TRILIO

Top 10 Requirements

For Cloud-Native Data Protection for Kubernetes

A Guide for Selecting the Right Data Protection for Kubernetes

January 2021

Abstract

CLOUD-NATIVE DATA PROTECTION IS CRITICAL FOR KUBERNETES

Kubernetes development is increasingly attracting DevOps teams. According to the CNCF, the worldwide number of cloud-native developers jumped almost 38% from 4.7 to 6.5 million in 2020!

It allows developers to deliver faster, more agile, lower-cost development; easier management; and better customer service. Plus, it helps avoid vendor lock-in by creating apps that can seamlessly move among different cloud platforms.

Interestingly, though, a fall 2020 report found that nearly half of IT professionals have delayed rolling out at least some containers and Kubernetes (K8s) applications because of concerns related to resiliency and data protection.

So, how do you go about choosing a cloud-native data protection solution that can provide backup integrity in these dynamic environments? This eBook covers the top 10 requirements for cloud-native data protection for Kubernetes.

CONTENTS

Cloud-Native Development

CONTAINERS

Developers are increasingly using containers, each of which includes coding and a library of resources (data), to run a workload or deliver a service in dynamic cloud environments.

Reasons for this evolution include:

- Cost savings
- Greater flexibility
- Faster disaster recovery
- Better application scalability

KUBERNETES FOR CONTAINER ORCHESTRATION

Kubernetes (K8s) builds and manages software containers. Individual containers can be organized into pods and clusters that are "orchestrated" to collaborate by the K8s platform.

Kubernetes is optimized to run applications in the cloud. And, like LINUX, K8s is open source and can be downloaded for free (bonus!) from GitHub.

Enterprisersproject.com reports that 78% of companies use K8s for managing container deployments. Adding Red Hat Openshift and Rancher, **Network Computing** puts Kubernetes usage even higher at 86%.

10 requirements to selecting data protection for Kubernetes

Because K8s is so hot, many data protection vendors are rushing to say they can fill the need for container backup. However, a shortage of true expertise in this evolving field is a huge problem! Playing on Newton's second law, if you're a developer charged with selecting the best path, it takes a lot of work to research the mass of mostly incomplete information to find the right data protection solutions for Kubernetes. But, no worries! We're here to help lighten your load and make the task of selecting data protection a lot easier.

Next, we'll explore top 10 requirements for cloud-native data protection for Kubernetes. it safeguards:

1. Kubernetes-native and cloud-agnostic
2. Application- and namespace-centric
3. Application discovery
4. Hooks, policies and security
5. Backup repository choice
6. Open backup schema
7. Intelligent restore
8. Integrations and visualizations
9. Certifications
10. Efficient data handling

1 Kubernetes-Native and Cloud-Agnostic

The tool you pick needs to be built specifically for Kubernetes and indifferent to the cloud on which it operates at any given time (e.g., hybrid, multi-cloud, or any ever-changing combination thereof). The backup solution you choose must work directly within the Kubernetes layer, drawing context from K8s and pushing policies back to it for orchestration and enforcement.

Data protection is best implemented directly within the platform it is to protect. It must also be cloud-agnostic to avoid restricting apps and services to a specific cloud provider.

Non-K8s tools can't adequately address multi-tenancy

A cloud-native data protection solution innately enables multi-tenancy, self-service and role-based access control without additional complexity. Importing legacy VM solutions into multi-tenancy applications will become a never-ending, uphill battle.

Non-K8s tools will always be the weakest link

K8s enables faster application development and delivery. Data protection and other products introduced within a K8s landscape cannot compromise applications nor limit the overall implementation of K8s schemas. Further, operations need to leverage abstractions directly within Kubernetes, without requiring further integrations with IT infrastructure.

Non-K8s tools can't scale without bottlenecks

K8s enables applications to scale up and down seamlessly based on workloads. Any legacy infrastructure solution will lead to bottlenecks. Only a K8s-compliant solution assures strong data protection while allowing applications to scale and migrate freely.

