



**SDC** 18

September 24-27, 2018  
Santa Clara, CA

[www.storagedeveloper.org](http://www.storagedeveloper.org)

# Bulletproofing Stateful Applications on Kubernetes

**Dinesh Israni, @disrani**  
**(Principal Software Engineer)**  
**Portworx Inc.**

# Agenda

- ❑ About Portworx
- ❑ Motivation for Stork
- ❑ Scheduling stateful services efficiently
- ❑ Storage Health Monitoring
- ❑ Disaster Recovery and Migration
- ❑ Demo
- ❑ Q&A

# About Portworx

- ❑ First production-ready software defined storage solution designed for microservices
- ❑ Container granular virtual storage
- ❑ Run your workloads local with your storage
- ❑ Snapshots and CloudSnaps for backup and DR
- ❑ Bring-Your-Own-Key Encryption
- ❑ Automate provisioning and control repeatedly on-prem and in any cloud
- ❑ Runs as a container itself on your agents!

# Storage Orchestration Runtime for Kubernetes (STORK)

- ❑ Started to help run stateful applications more efficiently on Kubernetes
- ❑ Manage lifecycle of stateful applications
- ❑ Plugin model, can be extended to work with any storage driver
- ❑ Open source:  
<https://github.com/libopenstorage/stork>

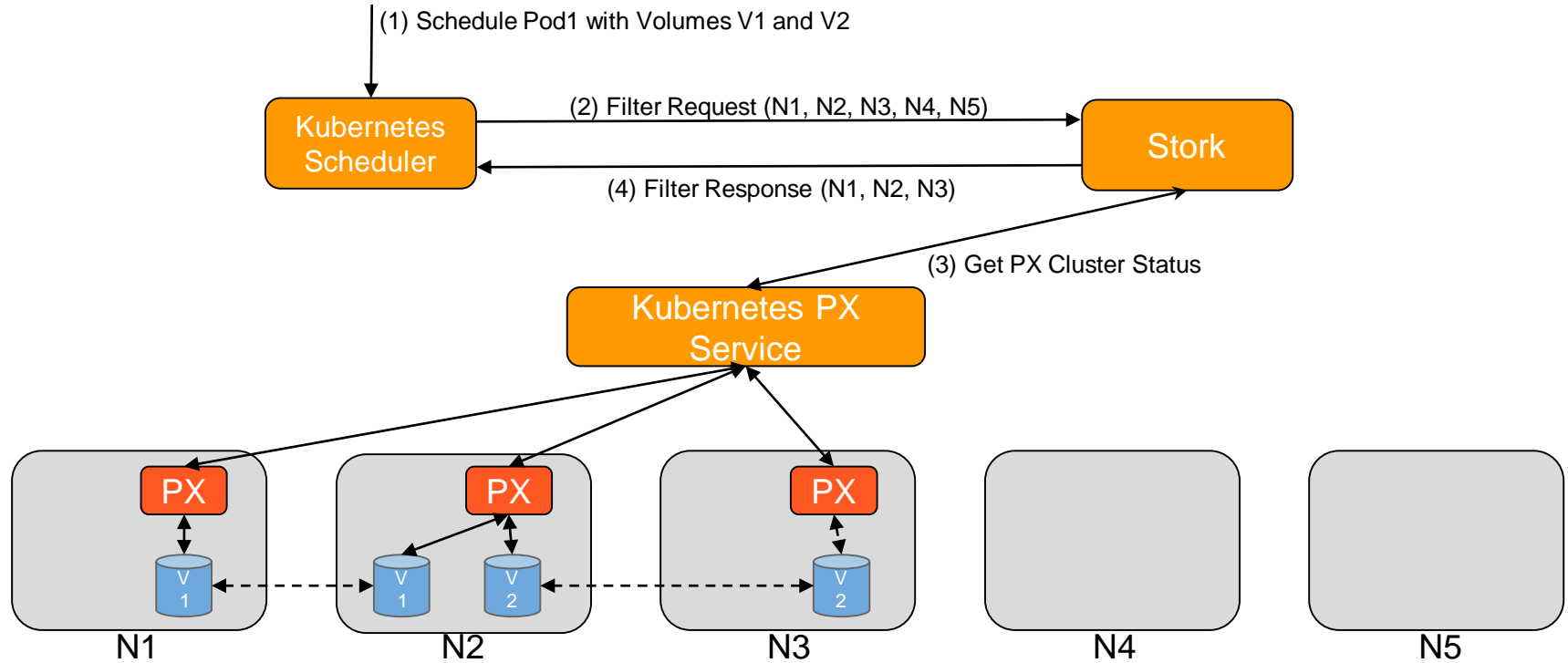
# Scheduling stateful services efficiently

