CSI Volume Snapshots: On the Way to Faster and Better Backups

Adnan Abdulhussein

Nolan Brubaker

vmware

Agenda

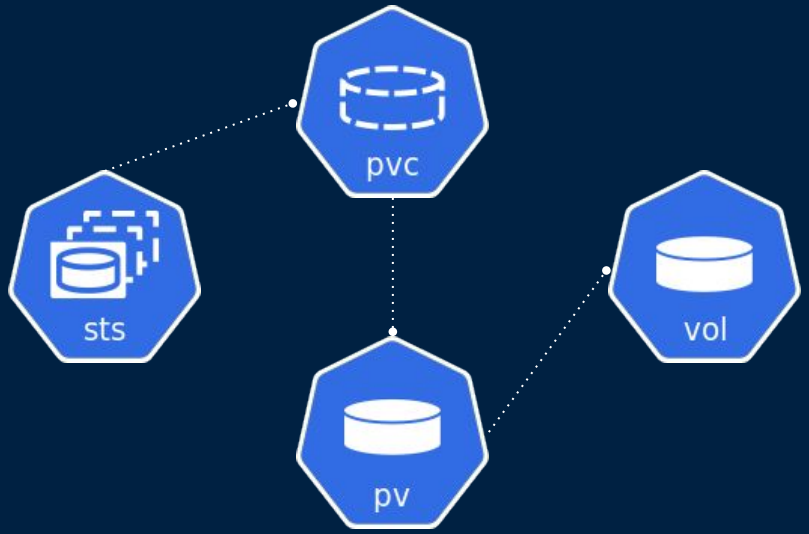
Intro to Container Storage Interface (CSI)

Kubernetes CSI API & Workflow

Demo

Future of CSI

The state of Stateful in Kubernetes



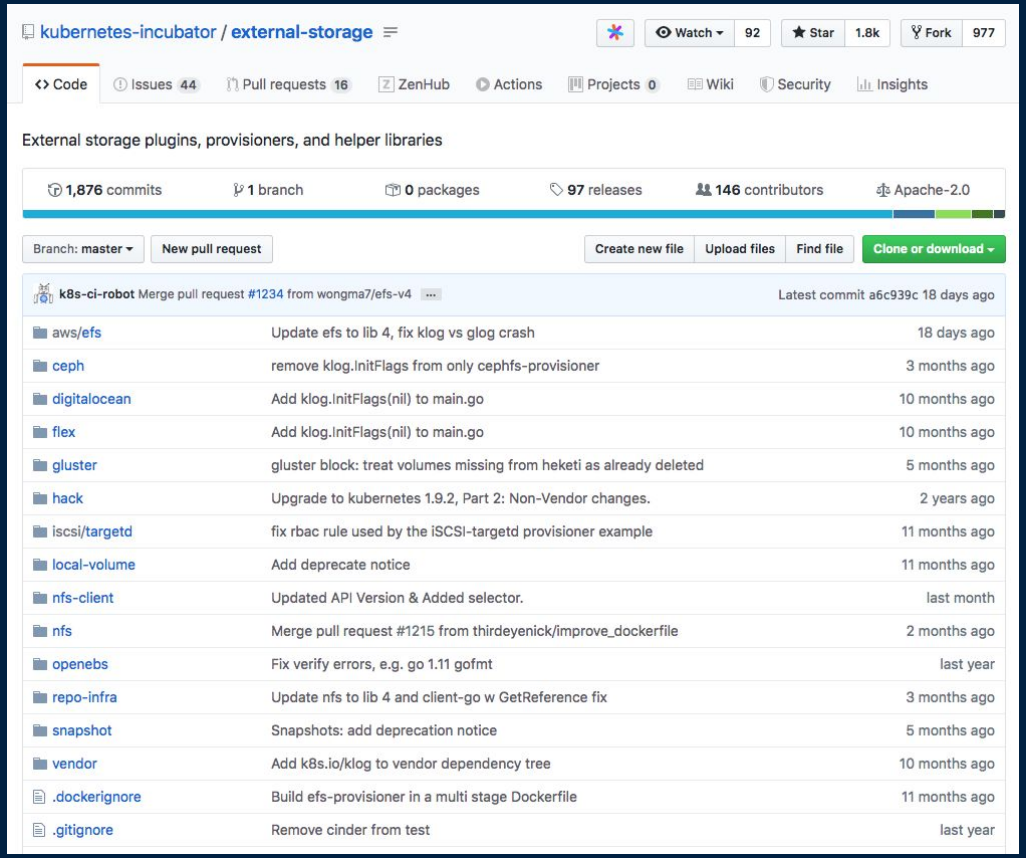
Stable APIs for storage provisioning

- StatefulSets
- PersistentVolumesClaims, PersistentVolume
- StorageClasses

Support for multiple storage providers

Volume Plugin	Internal Provisioner
AWSElasticBlockStore	✓
AzureFile	✓
AzureDisk	✓
CephFS	-
Cinder	✓
FC	-
FlexVolume	-
Flocker	✓
GCEPersistentDisk	✓
Glusterfs	✓
iSCSI	-
Quobyte	✓
NFS	-
RBD	✓
VsphereVolume	✓
PortworxVolume	✓
ScaleIO	✓
StorageOS	✓
Local	-

Not every provider can go into core



kubernetes-incubator / external-storage

Watch 92 | Star 1.8k | Fork 977

Code | Issues 44 | Pull requests 16 | ZenHub | Actions | Projects 0 | Wiki | Security | Insights

External storage plugins, provisioners, and helper libraries

1,876 commits | 1 branch | 0 packages | 97 releases | 146 contributors | Apache-2.0

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

k8s-ci-robot Merge pull request #1234 from wongma7/efs-v4 — Latest commit a6c939c 18 days ago

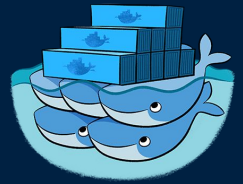
aws/efs	Update efs to lib 4, fix klog vs glog crash	18 days ago
ceph	remove klog.InitFlags from only cephfs-provisioner	3 months ago
digitalocean	Add klog.InitFlags(nil) to main.go	10 months ago
flex	Add klog.InitFlags(nil) to main.go	10 months ago
gluster	gluster block: treat volumes missing from heketi as already deleted	5 months ago
hack	Upgrade to kubernetes 1.9.2, Part 2: Non-Vendor changes.	2 years ago
iscsi/targetd	fix rbac rule used by the iSCSI-targetd provisioner example	11 months ago
local-volume	Add deprecate notice	11 months ago
nfs-client	Updated API Version & Added selector.	last month
nfs	Merge pull request #1215 from thirdeyenick/improve_dockerfile	2 months ago
openebs	Fix verify errors, e.g. go 1.11 gofmt	last year
repo-infra	Update nfs to lib 4 and client-go w GetReference fix	3 months ago
snapshot	Snapshots: add deprecation notice	5 months ago
vendor	Add k8s.io/klog to vendor dependency tree	10 months ago
.dockerignore	Build efs-provisioner in a multi stage Dockerfile	11 months ago
.gitignore	Remove cinder from test	last year

Storage doesn’t stop at provisioning

Backup and Disaster Recovery

Storage providers need to support each orchestration system as well as different backup and DR solutions

Container Storage Interface (CSI)



“Industry standard interface that will enable storage vendors to develop a plugin once and have it work across a number of container orchestration systems.”