Let's do a quick rundown. The right security tool you pick must include:

- K8s constructs and tools to manage and operate solutions
- Restful calls to the API server
- Custom controllers
- Custom resource definitions for the maintenance, packaging and lifecycle management of helm charts and operators

Any separately-managed solution can never be nimble or agile enough for dynamic Kubernetes environments.

TRILIO

2 Application- and Namespace-Centric

Your tools must support all applications in their various forms. It's critical!

Unfortunately, there's no formal definition of an "application," and that's a huge challenge (and paradox) in K8s implementations today. Any of the following can be called an application individually or collectively:

- Simple pod
- Deployment
- StatefulSet
- DaemonSet

Fortunately, there are several techniques for application discovery, deployment and lifecycle management:

Labels

These enable you to tag various components in single or composite applications. Backing up K8s by labels is a basic technique that helps you discover and protect associated applications.

Helm Charts and Operators

Helm charts and operators are the de-facto tools for deploying and maintaining K8s from day one on through day N. Since these provide ongoing value, it's essential that these tools are protected along with their applications as a single unit. Otherwise, you could miss out on critical options like "helm upgrade and rollback," which allows you to use a single command to upgrade or restore previous configurations. Helm charts can also be combined with operator apps to extend functionality for additional software and use cases. Data protection technologies must preserve all these disparate processes together, along with the app and associated metadata.

Namespace

Namespace backup is another popular method. Due to the lack of standard definitions, some customers establish 1:1 mapping between application domains and their namespaces. This provides fail-safe protection to make sure that critical application components aren't overlooked during backups. However, like all good things, this comes with a heavy (literally) price. The trade-off is that these backups are heavier than otherwise required, and mappings can be burdensome to maintain.

3 Application Discovery

Applications and their footprints in K8s clusters and pods are highly dynamic. It's basically impossible to manage any product or solution if you have to keep track of different components and then manually update labels, helm chart, and operator backups.

It's simply too much! Seriously.

The solution you choose must provide the capability to **visually** view/select application resources within a cluster through various lenses/views — and the ability to mix-match items across those provided views. This is really important!

In the absence of an application definition in Kubernetes, this enables a user to define her/his/their own unit of application. That way, she/he/they can protect and grow to that specific definition as the application footprint grows. An example would be the ability to pick a helm chart, and concatenate that with a label-based item (and reuse a protection definition created earlier).

Application discovery needs to be simple, automatic and encompass both vertical and horizontal extensions :

Consider service discovery

Kubernetes uses the service discovery process to connect to a service or application. It's the mechanism that discovers applications and presents them to users.

Tools have to consistently work together

A helm chart and operator app must be able to work together indefinitely. Otherwise, future changes in the scope of work may lead to an operator application relying on a simple label app for pre-process data.

Backups must preserve operational relationships

A backup must include helm and operator applications to allow scope extensions for newer applications that must be protected as combined units.

4 Hooks, Policies and Security

As we mentioned earlier, security/compliance issues are primary reasons for delayed K8s rollouts. Developers have to consider data protection and compliance from the beginning and at every stage of container development.

Hooks

Hooks ensure application consistency and let you add workflows before and after data protection operations. Hooks also help ensure compliance with policies and accommodate disparate applications.

And just to be clear, multiple options for injecting hooks are absolutely required! Applications may use hook injection in various ways (e.g., via a command-line library or a REST API).

Hooks are another example of why data protection solutions must offer multiple options to accommodate application variability.

Automated policies

You need to create policies for automatic, recurring backups in order to protect applications. Also, your policies (and those backups) need to be retained for compliance. Pick solutions with feature-rich, simple-to-manage scheduling and retention policies that are compatible with your K8s environments.

You must be able to manage data protection policies and compliance requirements with a few clicks or command-line instructions.

Security is paramount

Data protection is not just about storing data, but doing so reliably! Your data protection solution needs to support your compliance policies.

The policies need to be in alignment with regulations and also with your requirements for data security, transfer time, storage location, capacity and frequency.

Choose solutions that let you meet your organization's data protection and disaster recovery targets without compromises.

