



September 23-26, 2019  
Santa Clara, CA

# Simplifying Stateful Apps On Kubernetes: A Git Like Workflow

Jagadish Mukku, Robin.io  
Dhanashankar Venkatesan, Robin.io



# Agenda

September 23-26, 2019  
Santa Clara, CA

SDC<sup>19</sup>

- Stateful Application
- Background on Containers and Kubernetes
- Data Management of Stateful App
- Challenges and Solutions
- Q & A

# Spectrum Of Applications

Stateless  
Applications

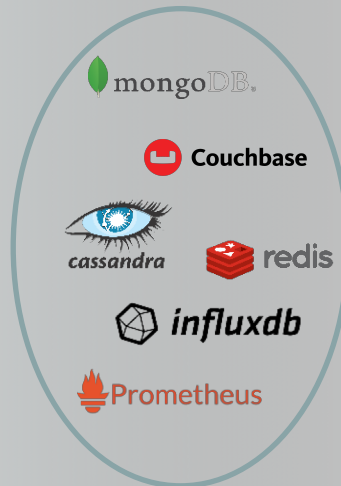


Web  
Apps

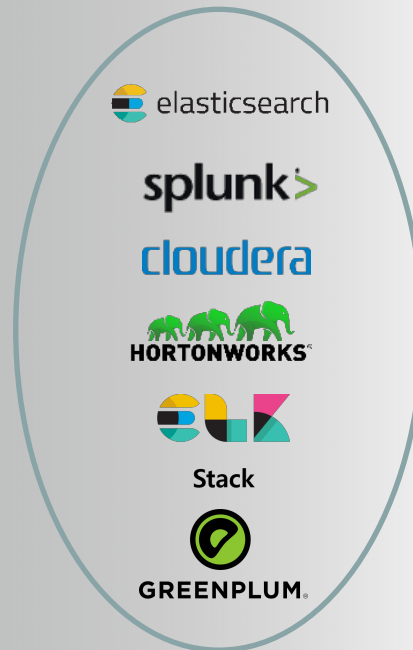
Stateful  
Applications



SQL  
Databases

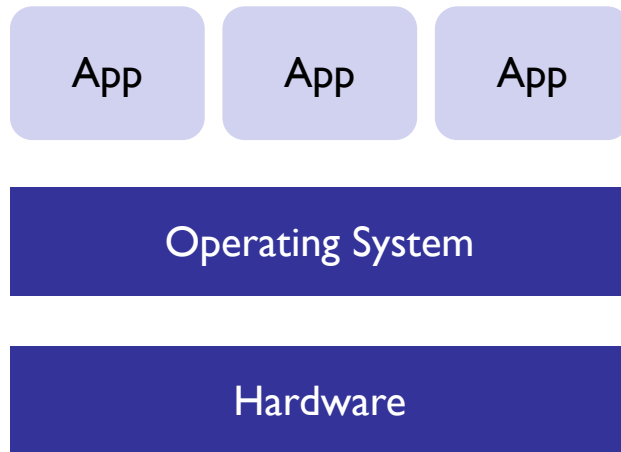


NoSQL  
Databases



Big Data

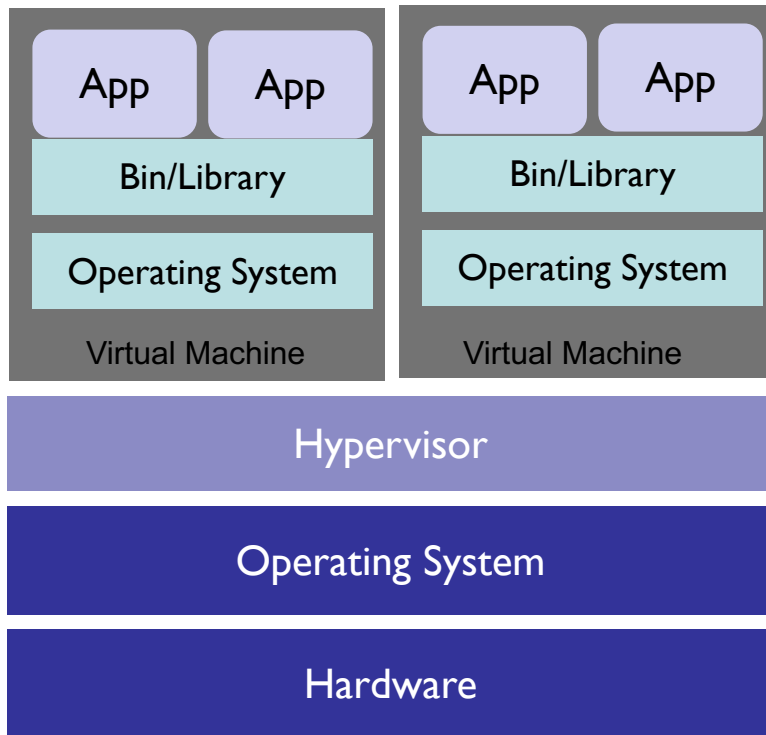
# Traditional Deployment



# Virtualized Deployment

September 19, 2019  
Santa Clara, CA

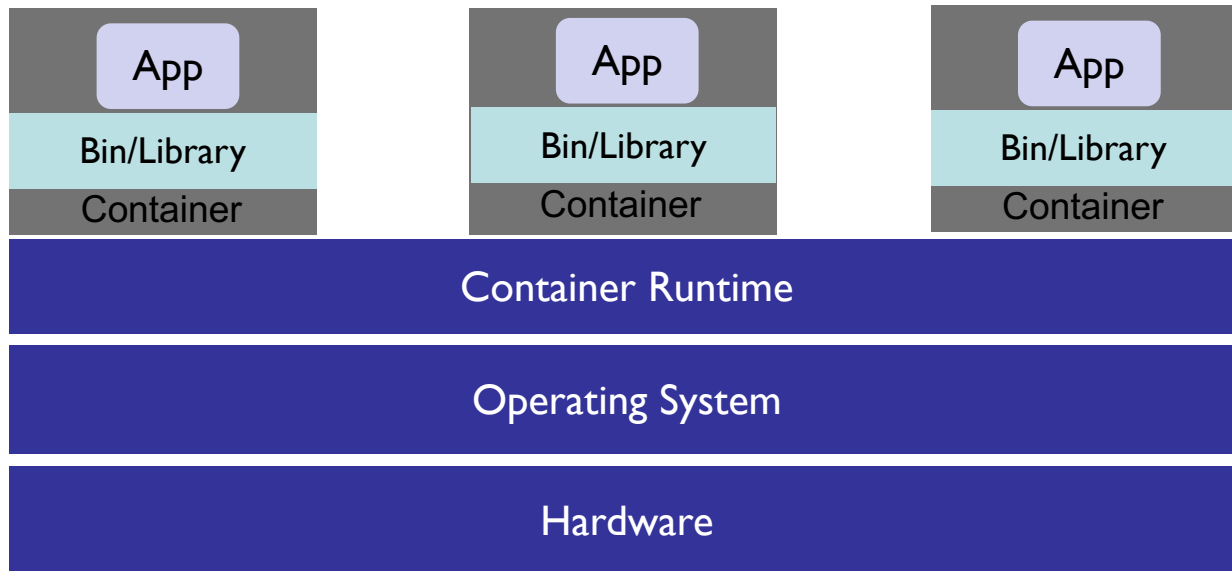
SDC<sup>19</sup>



# Containerized Deployment

September 24, 2019  
Santa Clara, CA

SDC<sup>19</sup>





# Kubernetes Overview

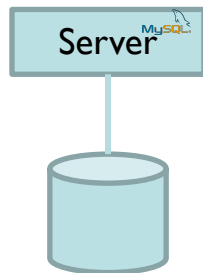
September 24-26, 2019  
Santa Clara, CA

SDC<sup>19</sup>

- Portable, Extensible, Open-Source platform for managing containerized workloads and services
- Agile
  - Modularity and Scalability
- Life Cycle Management
  - Health Checks, Automated Rollouts, Canary Deployment, Load Balancing
- Reduces Costs
  - Containerize apps and consolidate resources