CSI gRPC API

```
service Controller {
  rpc CreateVolume (CreateVolumeRequest)
    returns (CreateVolumeResponse) {}

  rpc DeleteVolume (DeleteVolumeRequest)
    returns (DeleteVolumeResponse) {}

  rpc ControllerPublishVolume (ControllerPublishVolumeRequest)
    returns (ControllerPublishVolumeResponse) {}

  rpc ControllerUnpublishVolume (ControllerUnpublishVolumeRequest)
    returns (ControllerUnpublishVolumeResponse) {}

  rpc ValidateVolumeCapabilities (ValidateVolumeCapabilitiesRequest)
    returns (ValidateVolumeCapabilitiesResponse) {}

  rpc ListVolumes (ListVolumesRequest)
    returns (ListVolumesResponse) {}

  rpc GetCapacity (GetCapacityRequest)
    returns (GetCapacityResponse) {}

  rpc ControllerGetCapabilities (ControllerGetCapabilitiesRequest)
    returns (ControllerGetCapabilitiesResponse) {}

  rpc CreateSnapshot (CreateSnapshotRequest)
    returns (CreateSnapshotResponse) {}

  rpc DeleteSnapshot (DeleteSnapshotRequest)
    returns (DeleteSnapshotResponse) {}

  rpc ListSnapshots (ListSnapshotsRequest)
    returns (ListSnapshotsResponse) {}

  rpc ControllerExpandVolume (ControllerExpandVolumeRequest)
    returns (ControllerExpandVolumeResponse) {}
}
```

```
service Node {
  rpc NodeStageVolume (NodeStageVolumeRequest)
    returns (NodeStageVolumeResponse) {}

  rpc NodeUnstageVolume (NodeUnstageVolumeRequest)
    returns (NodeUnstageVolumeResponse) {}

  rpc NodePublishVolume (NodePublishVolumeRequest)
    returns (NodePublishVolumeResponse) {}

  rpc NodeUnpublishVolume (NodeUnpublishVolumeRequest)
    returns (NodeUnpublishVolumeResponse) {}

  rpc NodeGetVolumeStats (NodeGetVolumeStatsRequest)
    returns (NodeGetVolumeStatsResponse) {}


  rpc NodeExpandVolume(NodeExpandVolumeRequest)
    returns (NodeExpandVolumeResponse) {}


  rpc NodeGetCapabilities (NodeGetCapabilitiesRequest)
    returns (NodeGetCapabilitiesResponse) {}

  rpc NodeGetInfo (NodeGetInfoRequest)
    returns (NodeGetInfoResponse) {}
}
```

What CSI means for Kubernetes

Move all storage providers out-of-tree by Q4 2020

Embrace CSI as primary storage integration

Introduce APIs for interacting with CSI drivers

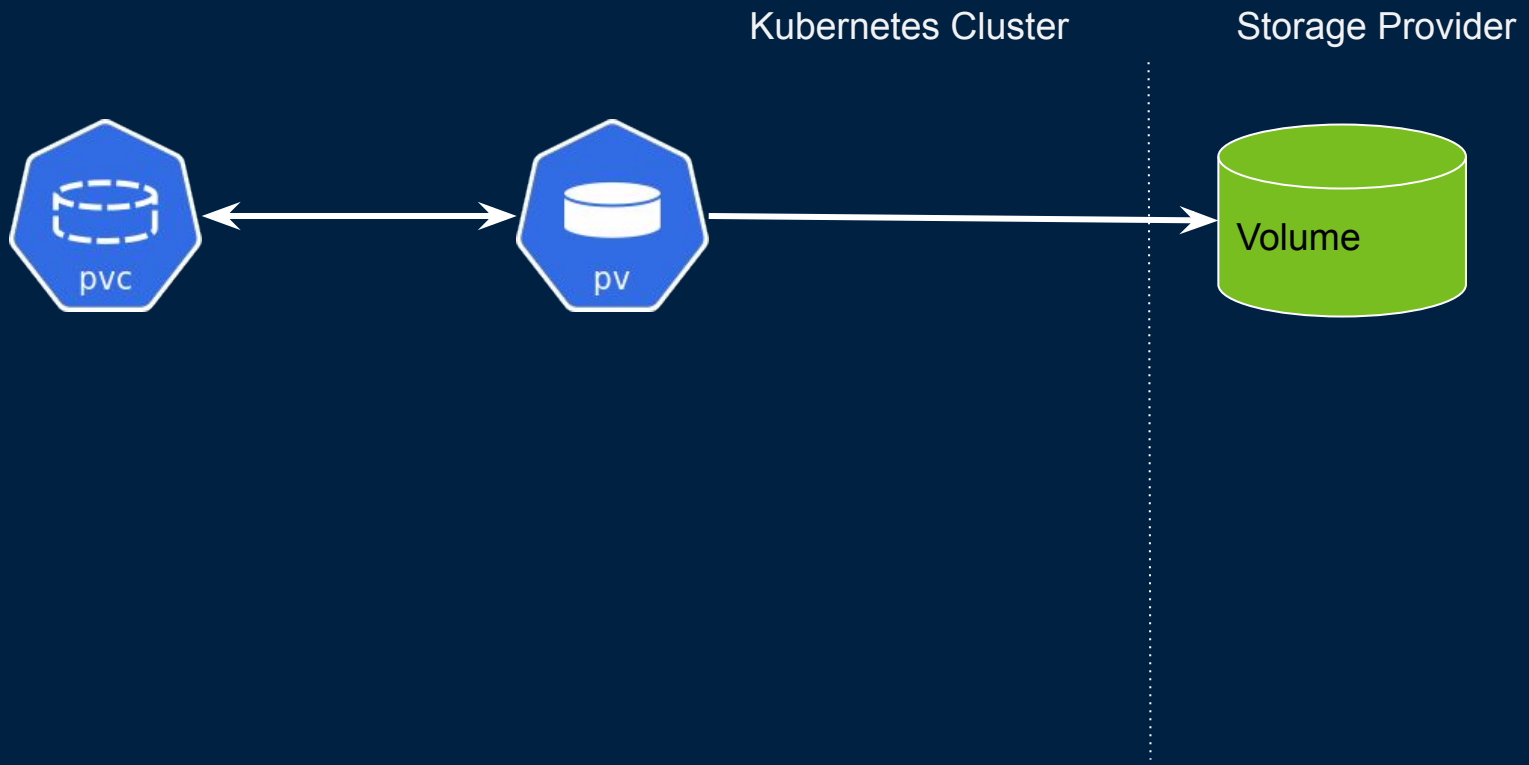
How do we use it?