5 Backup Repository Choice

Any tool you pick has to support APIs for varied backup platforms, such as S3 and NFS.

S3

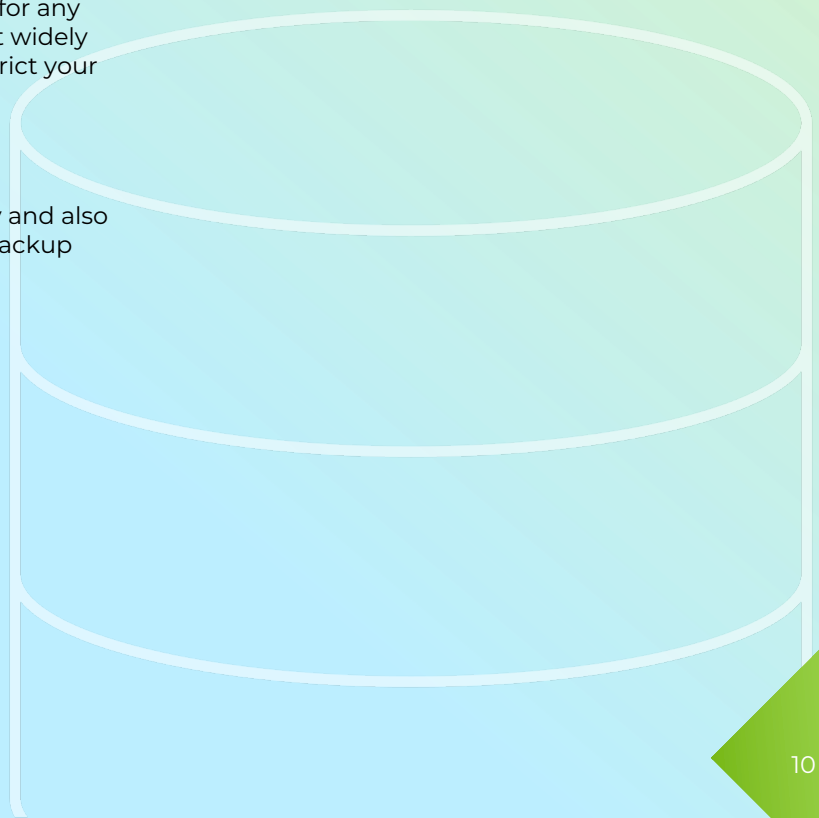
This is a simple, scalable and popular object storage target for backing up K8s. S3 guarantees 11 nines availability and enjoys wide acceptance. AWS is in strong demand from both enterprise and SMB clients. And S3 supports standard APIs adopted by many on-prem storage vendors as well.

Note: Offering options for S3 storage is a basic requirement for any storage vendor. Data protection solutions also must support widely available storage solutions. Avoid backup products that restrict your backup options and don't offer easy-to-deploy APIs!

NFS

Traditional file storage protocols like NFS are still used today and also need to be supported. Many companies already have NFS backup systems in place and expect their vendors to use them.

.



6 Open Backup Schema

Your chosen data protection tool needs to support an open schema so backups can be stored and accessed whenever and wherever you need.

Flexible architectures

IT today relies on highly distributed architectures. In an enterprise network, there may be hundreds or even thousands of containers in core data centers, on public clouds and in edge servers. Backups need to offer flexibility to cope with diversity.

Multi-cloud

Over 90% of organizations today have multi- or hybrid-cloud strategies. Your solutions should be able to move seamlessly between different cloud providers.

Supply chain diversification

Performance, cost and avoidance of single-points-of-failureare driving demands for vendor diversification. Another important consideration is avoiding lock-in.

Now more than ever, organizations need to make sure their data is available on demand, whenever and wherever they need it. **Don't stand for proprietary** data backup products that depend on a specific architecture or provider!

7 Intelligent Restore (selective, granular options)

Cloud-native environments are collections of various microservices. Many abstractions are necessary for applications to run seamlessly and enable Kubernetes clusters to run on any infrastructure.