# Simple Stateful App (mysql)

Traditional



Cloud Native

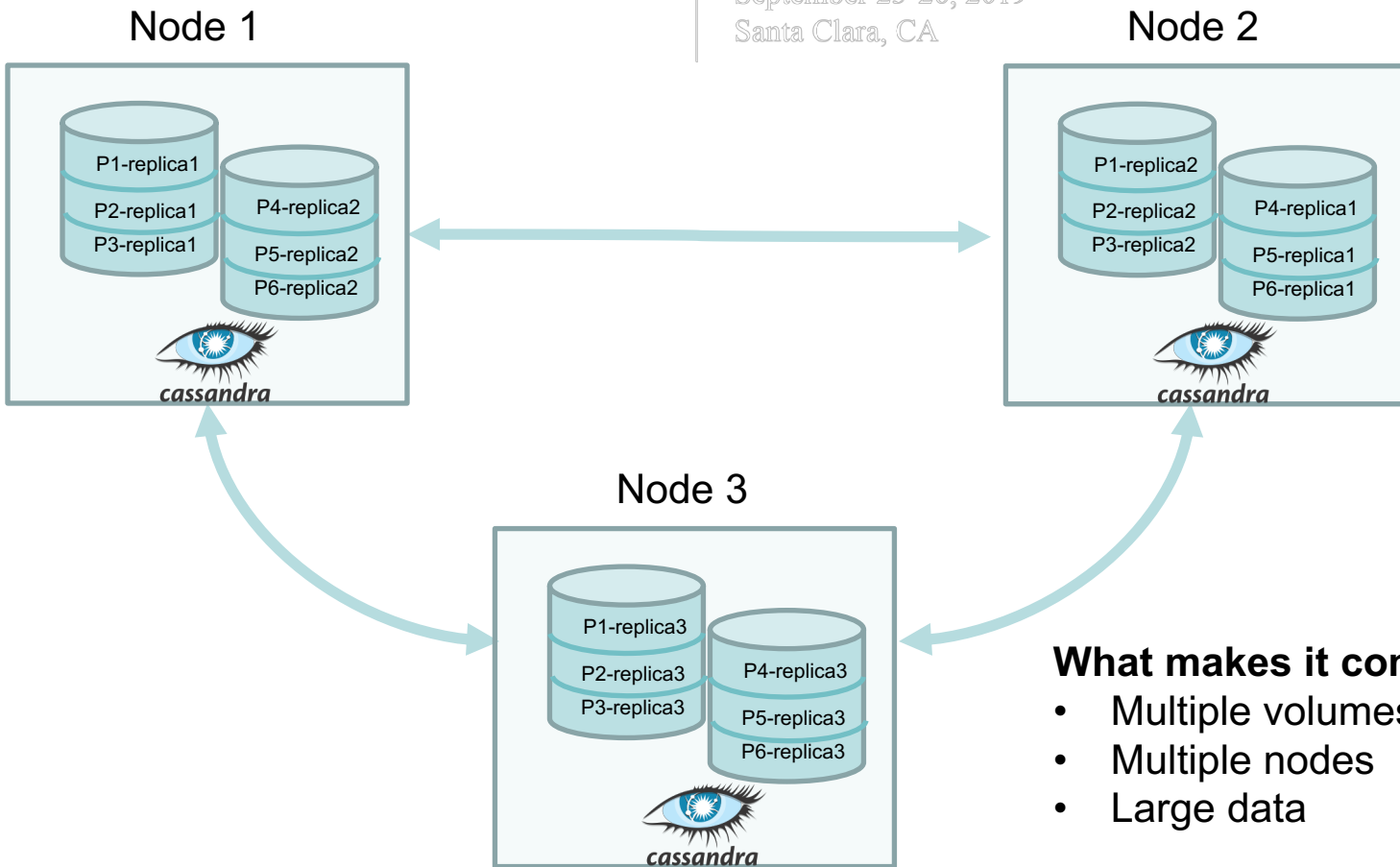




# Distributed Stateful App (Cassandra)

September 23-26, 2019  
Santa Clara, CA

SDC<sup>19</sup>



**What makes it complex?**

- Multiple volumes
- Multiple nodes
- Large data

# Distributed Stateful App (Cassandra)



kubernetes

Node 1