Key

Data relationship

Code acting

Cluster before taking a snapshot

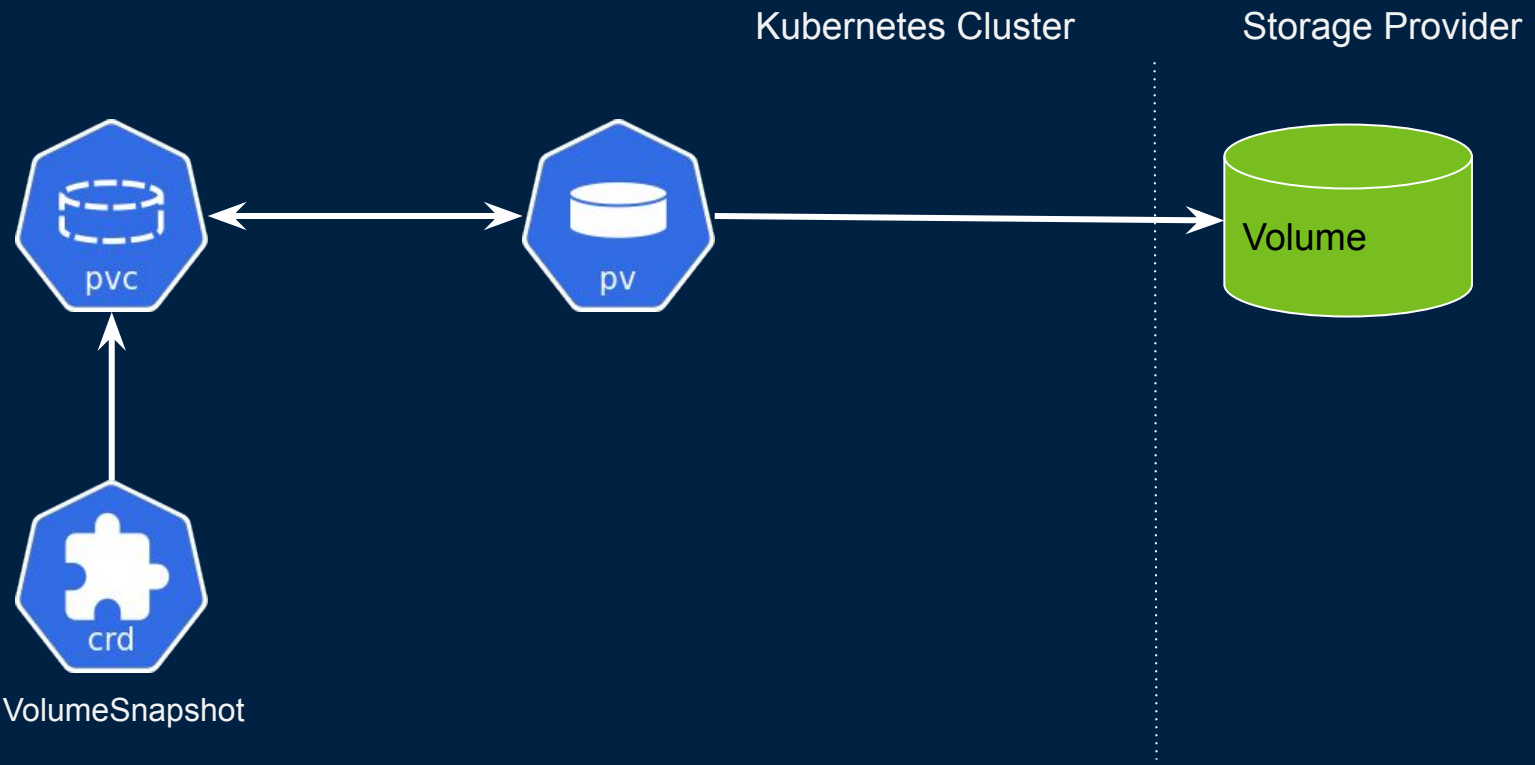


VolumeSnapshot CRD

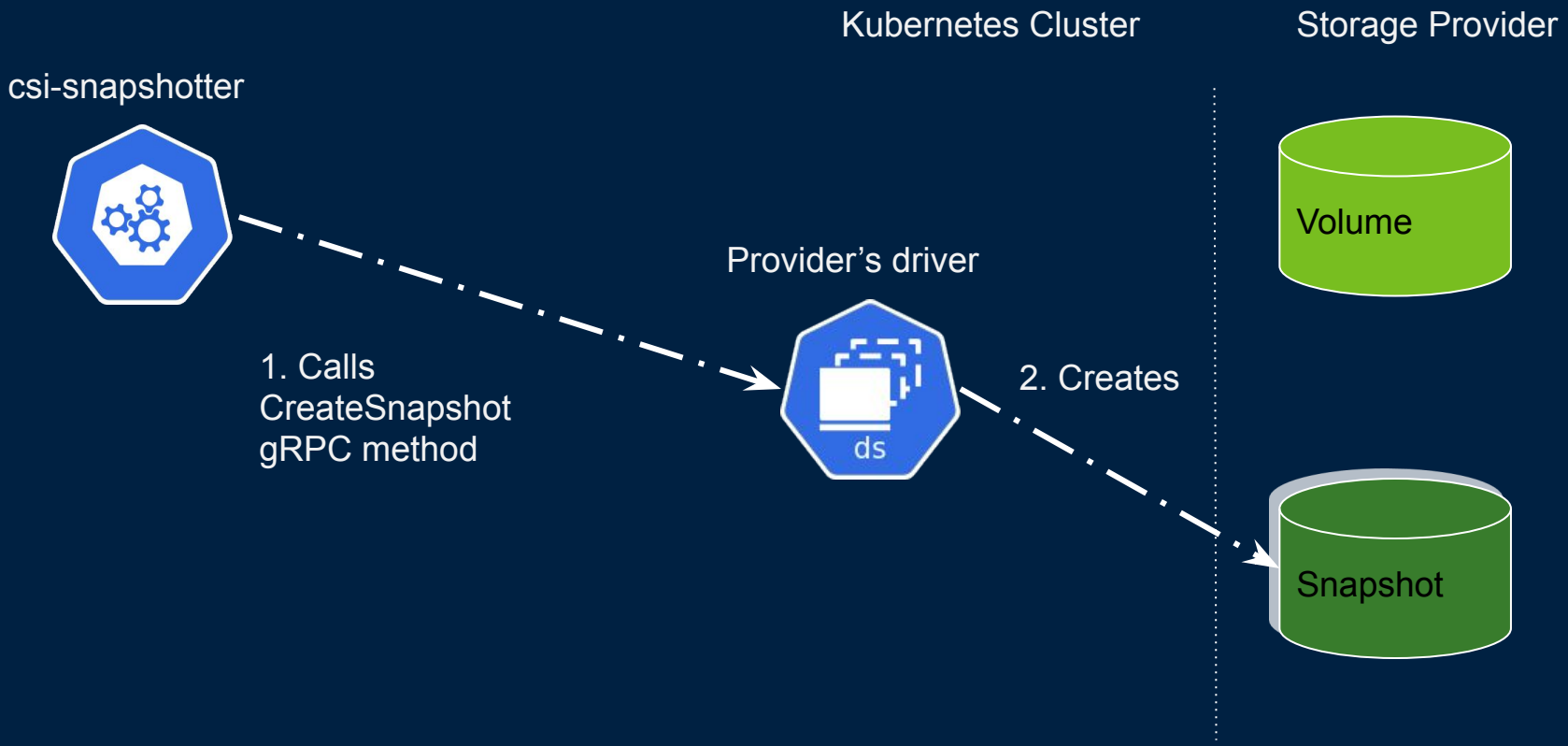
Namespaced CRD representing the snapshot

```
apiVersion: snapshot.storage.k8s.io/v1alpha1
kind: VolumeSnapshot
metadata:
  name: mysql-snapshot
  namespace: wordpress
spec:
  source:
    name: mysql-pv-claim
    kind: PersistentVolumeClaim
```

