the solution natively supports data compaction techniques for network traffic optimization. Backups are stored in the open QCOW2 format and use the open-source LUKS for data-at-rest and data-in-flight encryption on a per-backup file basis, not on a per-repository basis. It also applies immutability on a per-file basis but only on S3-compatible storage.

The management console includes workflows for disaster recovery plans and workflows to migrate applications to disparate clusters, as well as restore hooks for custom scripting during restores, transformations (like storage class mappings), and exclusions to granularly restore data. Continuous Restore, while still snapshot-based, offers low RTO and RPO by using built-in replication functionality, enabling users to to stage data to multiple heterogeneous clouds continuously.

Trilio supports bring-your-own-keys for data encryption and can be integrated with third-party key management systems like HashiCorp Vault.

Besides the enterprise subscription (licensed per worker node, vCPU, or cluster), the product is available for a free 30-day trial with an unlimited number of nodes. There is also a basic edition with a 160-vCPU limit aimed at testing, small organizations, and developers.

Strengths: Trilio has a balanced solution with broad support for managed cloud solutions, distributions, major databases, and application platforms and very deep integration with OpenShift. The disaster recovery feature set is very mature and is storage agnostic.

Challenges: Usability is somewhat lacking compared to the rest of the pack. Proactive security measures like anomaly detection and security scanning are immature though being actively developed. No SaaS option makes adoption slightly more cumbersome for some.

Veritas

A well-known brand in the data protection industry is Veritas, which has included support for Kubernetes in recent releases of NetBackup. NetBackup's Kubernetes support is an extension of the regular NetBackup product, requiring a full NetBackup installation. Kubernetes-specific components are deployed to each cluster via an operator.

NetBackup currently supports backing up entire namespaces only, with support for more granular selection using labels coming in a later release. There is no autodetection of applications currently.

Backup jobs are policy-based. While NetBackup supports much more than just containers, so far, there is no logical construct to define an application across, say, an RDS database, a container, and a file share. NetBackup supports creating backups of Amazon RDS, though, as part of its CloudPoint feature set, which is integrated into NetBackup and usable from the same UI as its Kubernetes data protection.

The product supports Red Hat OpenShift, AKS, EKS, GKE, and VMware Tanzu currently. The console supports protecting multiple clusters. Heterogeneous restores are supported for migrations across different distributions. Backup targets currently include only S3-compatible storage. NetBackup supports many other targets for non-Kubernetes backups, and it aims to bring that support to Kubernetes data protection in 2023. Current support includes object lock functionality.

NetBackup's operational security features—those that can be expected from a mature solution—include multifactor authentication, encryption at rest and in flight, and RBAC across the product, not just within the scope of Kubernetes. Its ransomware functionality includes anomaly detection.

While there is no specific support for Kubernetes at the edge, Veritas does have a physical backup appliance aimed at edge use cases.

Strengths: NetBackup protects more than just containers, making it a good solution for protecting Kubernetes-based applications when you're already using NetBackup. It has broad support for cloud databases, VMs, and more. It has mature operational security capabilities, including RBAC and MFA.

Challenges: NetBackup's Kubernetes support is not yet mature; its roadmap shows promise but needs more time to materialize.

VMware

VMware is the steward of the open-source Velero project, a cloud-native backup and recovery solution. It underpins VMware's own Tanzu Mission Control, aimed at operating Kubernetes. To use its data protection features, clusters must be onboarded into Tanzu Mission Control. This makes Tanzu Mission Control less interesting to use exclusively for its data protection features. Rather, Tanzu Mission Control and its umbrella product Tanzu for Kubernetes Operations are bigger platforms (that include data protection) and a great choice for those looking for more than just data protection.

The Velero open-source project is a core component used by other data protection vendors, including Veritas, Dell, Red Hat, and Catalogic, underpinning many of the commodity features of those solutions. In this context, VMware positions Velero more as a standard API-based framework to standardize backup operations in Kubernetes environments, both at the back end as well as at the front end, with APIs designed to simplify the interaction among different data protection solutions, orchestrators, management consoles, data movers, and abstraction layers offered by third-party vendors.