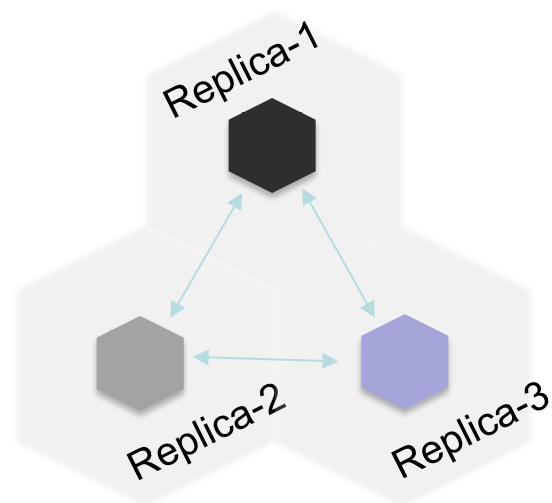
Node 2



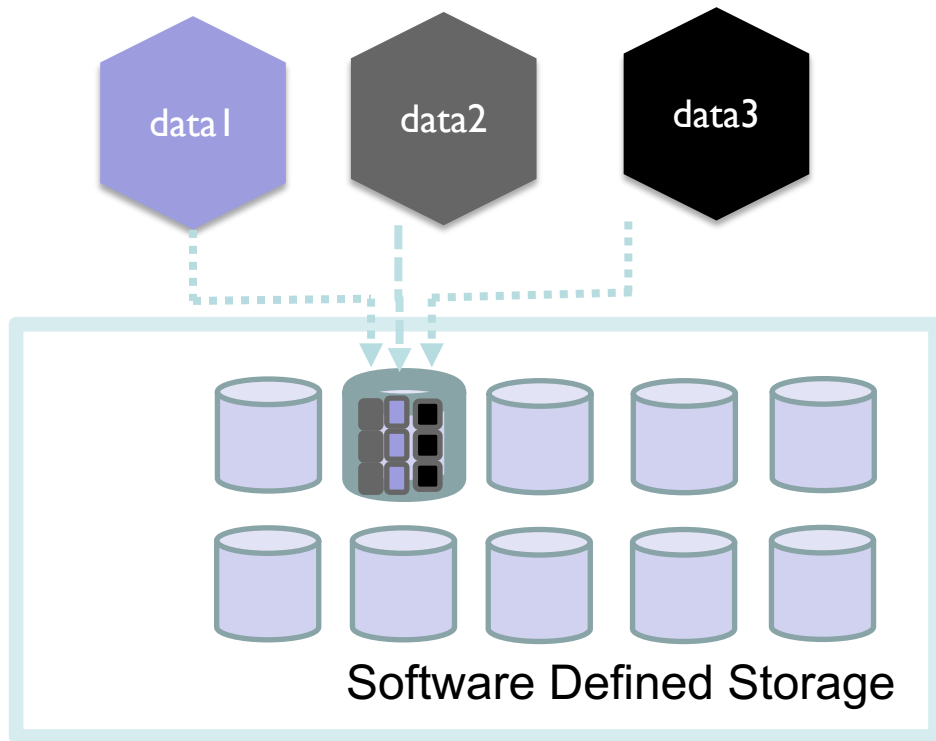
Node 3



# Storage Allocation Challenges



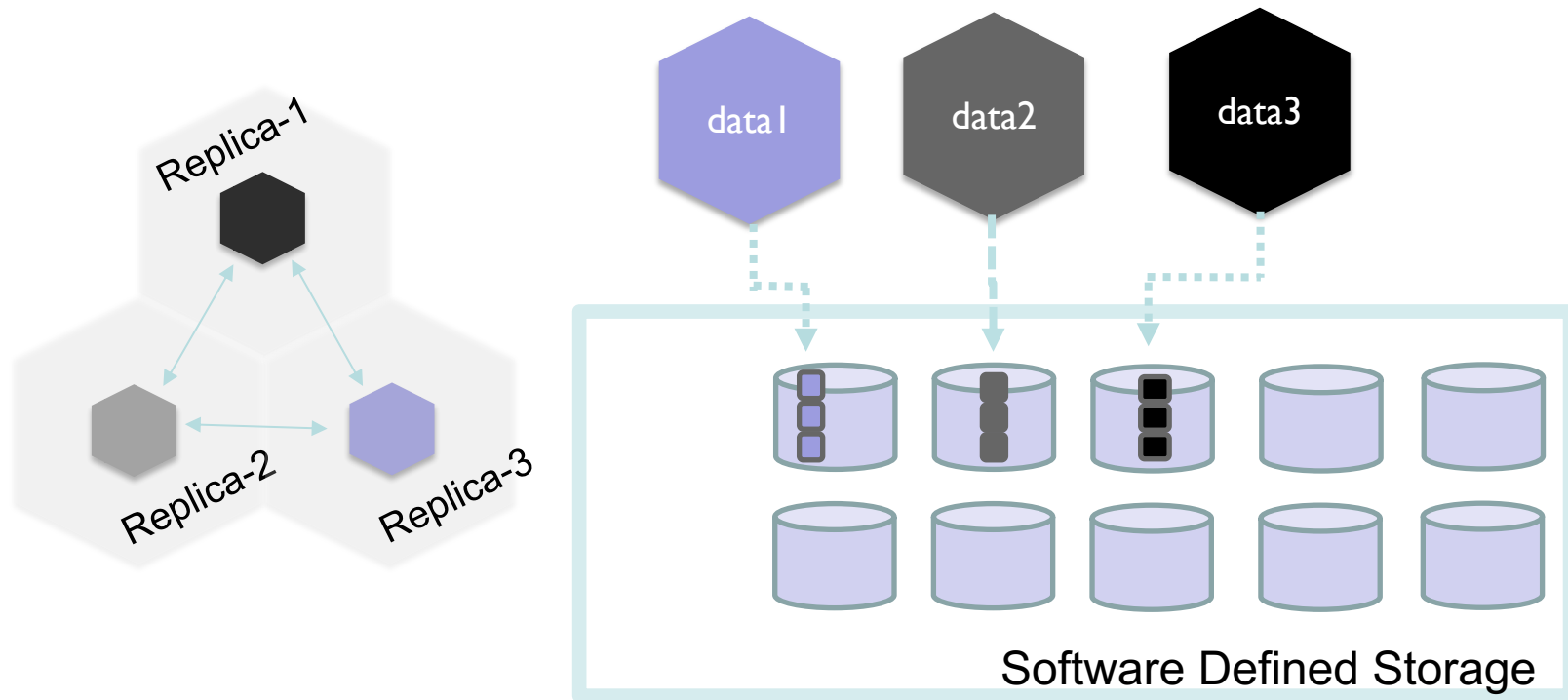
Still resilient to disk failure  
???



# Let us fix it ...

September 24-26, 2019  
Santa Clara, CA

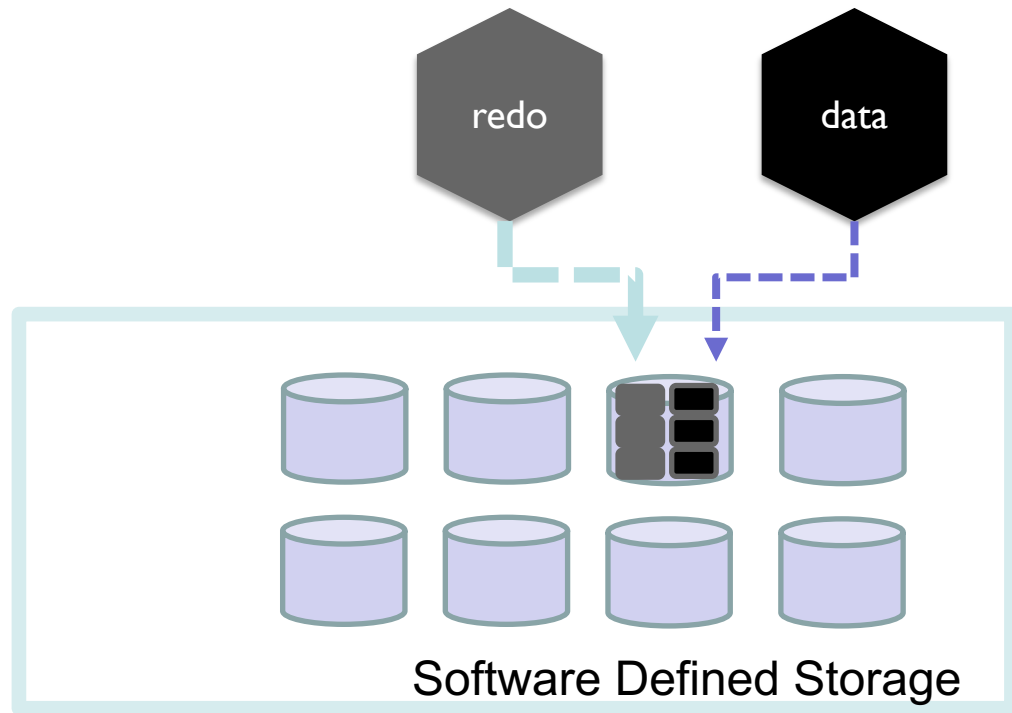
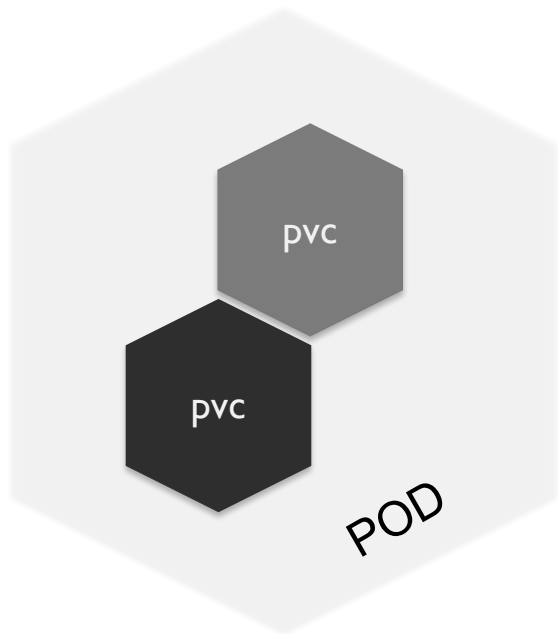
SDC<sup>19</sup>