- ❑ How do you start services close to where data is located?
- ❑ Wide use of labels and affinity rules
  - ❑ Doesn't scale
  - ❑ Doesn't work with stateful sets
  - ❑ Error prone

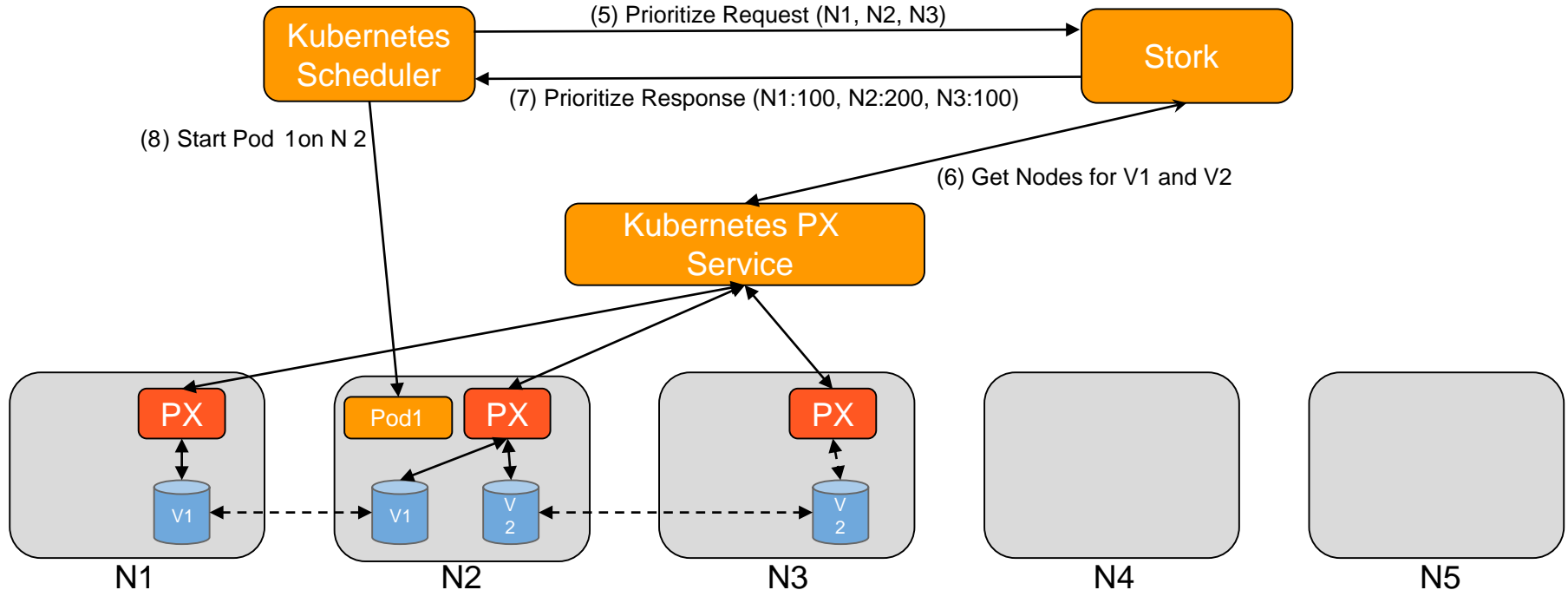
# Scheduling stateful services efficiently

- ❑ Solution: Use scheduler extenders
- ❑ Kubernetes allows extending the default scheduler
- ❑ Can be used to
  - ❑ “filter” out nodes where storage isn’t available
  - ❑ “prioritize” nodes where data is local
- ❑ Simple to use
  - ❑ Either configure default scheduler with extender
  - ❑ Or, start new instance of scheduler and use in your apps