Velero is tightly integrated within the Tanzu stack. Data protection operations can be managed through Tanzu Mission Control and take advantage of VMware's cloud native storage functionalities. Velero data protection is included in Tanzu, and users have access to it at no additional cost. It has cross-cluster and cross-cloud support and supports backing up full clusters, individual namespaces, or applications using labels.

Tanzu Mission Control supports snapshot backups for major cloud providers (supporting AWS and Azure) and Kubernetes clusters hosted on vSphere and is compatible with CSI-based storage via plug-ins, or restic and kopia.

Notably, VMware wants to open up Tanzu Kubernetes for Operations to third-party backup solutions via its marketplace.

Strengths: Tanzu Mission Control is a well-featured Kubernetes management solution that includes adequate data protection features. It has a compelling roadmap and new projects that aim to simplify, standardize, and automate data protection operations. It is included in and well integrated with VMware Tanzu.

Challenges: Tanzu Mission Control is much more than just a data protection solution, making adoption for just its data protection features a little excessive for some. Velero, as a standalone backup solution, is currently missing several key features, including ransomware protection, environmental awareness, strong encryption, self-service, mature data migration, multiple-region disaster recovery, and analytics functionalities.

Zerto

Zerto for Kubernetes is the cloud-native successor of Zerto's popular disaster recovery solution for virtual environments. Zerto for Kubernetes reuses the same data replication technology, with the benefit of per-second journal-based recovery of persistent volumes.

The Helm-based installer deploys the application as a Stateful DaemonSet into a cluster. It supports VMware Tanzu, Red Hat OpenShift, and native Kubernetes. In the public cloud, Zerto supports Amazon EKS, Azure Kubernetes Service, Google Kubernetes Engine, Oracle Container Engine, HPe Ezmeral, and IBM Cloud Kubernetes Service. It supports heterogeneous, one-to-many replication among these.

In addition to the per-cluster components, an instance of the Zerto for Kubernetes Manager must be deployed to either the cluster or a third site (like a separate VPC account). It has a minimal resource footprint on clusters by auto-scaling components and instantiating parallel instances based on running jobs. Licensing of Zerto for Kubernetes includes perpetual as well as subscription options.

Zerto's ZKM-PX components sit in the data path and intercept all I/O operations for replication. The product is designed to support local backups, remote disaster recovery, and long-term archival storage to S3, but it can be very effective for simplifying data and application migrations as well as the CI/CD process by adding tags to the Zerto journal during deployments of new application versions to capture that moment in time.

Zerto for Kubernetes does not include a UI but works via an extension of kubectl. It is limited to protecting the cluster it is deployed to, though the manager can manage multiple clusters. Access to the SaaS-based Zerto Analytics platform is included in the product, which provides some overview of multiple deployments. There is no SaaS GUI for the product itself, but since its acquisition by HPe, an integration into Greenlake is being considered.

Zerto for Kubernetes is able to discover applications and their resources to protect them as a whole, making both local and remote restores easy to perform. It currently supports backup selection through annotations. It does not support CRDs (right now, but it's on the roadmap), any database-aware processing, or external databases like AWS RDS.

Zerto for Kubernetes is very efficient and replicates only those blocks necessary to the secondary replication pod, which can run in a separate cluster for disaster recovery purposes, or a local cluster for backup purposes. The journal-based approach enables granularity in capturing checkpoints every couple of seconds and storing them for up to 30 days.

Checkpoints can be copied to a long-term retention repository (such as an S3- or Azure Blob-compatible service) daily, weekly, and monthly, and include immutability features based on the underlying storage service. It has no specific features for ransomware protection right now, but inline ransomware detection is coming in a future release.