# Storage Allocation Challenges

September 2019  
Santa Clara, CA

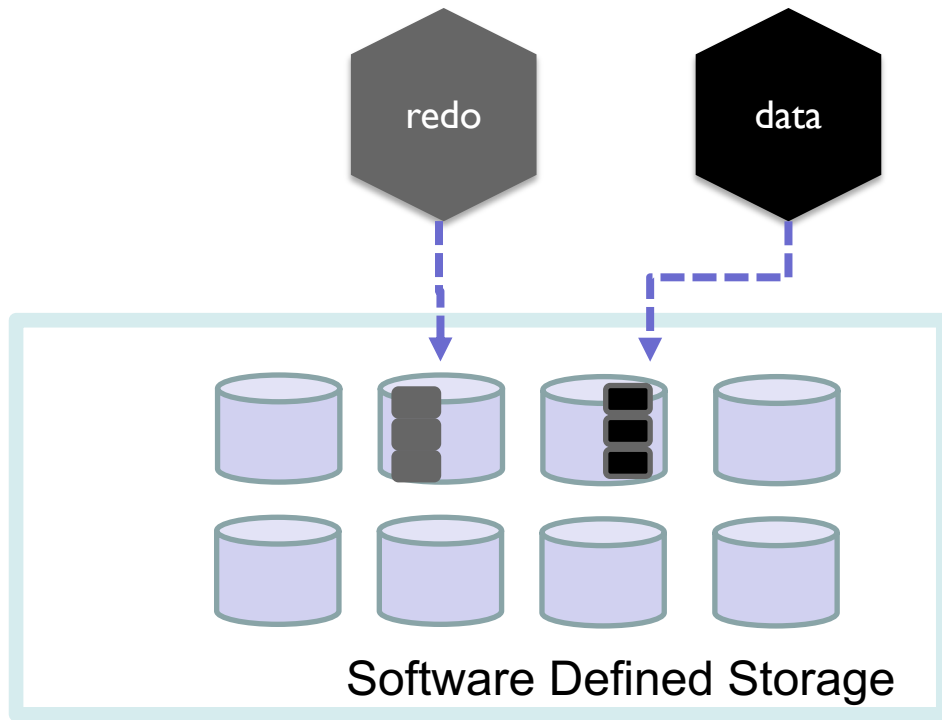
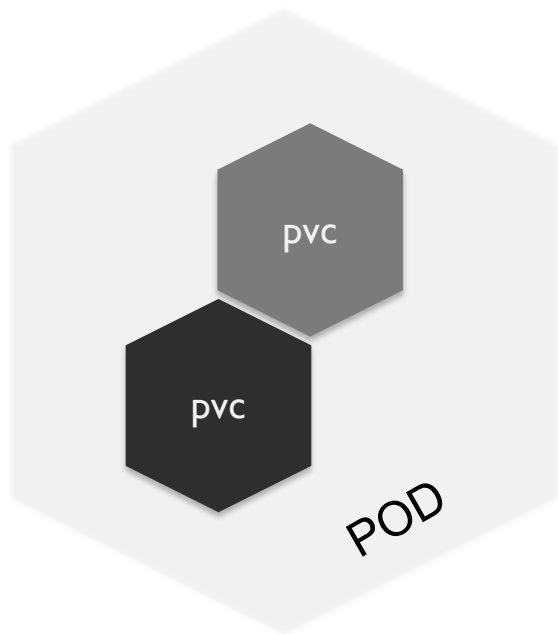
SDC<sup>19</sup>



# Let us fix it...

September 24-26, 2019  
Santa Clara, CA

SDC<sup>19</sup>



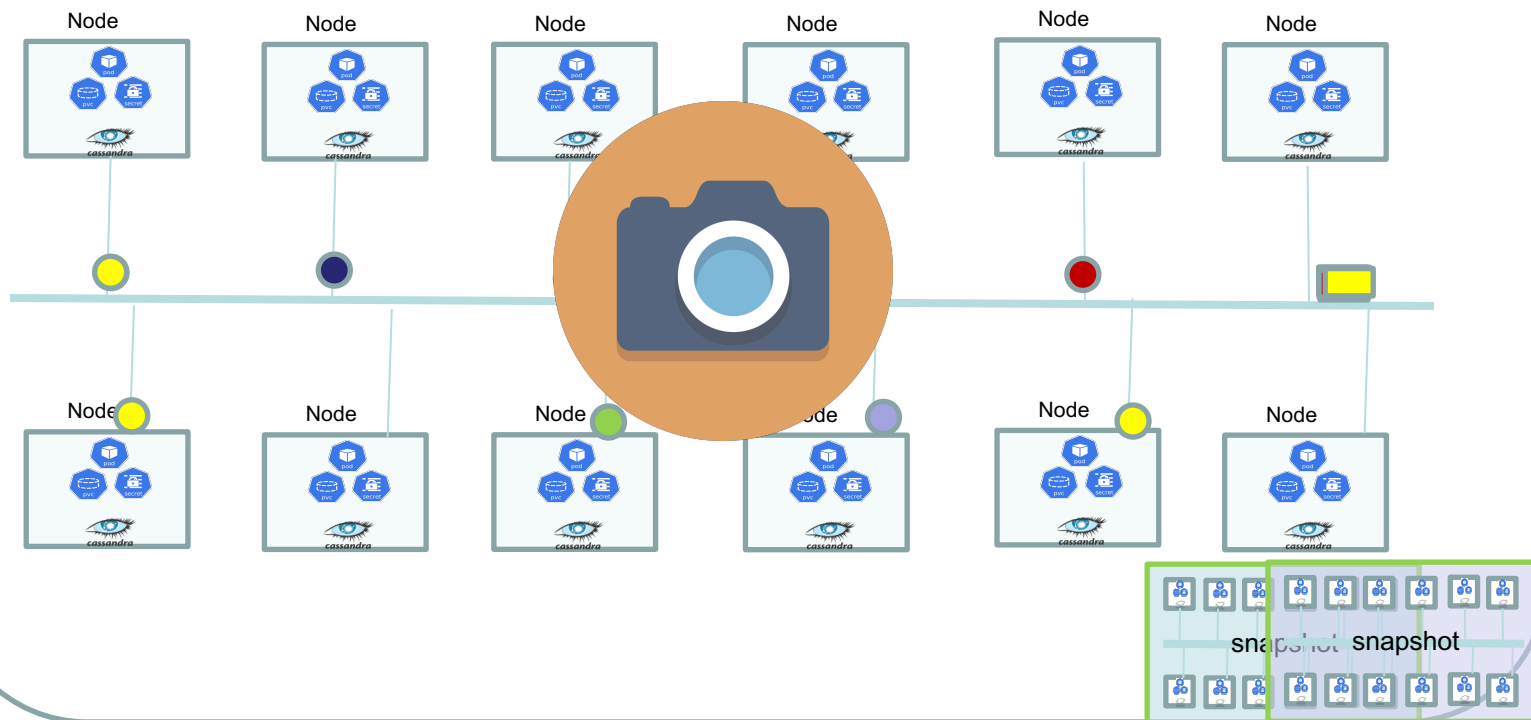
# Challenges of running Stateful Apps on Kubernetes

- Servers/Disks fail and humans make errors → Apps become unavailable, Data is lost
- Data must be protected and secured → Expects users to become storage experts
- Unpredictable performance from sharing resources → Application SLAs not met, unhappy users
- Unable to share data between prod and dev/test → Low productivity and slower innovation
- Data gravity prevents app mobility across environments → Cloud or Infrastructure Lock-in

# Distributed Stateful App (Cassandra)



kubernetes





# Git Operations

Santa Clara, CA

SDC 19

```
$ git commit -m "[descriptive message]"
```

```
$ git reset [commit]
```

```
$ git branch [branch-name]
```