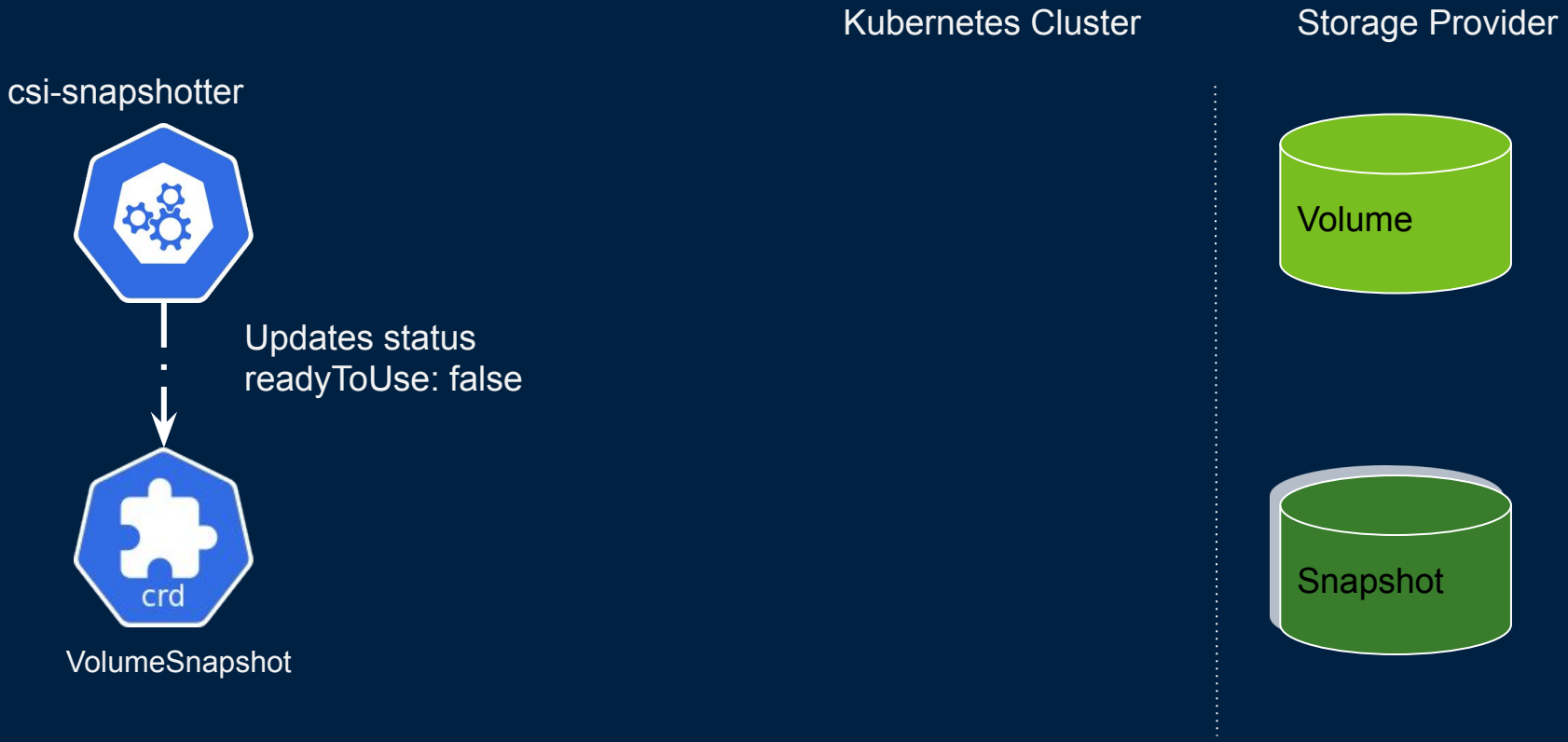
Cluster after submitting the snapshot



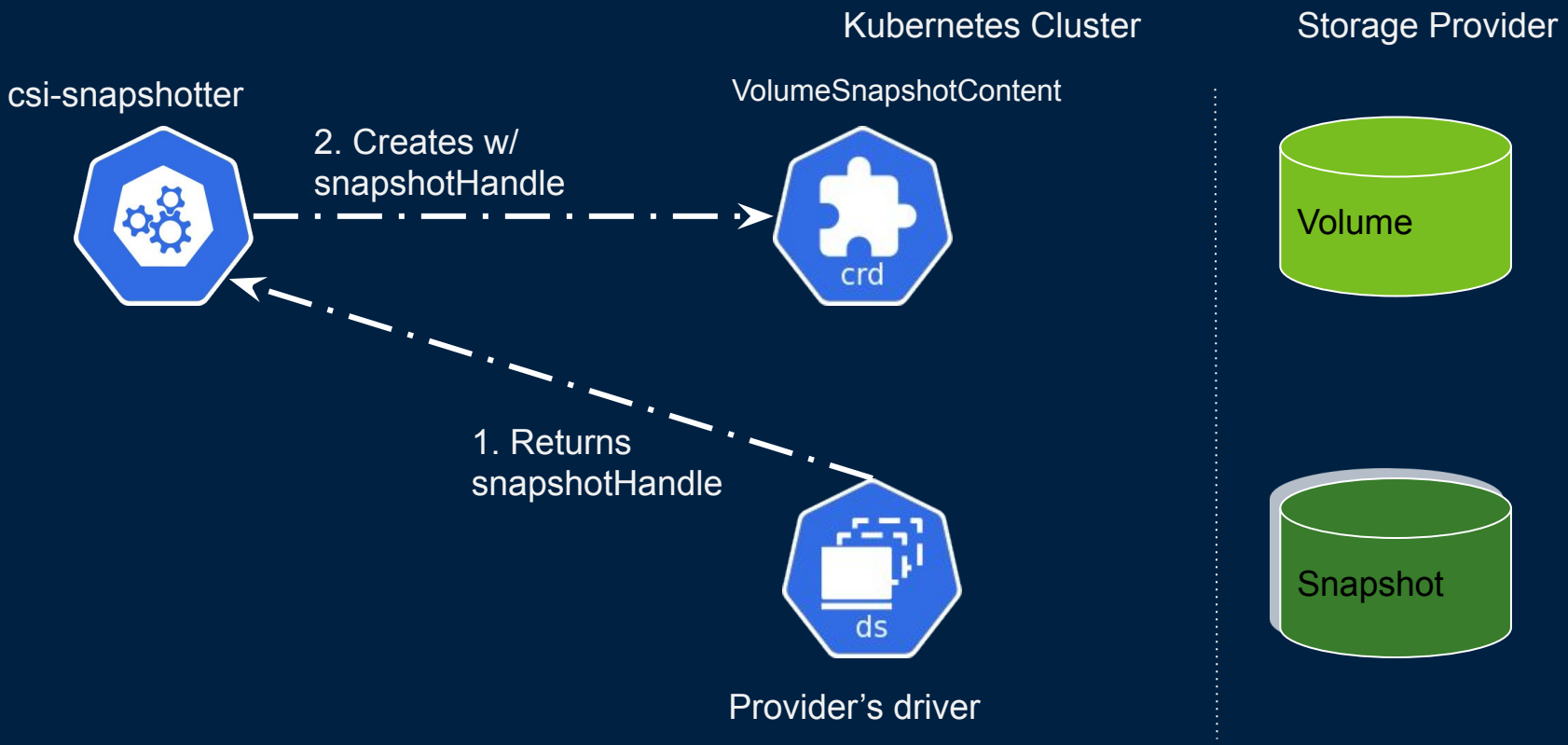
Creating the snapshot and reconciling state