Using the same journal-based approach, Zerto for Kubernetes can be used as an application migration tool with very short cut-over windows and includes failover and migration testing, deploying the application into a temporary new namespace, and allowing verification. This is also a proposed workflow for creating temporary copies for testing and validation during application development processes.

Strengths: The remote replication capabilities of Zerto for Kubernetes enable users to achieve very low RTOs and RPOs, offering potential for enhancing data mobility and simplifying migration activities.

Challenges: Even though the core technology is rock solid for disaster recovery and application migration scenarios, the product is rough around the edges for data protection needs. Not all customers accept it being in the data path for production workloads, and the requirement of Zerto on a target site may lock customers in for once-only migrations, even though a time-limited migration-only license exists.

6. Analyst's Take

Data protection for Kubernetes is a very hot topic for practically every organization with container-based applications in production. These apps turn out to be more stateful than the "containers are stateless" hype made us believe and more than Kubernetes can effectively support. Many users tried to transform their applications to stateless architectural models, but the technical and organizational complexity of enterprise applications put a wrench in the works.

With Kubernetes not natively capable of handling data protection of stateful applications, many vendors have stepped in to offer a breadth of solutions to solve the data protection and, by extension, data management problem. These range from table stakes backup and restore to more sophisticated multiple-region disaster recovery and even fully fledged data copy management solutions for test/dev environments, as well as data storage solutions. We see an uptick in support for native cloud database services like Amazon's RDS.

No single solution is perfect for all use cases. That means that if you're looking for a data protection solution for Kubernetes, you need to assess your existing application and infrastructure situation, taking into account existing investments in data protection and considering what kind of data you need to protect across VMs, containers, cloud databases, and other cloud services. That inventory and analysis will allow you to define what specific data protection features you're looking for. The best solution for your organization isn't necessarily the one that ticks the most boxes in our research for this Radar, or even all of them; it's the one that ticks the right boxes at the right price point.

Especially in the still-maturing stateful Kubernetes ecosystem, selecting the right data protection vendor is a non-trivial task because many optimizations and features at our disposal in virtualized or bare metal environments are only now making their way into data protection solutions for Kubernetes-based workloads in terms of storage snapshot support and data optimization techniques such as deduplication.

We see feature parity beginning to emerge between traditional and Kubernetes-specific solutions, including with respect to features for performance and data integrity, two areas where GigaOm noted discrepancies in previous reports.

This evolution suggests that the market for Kubernetes data protection is maturing, and the yearly Radars show many net-new and older features becoming commonplace. 2023 will be the year of differentiation beyond the table stakes and key criteria outlined in this report, with vendors looking at natively supporting more cloud database and storage services and enabling more advanced heterogeneous data copy and migration use cases.

7. About Joep Piscaer

Joep Piscaer
Joep is a technologist with team building and tech marketing skills. His background as a CTO, cloud architect, infrastructure engineer and DevOps culture coach. He has built many engineering and architect teams and culture.

Founder of TLA Tech, a tech marketing firm focusing on cloud-native. Co-hosts TheCUBE sometimes. Blogs at VirtualLifestyle.nl

8. 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.

9. Copyright

Our Research

- Research Calendar
- Cloud, Infrastructure, & Management
- DevOps
- Data, Analytics, & AI
- Security & Risk
- Network and Edge
- People, Processes, & Applications

For Practitioners

- Research Subscription
- Analyst Videos
- TCO & Benchmark
- Radars
- Advisory Services
- Key Criteria
- Business & Technology Impact
- Sonars
- GigaBrief

For Vendors

- TCO & Benchmark
- Radars
- Key Criteria
- Business & Technology Impact
- Advisory Services
- Sonars
- Analyst Videos
- Research Subscription
- GigaBrief
- Value Engineering

Resources

- Blog
- Case Studies
- On-Demand Webinars
- GigaOm Research FAQs
- Guides

Company

- Why GigaOm
- Our Team
- Analysts
- Partners
- Press Room
- Careers
- Contact Us