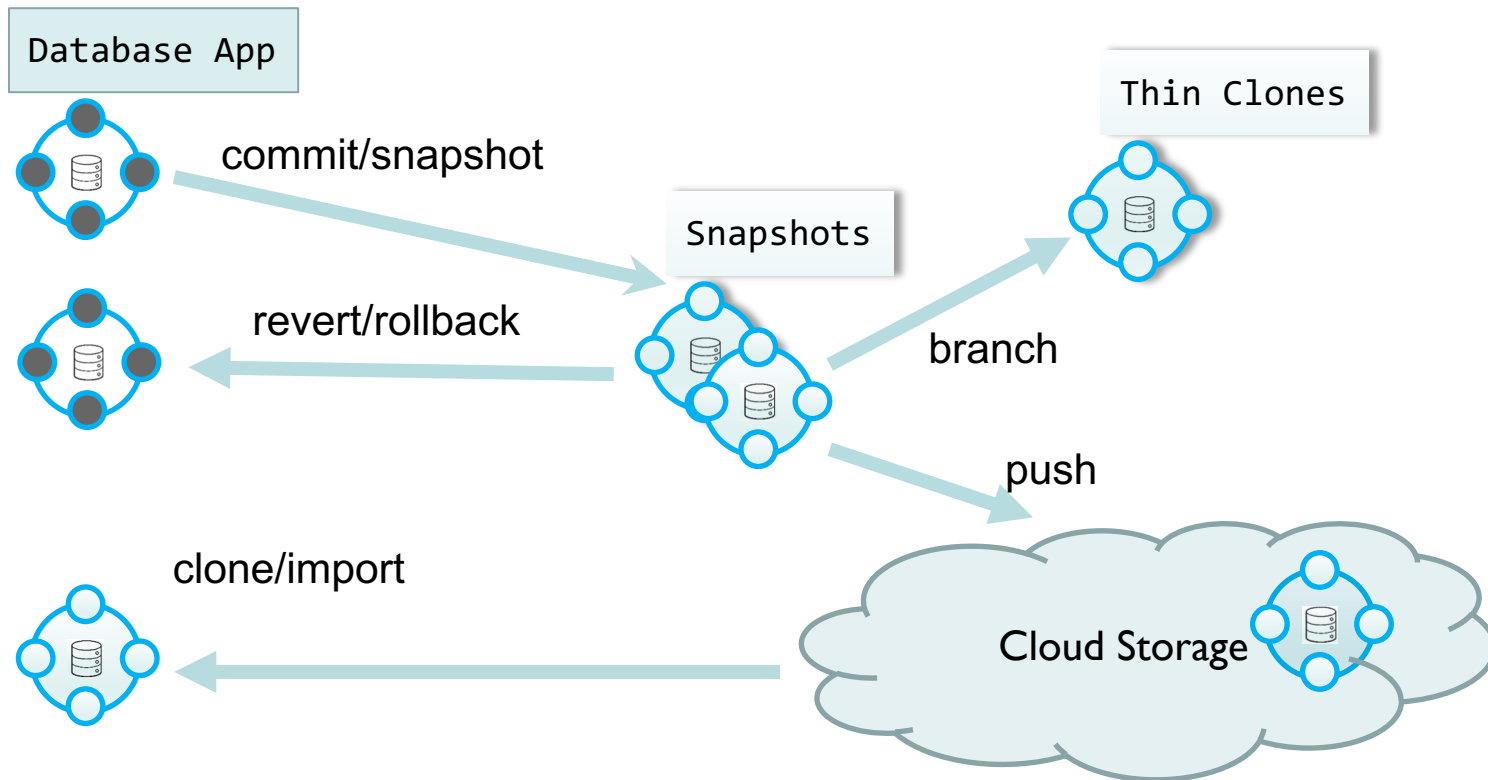
```
$ git push [branch]
```

```
$ git clone [url]
```

# Manage App As If Its In Git...

September 28-29, 2019  
Santa Clara, CA

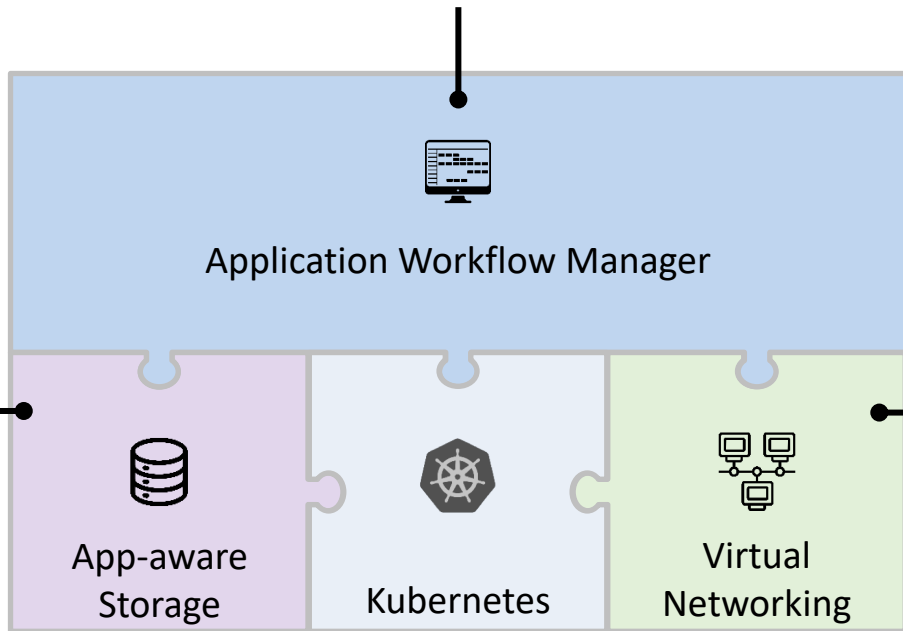
SDC<sup>19</sup>



# Robin Architecture Overview

## 1-click application Deploy, Snapshot, Clone, Scale, Upgrade, Backup

Application workflows configure Kubernetes, Storage & Networking



### Robin's built-in enterprise-grade storage stack

Snapshots, Clones, QoS, Replication, Backup, Data rebalancing, Tiering, Thin-provisioning, Encryption, Compression