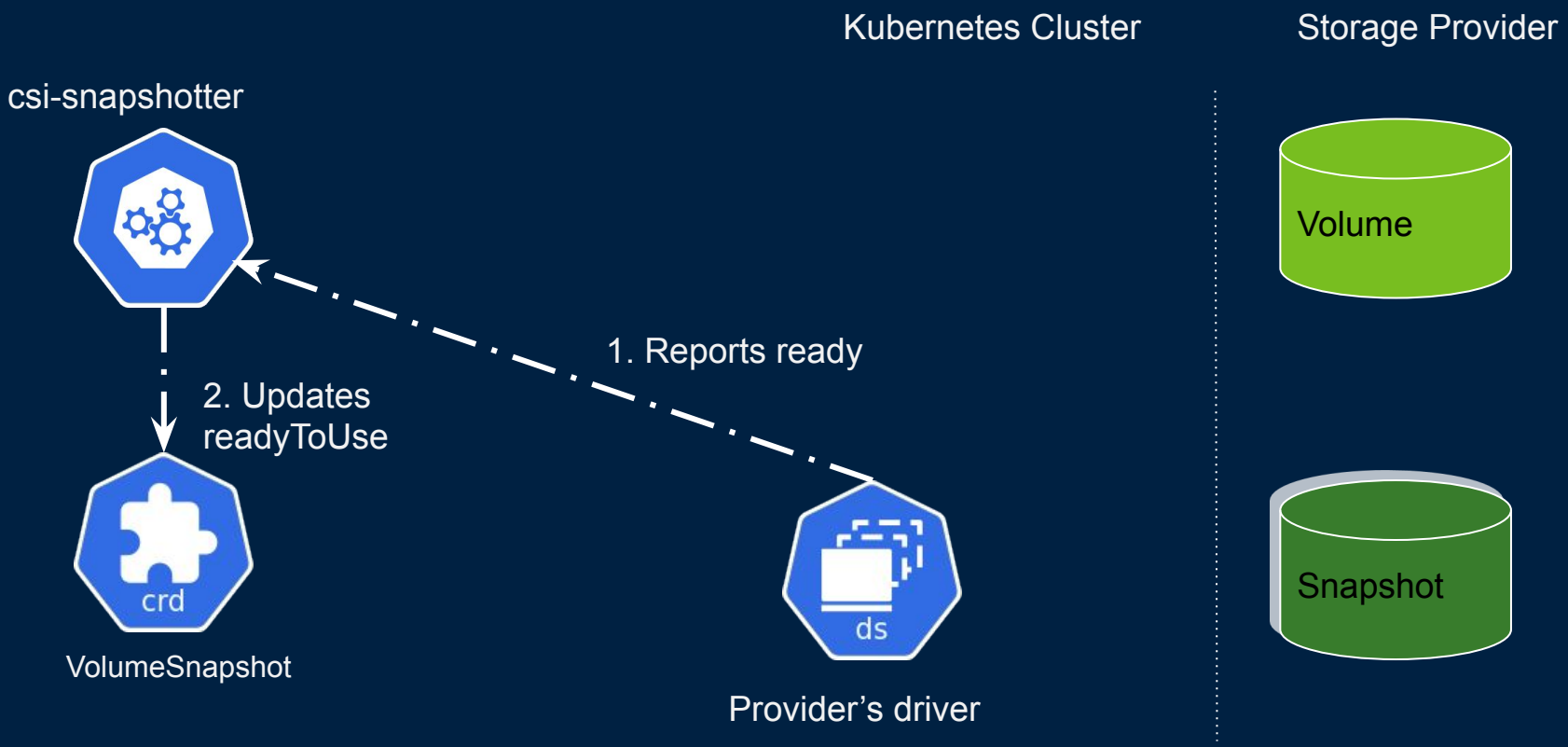
Creating the snapshot and reconciling state



Creating the snapshot and reconciling state



Creating the snapshot and reconciling state



VolumeSnapshotContent

Cluster-scoped, links to the storage provider's snapshot

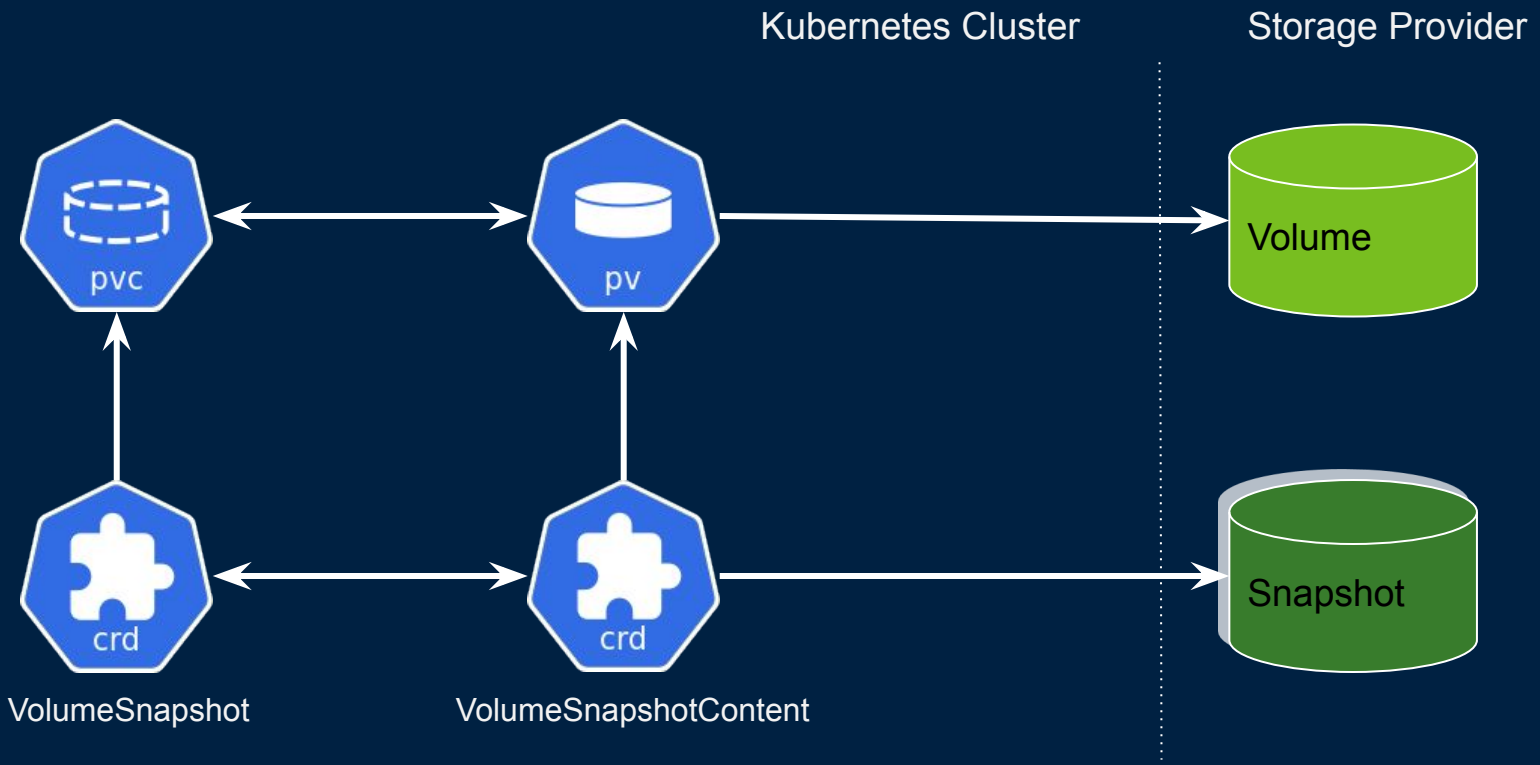
```
spec:
    deletionPolicy: Retain
    snapshotClassName: csi-aws-vsc
    csiVolumeSnapshotSource:
        restoreSize: 21474836480
        snapshotHandle: snap-0fcb35b19b90955e3
    persistentVolumeRef:
        apiVersion: v1
        kind: PersistentVolume
        name: pvc-8bd21b64-024b-11ea-bb60-02c4fcac63ba
    volumeSnapshotRef:
        apiVersion: snapshot.storage.k8s.io/v1alpha1
        kind: VolumeSnapshot
        name: mysql-snapshot
        namespace: wordpress
```

VolumeSnapshot CRD (Ready to Use)

Namespaced CRD representing the snapshot

```
metadata:
  name: mysql-snapshot
  namespace: wordpress
spec:
  snapshotClassName: csi-aws-vsc
  snapshotContentName: snapcontent-ba96923a-...-c4fcac63ba
  source:
      apiGroup: null
      kind: PersistentVolumeClaim
      name: mysql-pv-claim
status:
  creationTime: "2019-11-11T19:37:43Z"
  readyToUse: true
  restoreSize: 20Gi
```