Protecting data in a fine-grained environment requires options, like:

- **Selective restore**
  Restore specific items (K8s objects) from the backup.
- **Granular restore**
  Restore specific items from a data volume.
- **Restore configuration**
  Massage/tweak object names and specific YAML definitions prior to restore.

In order to accomplish this, your solution must:

- **Back up individual microservices properly**
- **Ensure abstractions from new apps or different clusters can be leveraged post-restore**
  Perhaps it's less relevant when you're restoring to the same source cluster, but otherwise, portability across namespaces, migrations and clouds is critical!
- **Allow exclusions of items from backups prior to restore**
  Take this example: If persistent volumes aren't required by the target apps, excluding them can significantly improve RTO.

8 Integrations and Visualizations

The diverse and rapidly evolving worlds of IT and DevOps employ many tools.

SO. VERY. MANY.

Enterprise IT systems demand smooth operations and interoperable products. Any tool you add to the mix simply MUST provide out-of-the-box integrations to widely-used monitoring and visualization tools. Prometheus, Grafana, and FluentD are de facto tools for K8s environments.

Your data protection solution needs to integrate directly with them and maintain consistency with the rest of your clusters, applications, objects and resources:

- Readily allow metrics and endpoints to be exported to a Prometheus database
- Visualization preferences vary among organizations, but, as a starting point, Grafana should be provided to visualize backups, restores, targets and other key metrics (important because CLI-managed products that lack custom user interfaces may fail to display important metrics)
- The tool must also be capable of forwarding well-formed logs to a central repository for single-pane-of-glass management (like FluentD)

Hey, don't forget!

The cloud-native ecosystem will continue to evolve. New data visualization solutions (like all other applications) must uphold Kubernetes and cloud-native principles.