### Built-in flexible networking

OVS, Calico, VLAN, Overlay networking, Persistent IPs

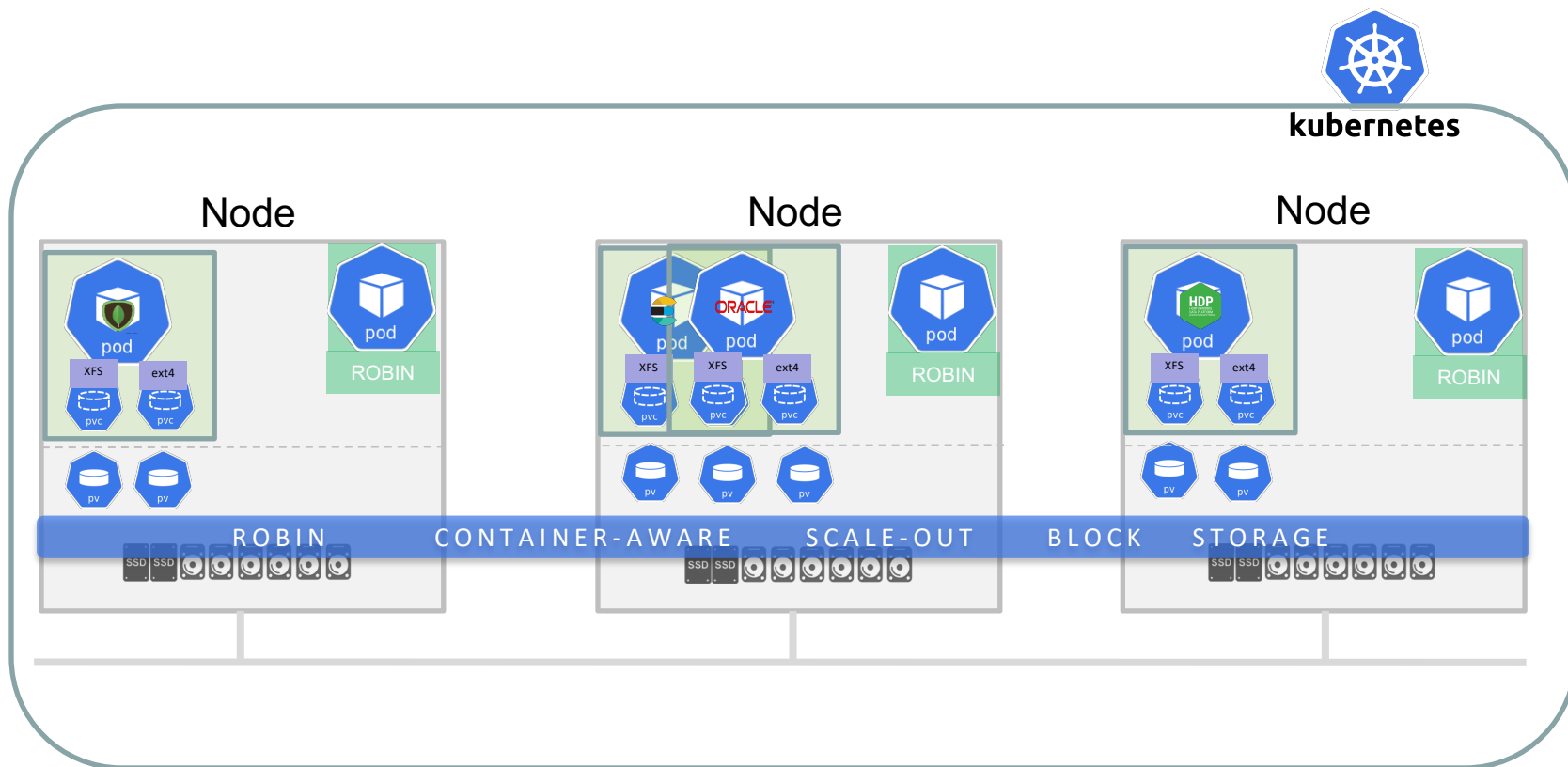


Works any where

# Robin Architecture Overview

Santa Clara, CA

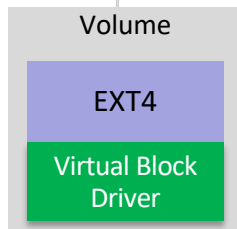
SDC 19



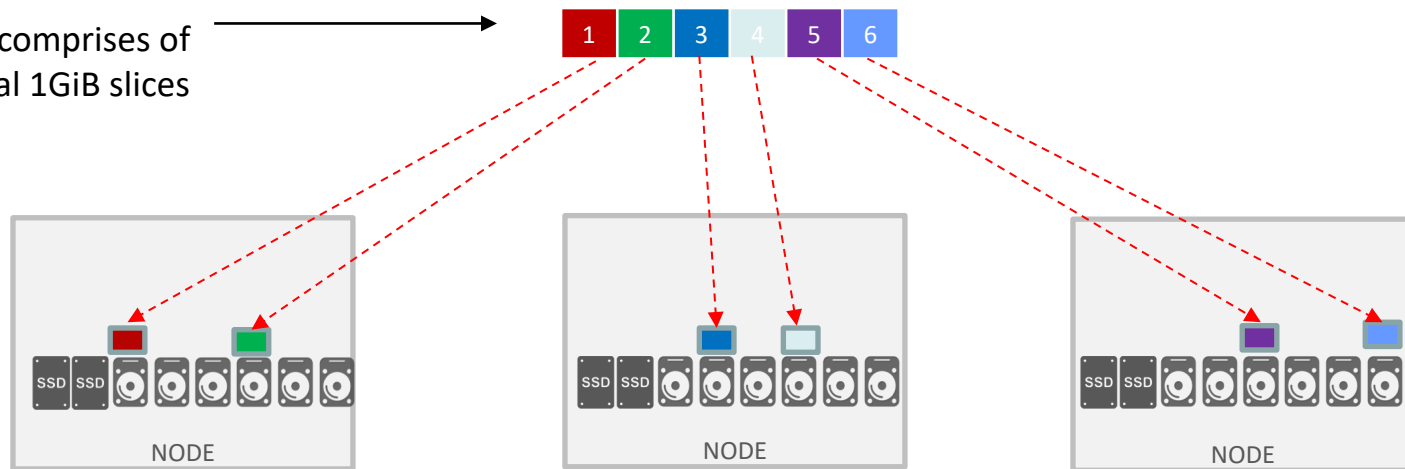
# Logical Volume Allocation

September 28-29, 2019  
Santa Clara, CA

SDC<sup>19</sup>



Volume comprises of  
logical 1GiB slices



# Physical Space provisioning

September 26-27, 2019  
Santa Clara, CA

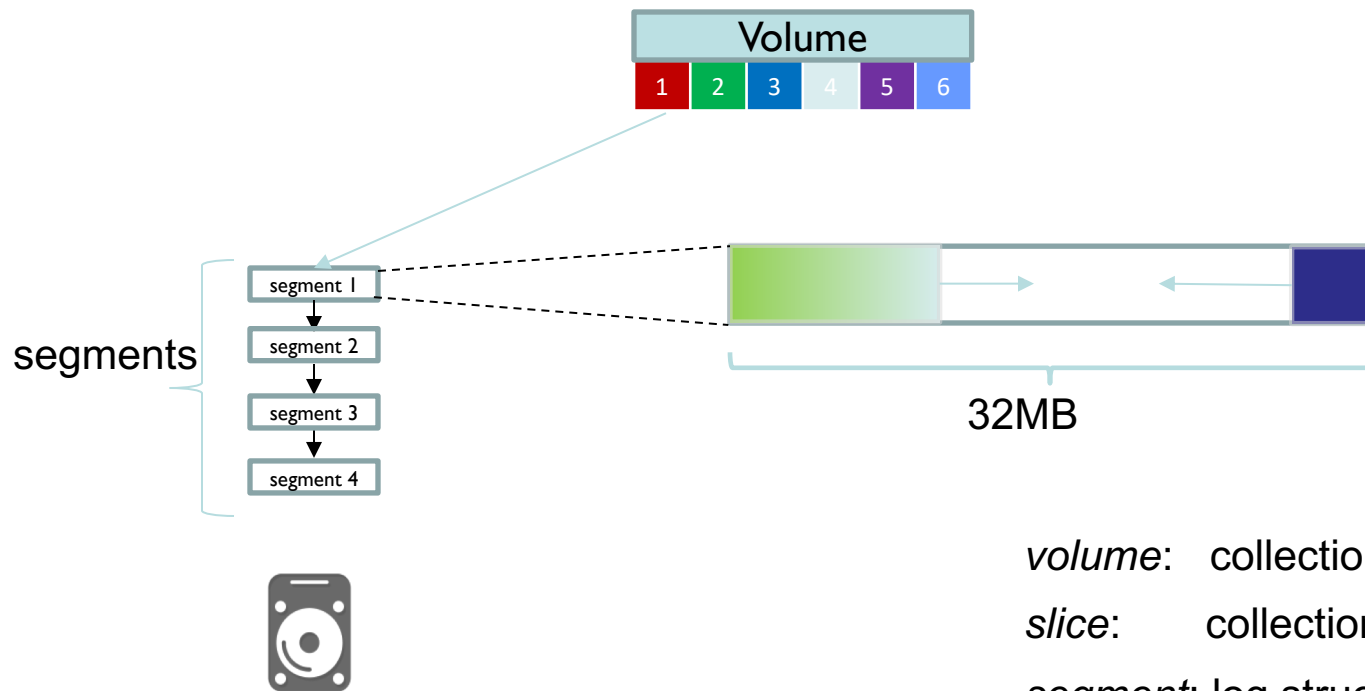
SDC 19

- Multiple volumes of app
- Affinity requirements
- Anti-affinity requirements
- Latency sensitive, Bandwidth intensive
- Node, Rack, Datacenter awareness