Cluster once snapshot is ready to use



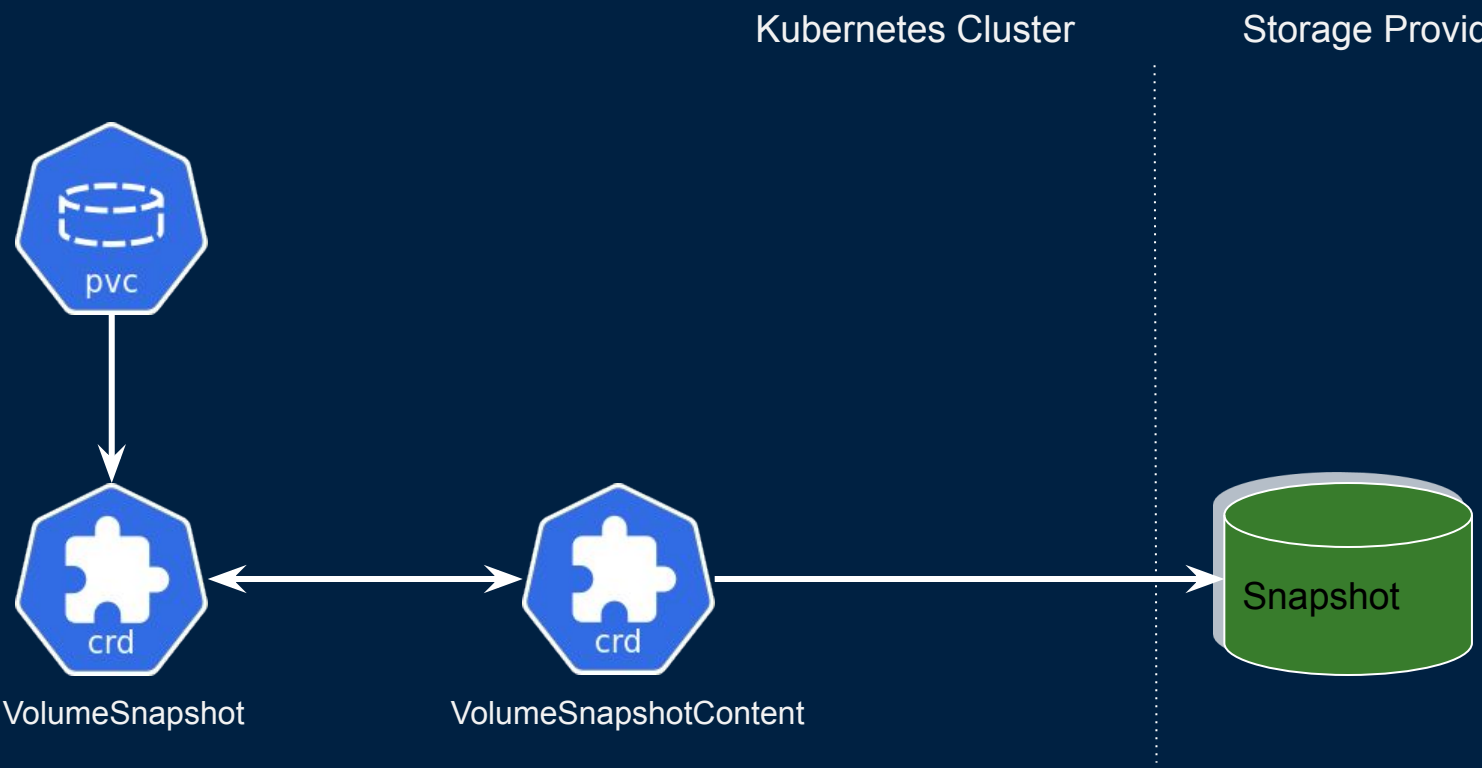
Restoring

Create a new PVC

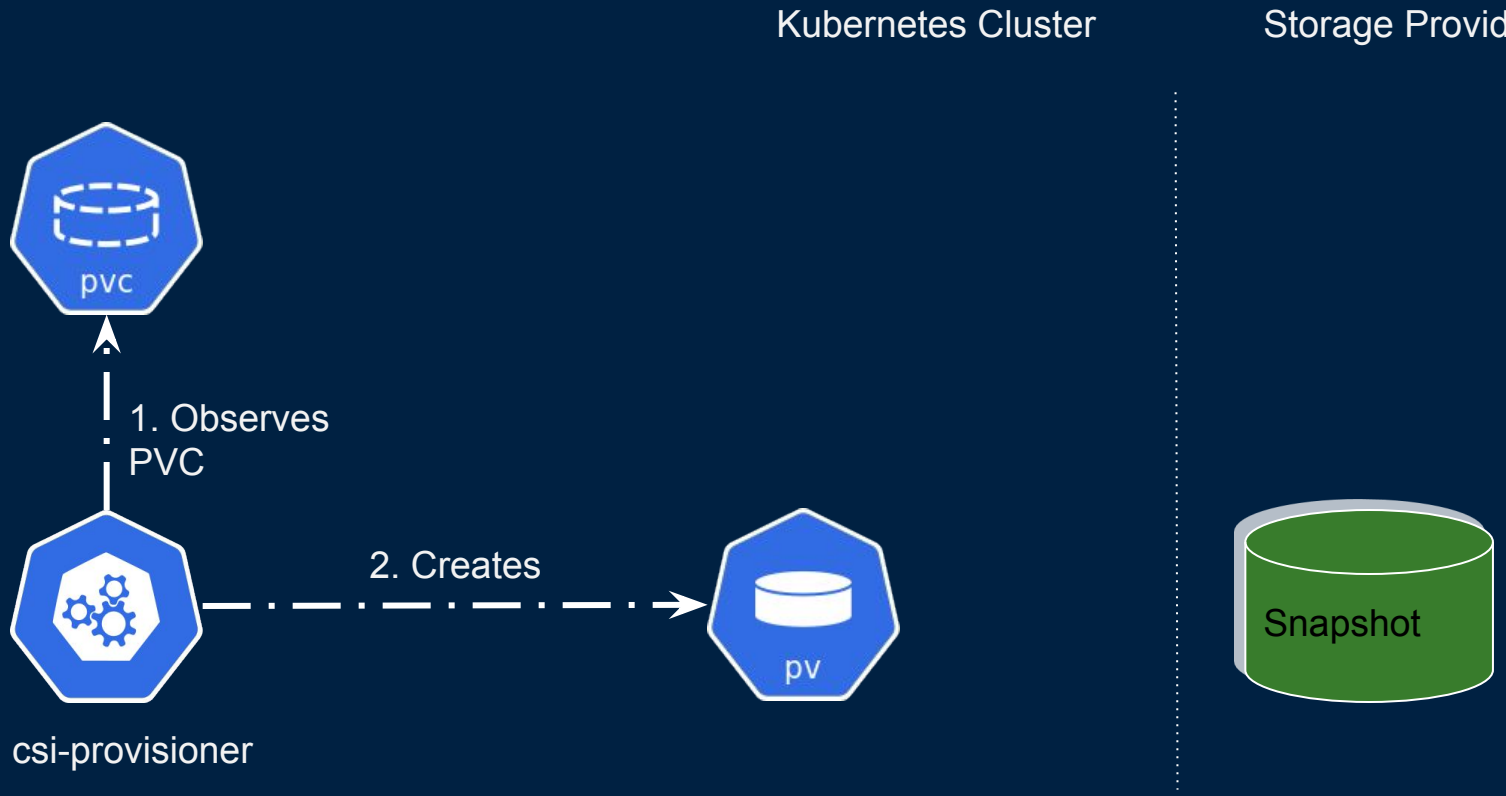
```
kind: PersistentVolumeClaim
metadata:
  name: mysql-pv-claim
  namespace: wordpress
spec:
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
    storage: 20Gi
  dataSource:
    name: mysql-snapshot
    kind: VolumeSnapshot
    apiGroup: snapshot.storage.k8s.io
```