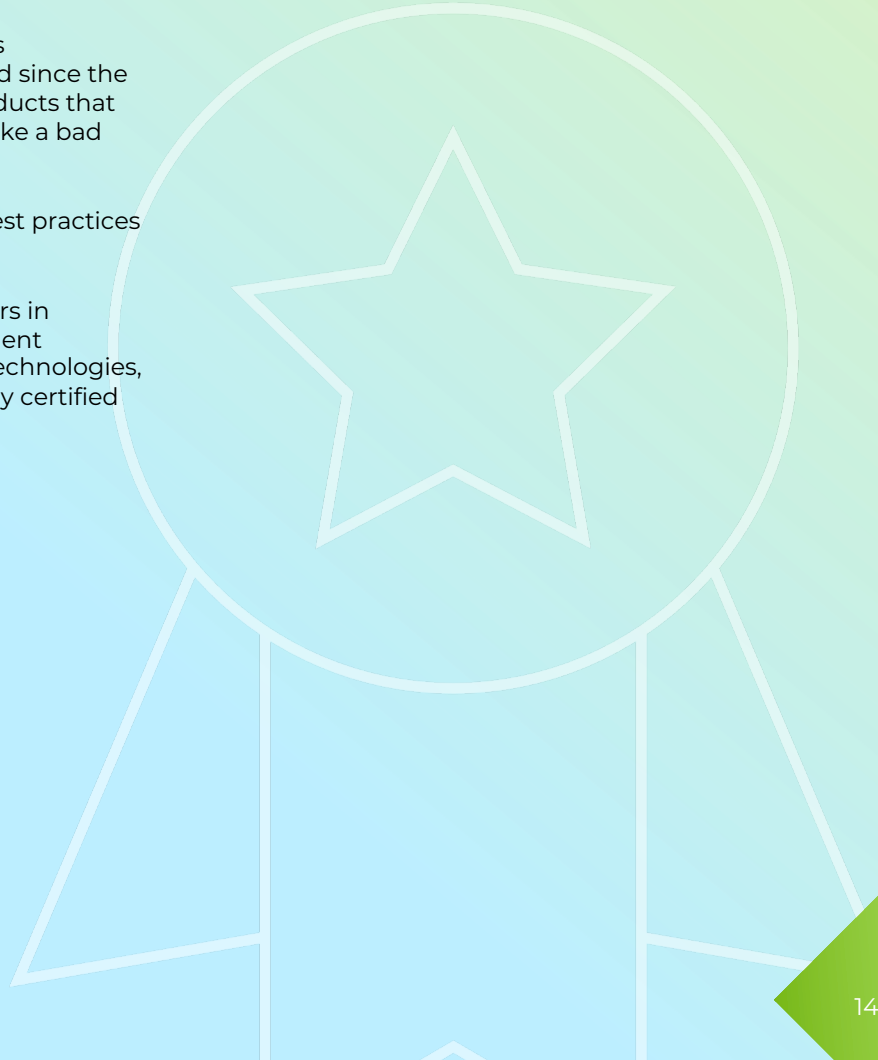
9 Certifications

Vendor's certifications are essential to help verify your solution was built by experts.

As we mentioned, experts in cloud-native technologies (particularly in Kubernetes) are highly sought after. And since the K8s space is evolving so rapidly, you need to avoid products that could become limited or obsolete a few months later like a bad fashion trend.

The strongest method is to pick vendors who follow best practices and recognized data protection principles.

Companies like Red Hat and IBM are recognized leaders in container innovation. These companies offer independent software vendors (ISVs) several ways to validate their technologies, product designs and conformance. Selecting a properly certified vendor can help you avoid unpleasant surprises.



10 Efficient Data Handling (simple, fast and reliable)

The Kubernetes market is the current Gold Rush in the IT industry! Many vendors are hurrying to stake their claims and assert that their solutions are the best available.

Kubernetes has fundamentally changed how computing is done and how data is created, stored and served to applications. With that, the processes for protecting, backing up and restoring data have also changed! Traditional backup solutions simply cannot meet the needs of these dynamic environments.

The data protection solution you select must:

- Accurately collect, store, backup and restore data without errors, modification or corruption at any time
- Not create any security loopholes that could compromise data across the cluster or namespace
- Use data storage space and storage media efficiently
- Provide validated, industry-leading features and benefits, including efficient installation, operations and use
- Be cost-effective to purchase and maintain without worries that it will have limitations of traditional backup solutions

Do your homework here! Meeting these goals when you pick a new data protection product is paramount! Carefully evaluate your potential solutions.

Getting Going

As a hot new market, the cloud-native industry offers exciting new opportunities for digital transformation using Kubernetes, Agile development and microservices. Demand for container products has fueled a rapidly growing market filled with new vendors churning out tons of products.

In older, more established industries, there are clear levels of competency defined by maturity models and industry standards. But, because the dynamic cloud-native industry is still just a baby (constantly innovating and evolving) maturity is subjective and standards are still being defined.

The current landscape includes clear risks in picking the wrong data protection strategies! Lack of confidence in data protection has been holding up hundreds of deployments across the industry, but you now have this guide to help you navigate.

Let Trilio Help

No matter if you're an IT executive, developer or business manager, to thrive in the new age of the cloud you need to know how containers can help you build better, faster and more productive solutions.

Kubernetes container orchestration technologies offer huge opportunities for your company to employ new digital transformation strategies to develop a more agile, competitive and resilient business. And they open pathways to explore new developments such as big data, machine learning, blockchain and AI.

But like with any open-source technology, there are risks that you need to understand and manage. Your best bet is choosing a certified data protection product from a trusted vendor who:

- Is committed to following cloud-native practices
- Has clearly demonstrated capabilities across the above 10 requirements

We hope this ebook helps you analyze your options and make the best choice for your cloud-native projects!

We at Trilio are experts in this field, and have purpose-built a solution that meets all of the above requirements and more!

To learn more about our certified, industry-leading TrilioVault data protection solution, check out our:

- Website
- Videos
- Online labs
- Free trial
- Documentation
- Demo

Learn More

Want to know more?

WATCH DEMO

GITBOOK DOCUMENTATION

ONLINE LAB

DOWNLOAD NOW

TRILIO

WWW.TRILIO.IO