# Log Structured Data Layout

September 12, 2019  
Santa Clara, CA

SDC<sup>19</sup>



*volume*: collection of slices

*slice*: collection of segments

*segment*: log structured data blocks

# Robin IO path

September 23-26, 2019

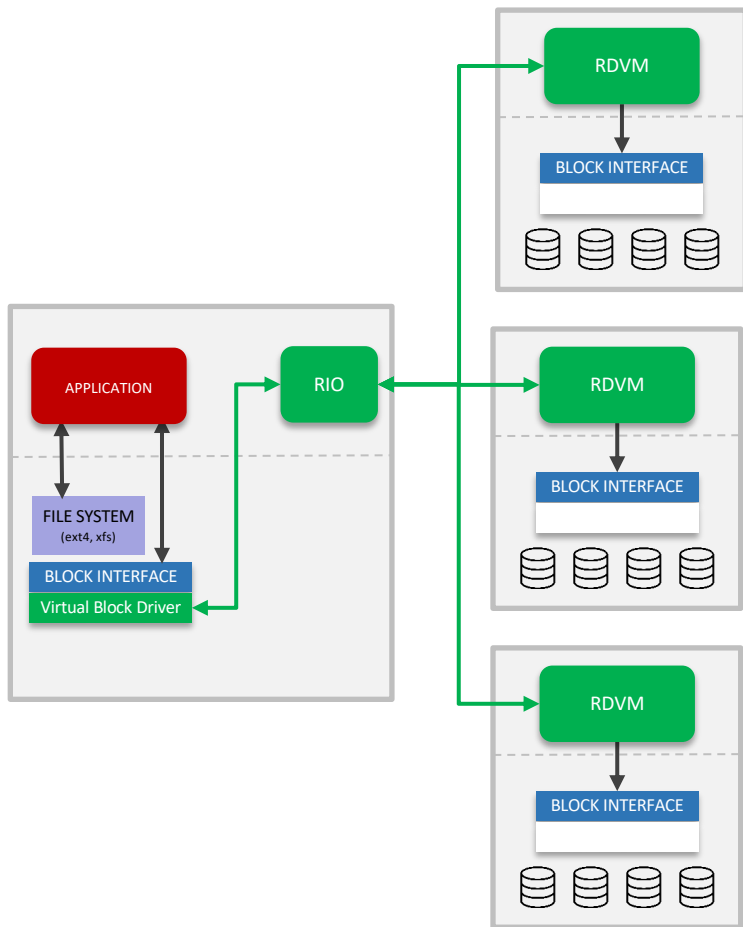
Santa Clara, CA

## COMPONENTS

- › RIO – Robin IO Manager
- › RDVM – Robin Distributed Volume Manager

## FEATURES

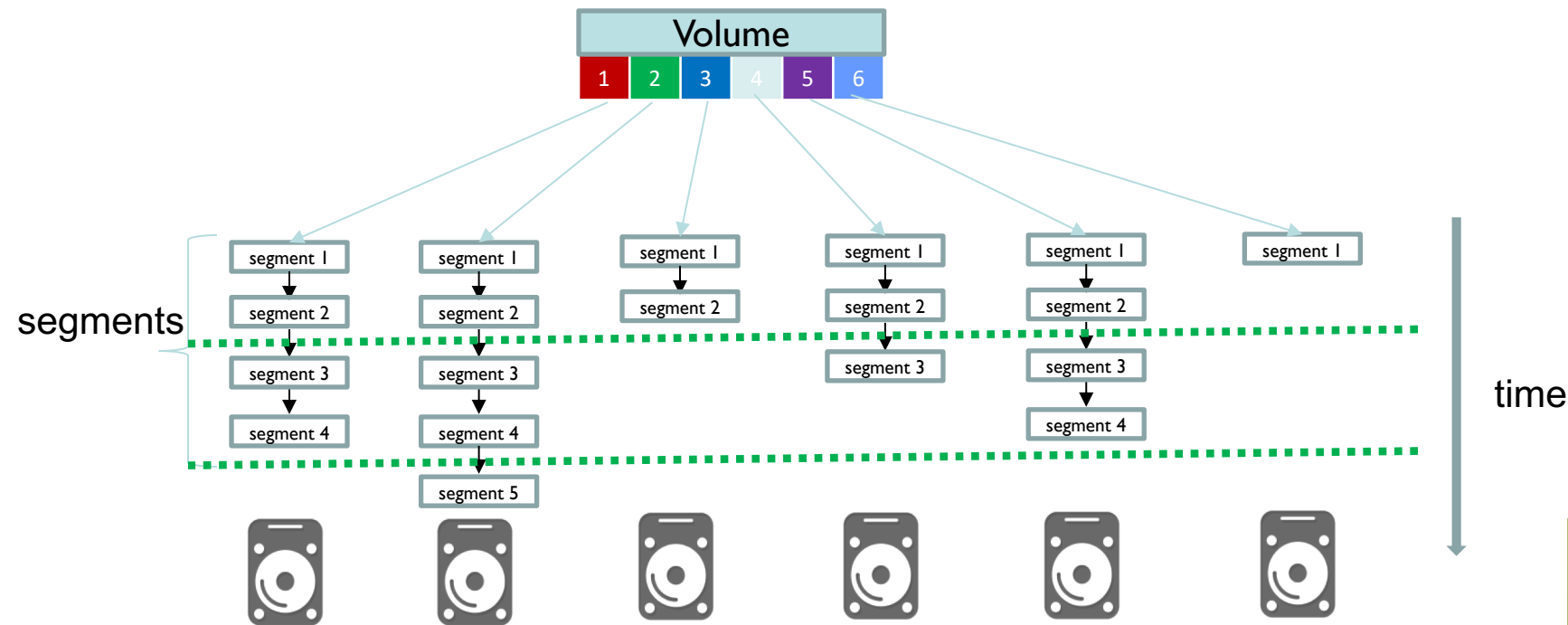
- › High performance IO stack
- › IO tagging helps with end-to-end QoS enforcement
- › Application-aware data placement
- › Supports multi-node volume access (clustered volume manager)
- › Block-level data management (Snapshots, Clones, Compression, Encryption, Tiering etc) enables a common workflow for any and all applications