Beta as of 1.17
Alpha since 1.12

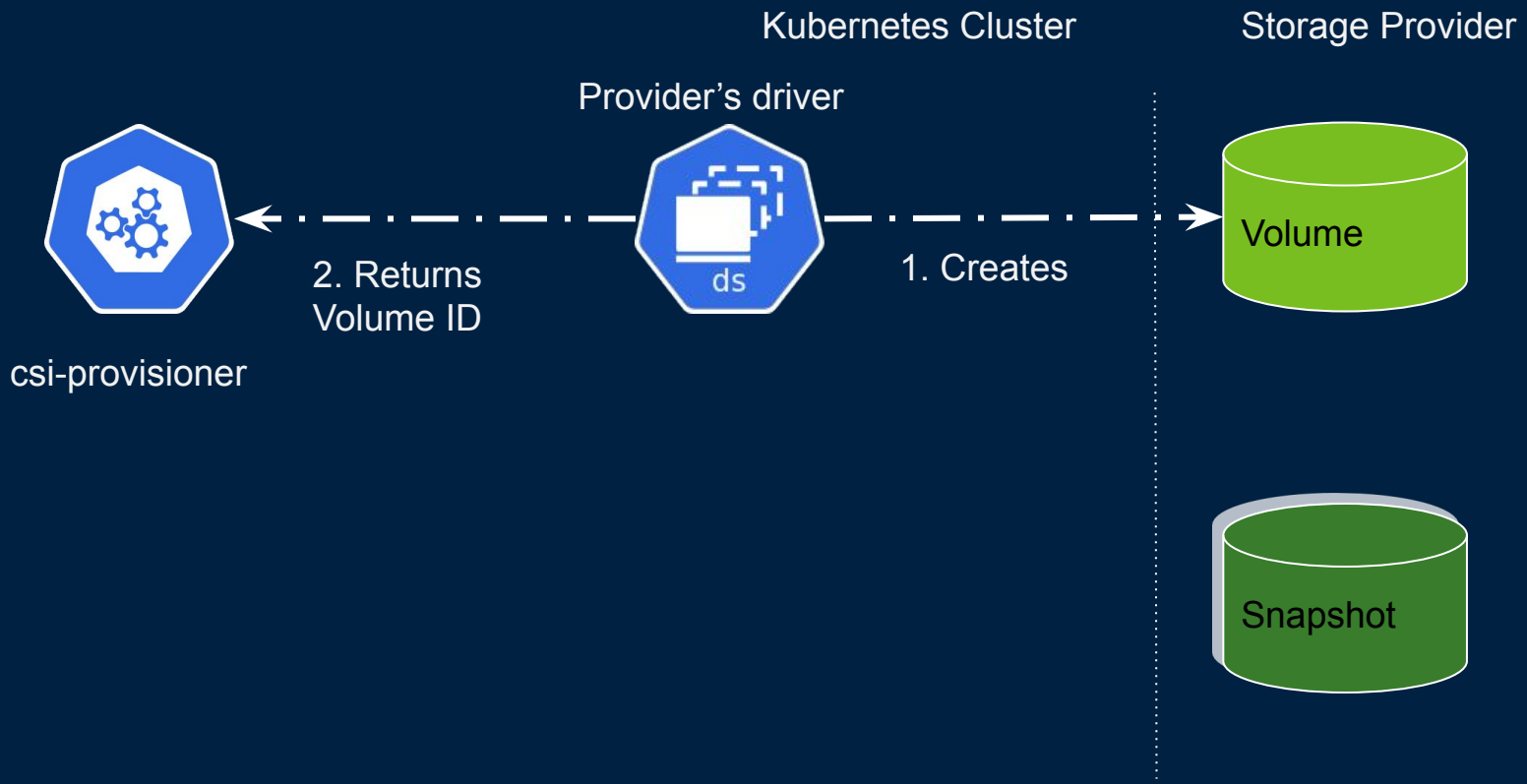
CSI Snapshot CRD Relationships (Restore)