# Volume Time Travel

September 24-26, 2019  
Santa Clara, CA



# Protect An entire Application, Not Just Storage Volumes

September 23-26, 2019  
Santa Clara, CA

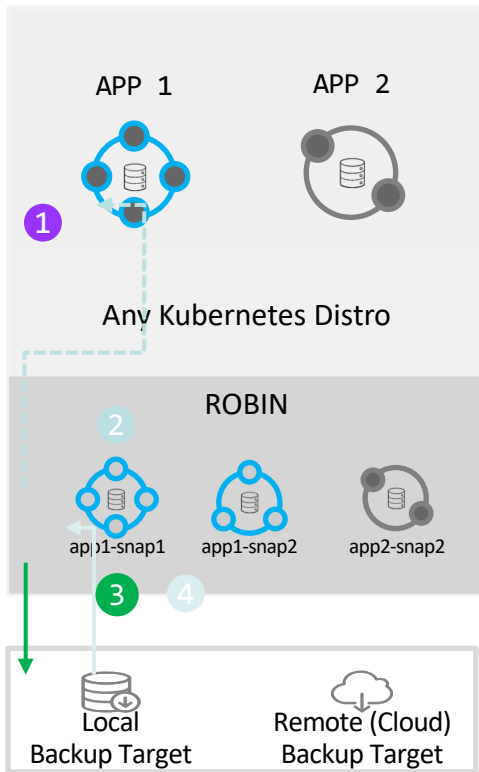
- 1 Maintain periodic checkpoints of your entire app with data

```
$ robin snapshot app1 snap1
```

1	DATA	PersistentVolumeClaims
2	CONFIG	ConfigMap, Secret, Labels, ...
3	METADATA	Pods, StatefulSets, Services, ...

- 2 Rollback entire app+data to healthy state to recover from corruptions or user errors

```
$ robin rollback snap1 app1
```



- 3 Backup entire app+data as into external backup targets

```
$ robin push snap1 target
```

- 4 Restore entire app+data to healthy state from catastrophic hardware and datacenter failures

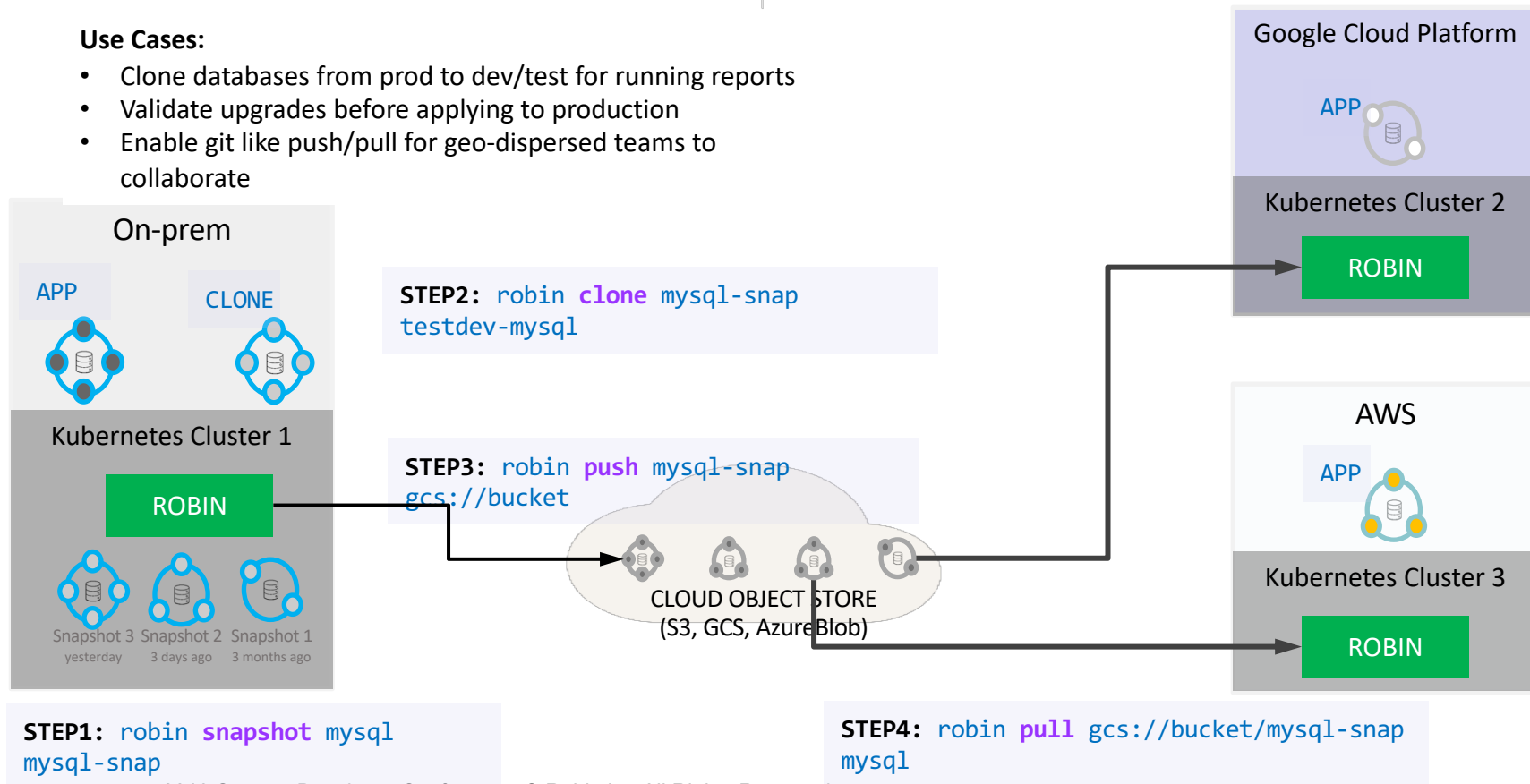
```
$ robin pull target snap1
```

- › ROBIN Backups are fully self-contained
- › Entire app resources can be restored in the same or different data center or cloud even if the source is completely destroyed

# Make DevOps Productive With Stateful Apps

## Use Cases:

- Clone databases from prod to dev/test for running reports
- Validate upgrades before applying to production
- Enable git like push/pull for geo-dispersed teams to collaborate



# App Snapshot

September 26-27, 2019  
Santa Clara, CA

SDC<sup>19</sup>

```
$ git commit -m "[descriptive  
message]"
```

```
$ robin snapshot <snapname>  
<appname> -m "[descriptive message]"
```

- Application Consistent
- Quiesce/Unquiesce volumes
- Complete app state snapshotted
  - Kubernetes Objects (Metadata)
  - Volumes (Data)

# App Rollback/Reset

Storage Developer 2019  
Santa Clara, CA

```
$ $ git reset [commit]
```

```
$ robin app rollback <snapname> <appname>
```

- App config and data restored
- Handle scale-in/scale-out scenarios
- Revert on corruption, user errors, upgrade failures

# Thin Clone

September 16, 2019  
Santa Clara, CA

```
$ git branch [branch-name]
```

```
$ robin app create <appname> <snapname>
```

- Instant app deployment from snapshot
- No data copied
- Customize cloned app config
- Blue/Green deployment
- Test upgrades
- Run analytics

# Push to Cloud

September 24, 2019  
Santa Clara, CA

SDC<sup>19</sup>

```
$ git push [branch]
```

```
$ robin app push <snapname> <reponame>
```

- Uses app snapshot
- Self sufficient copy of app
- Incremental transfer to repos
  - AWS S3, Google GCS, Azure Block blob, NFS share
- Encrypted and Compressed

# Pull from Cloud

Santa Clara, CA

```
$ git clone <url>
```

```
$ robin app pull/create <snapname> <reponame>
```

- Repair app from repo copy
- Create new app from repo copy
- App mobility
  - Run app in the cloud
- Protect from site failures
  - Use cloud copy as backup





**Demo**



# Q & A



# Thank You

Please visit ***robin.io*** for more information  
Download the software at ***get.robin.io*** and  
try it out