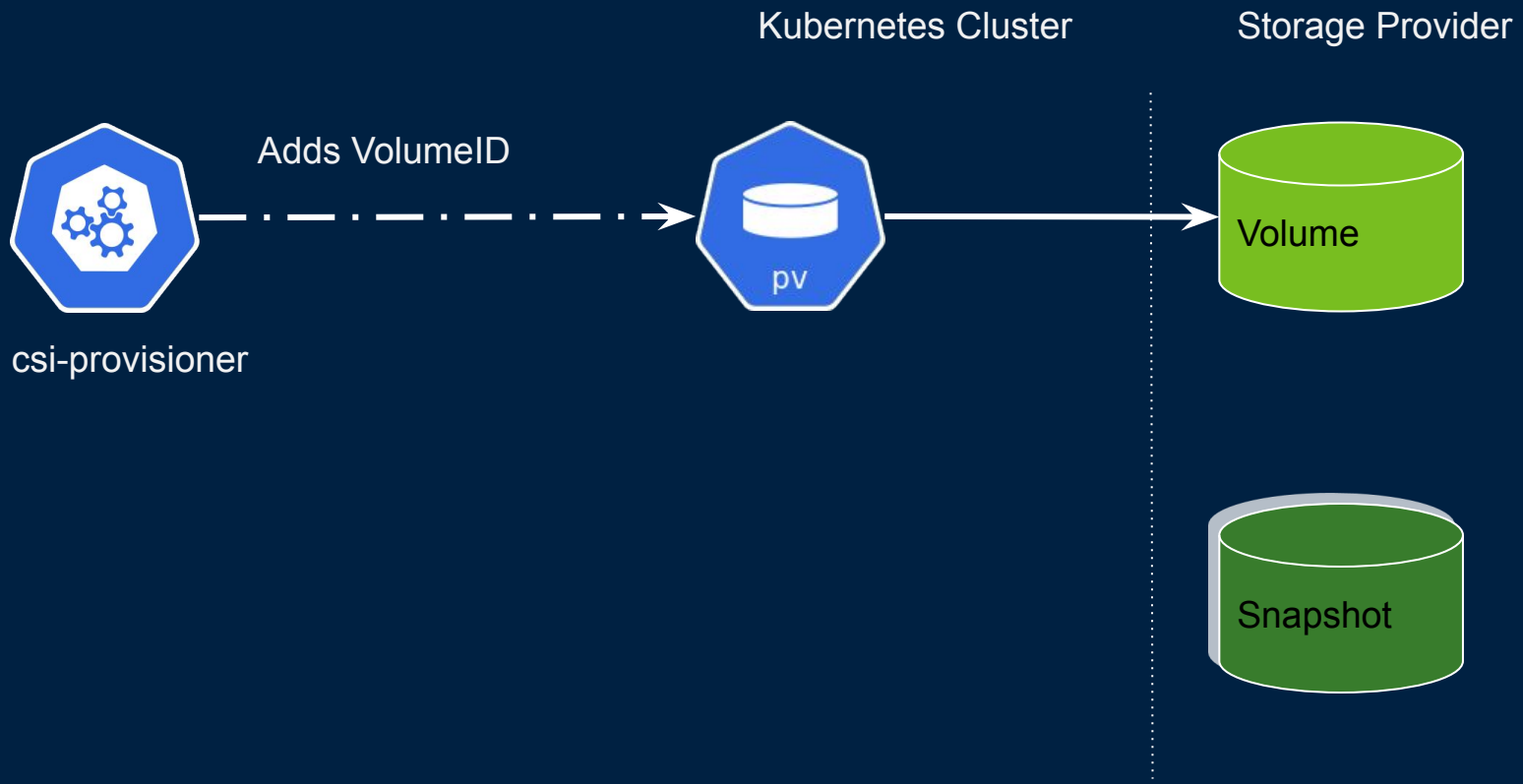
Restoring and reconciling state



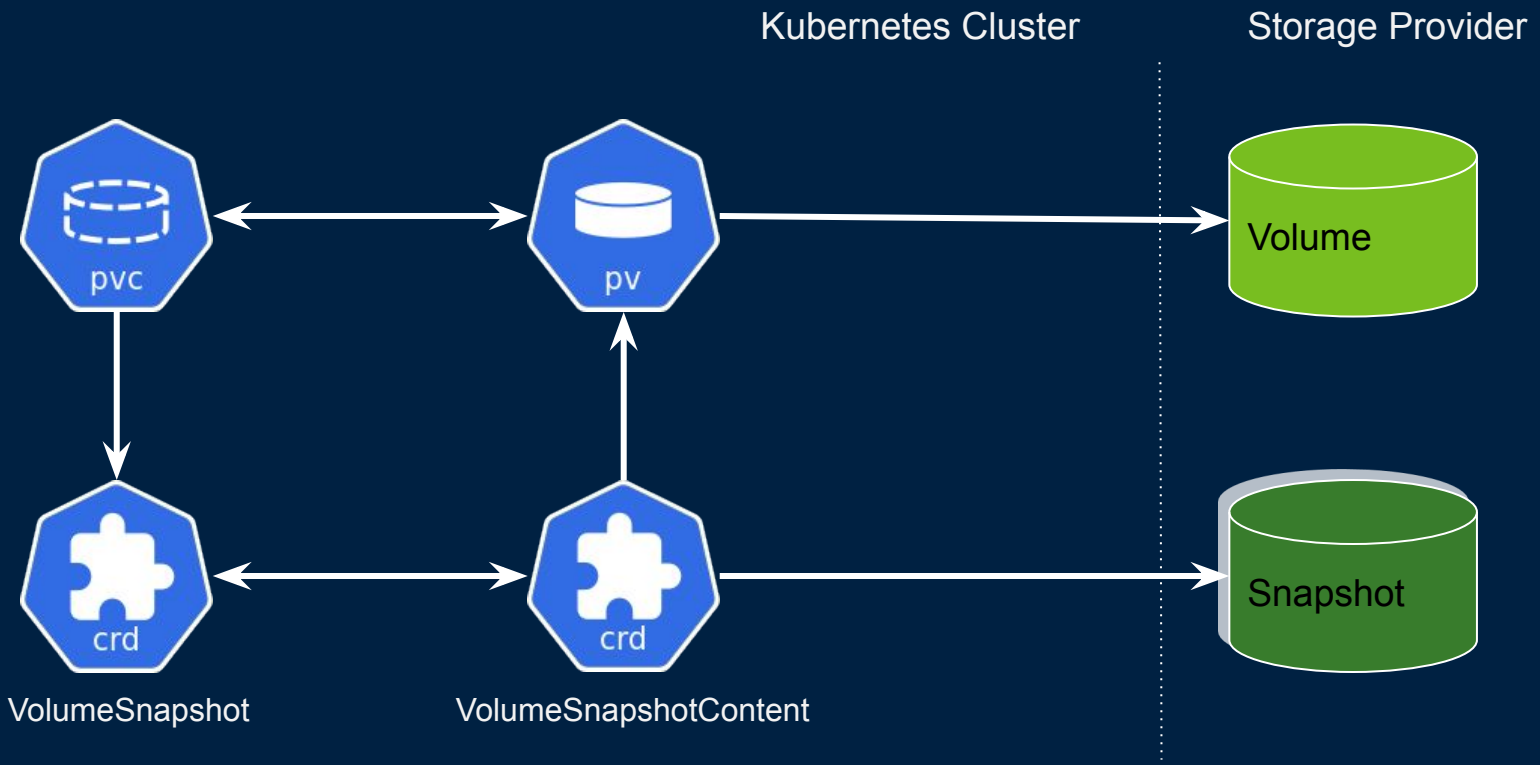
Restoring and reconciling state



Restoring and reconciling state



CSI Snapshot CRD Relationships (Restore)



VolumeSnapshotClass CRD

General Settings for Snapshots

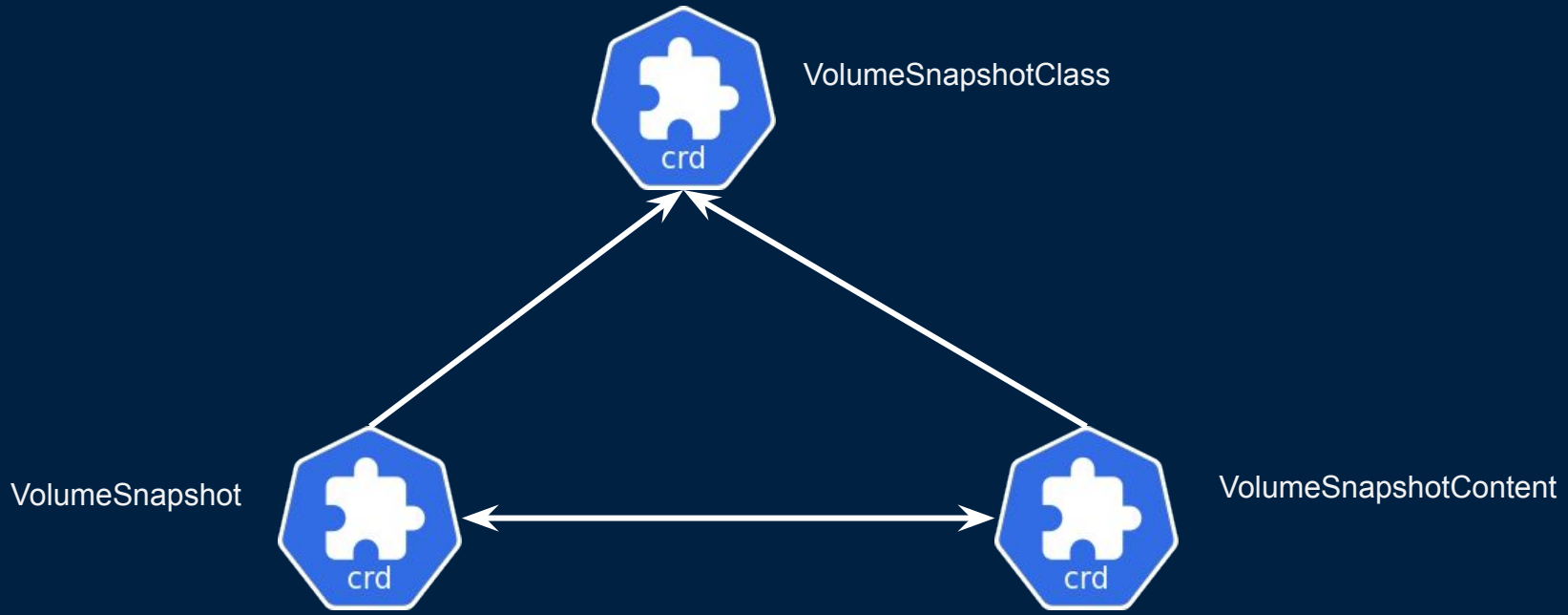
Most significant fields:

Snapshotter
- Name of the CSI plugin that does the work

Parameters
- Key/values pairs to pass to the snapshotter plugin
- Varies based on the plugins

DeletionPolicy
- Retain or Delete

CSI Snapshot CRD Relationships



Demo!

Future of CSI

More CSI drivers with snapshot capabilities

Bring Snapshot CSI APIs to GA

Tooling integration (e.g. Velero CSI plugin)

ExecutionHook API

- Enable application quiescing and unquiescing

Stateful Application Data Management API

- App-level snapshot, cloning and restore APIs

Where to contribute

SIG Storage

- KubeCon SIG Storage Intro/Deep Dive - tomorrow @ 2:25pm
- #sig-storage on k8s Slack
- Thurs 9am PT biweekly meeting
- Mailing list, GitHub

CSI

- #csi on k8s Slack
- https://github.com/kubernetes-csi
- https://github.com/container-storage-interface

Demo: https://github.com/nrb/kubecon2019-demo/

Thank You

vmware®