# Scheduling stateful services efficiently



# Scheduling stateful services efficiently





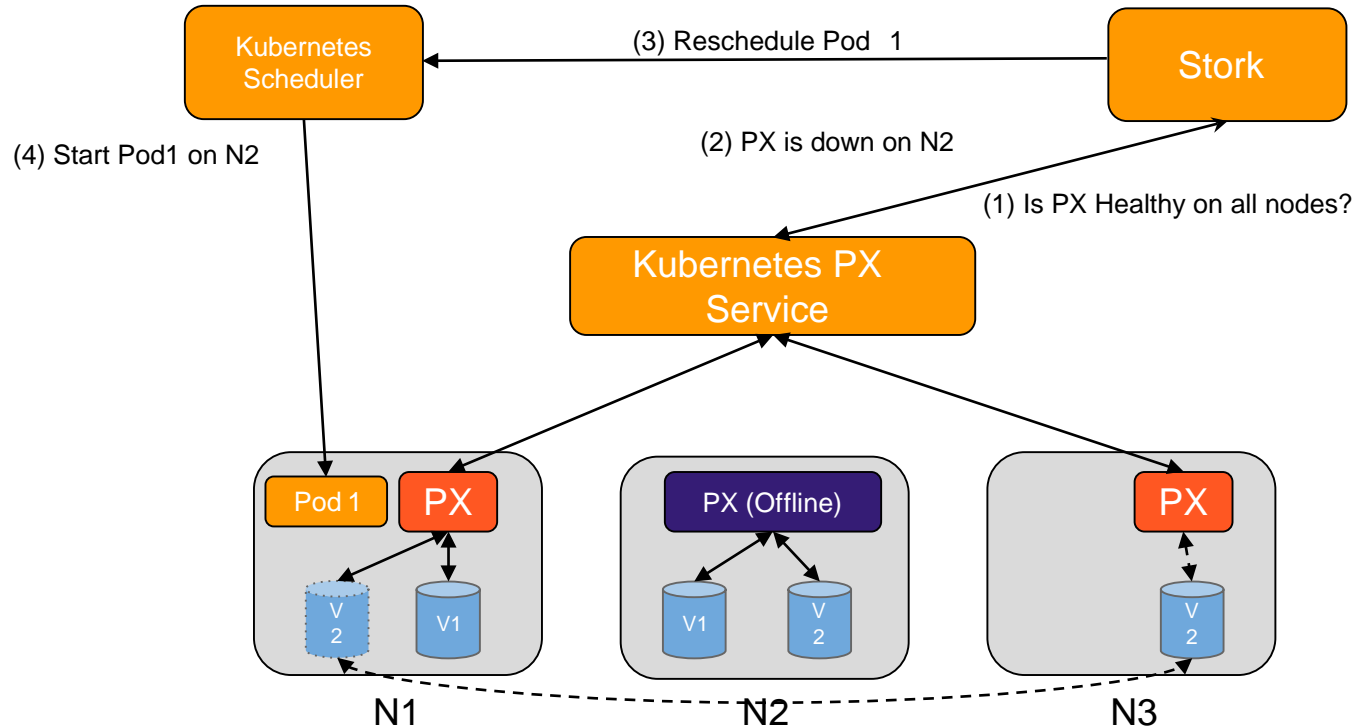
# Storage Health Monitoring

- ❑ All good when everything is online
- ❑ Dealing with failures is difficult, especially with state
- ❑ What if storage driver goes offline on a node?
  - ❑ Storage degradation
  - ❑ Software bugs/crashes
- ❑ What happens to pods on that node?
  - ❑ Kubelet is still running
- ❑ Usually requires manual intervention

# Storage Health Monitoring

- ❑ Monitors the health of storage driver on all nodes
- ❑ Storage driver offline?
  - ❑ Reschedule pods using storage driver
- ❑ Rescheduled on another node with volume replica
  - ❑ Continue with local disk performance
- ❑ Without this, pods will get stuck in Pending, or not able to access storage

# Storage Health Monitoring



# Disaster Recovery

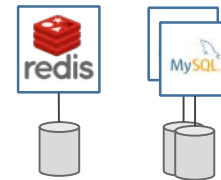
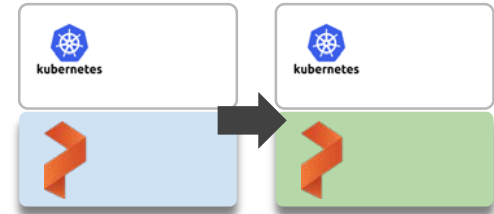
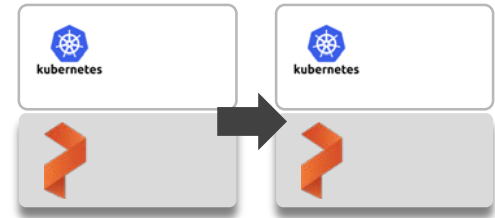
- ❑ Need a way to manage lifecycle of storage natively in Kubernetes
- ❑ No in-tree mechanism to take snapshots of PVCs in Kubernetes yet
- ❑ Adds support for Snapshots (based on Kubernetes Incubator project)
  - ❑ Local Snapshots
  - ❑ Cloudsnaps to objectstore (Any S3 compliant, Azure and Google)
- ❑ Application consistent snapshots
  - ❑ Quiesce or flush applications before taking snapshots using pre/post snapshots hooks
  - ❑ Also works over a group of PVCs / Volumes for distributed apps

# Cluster Migration

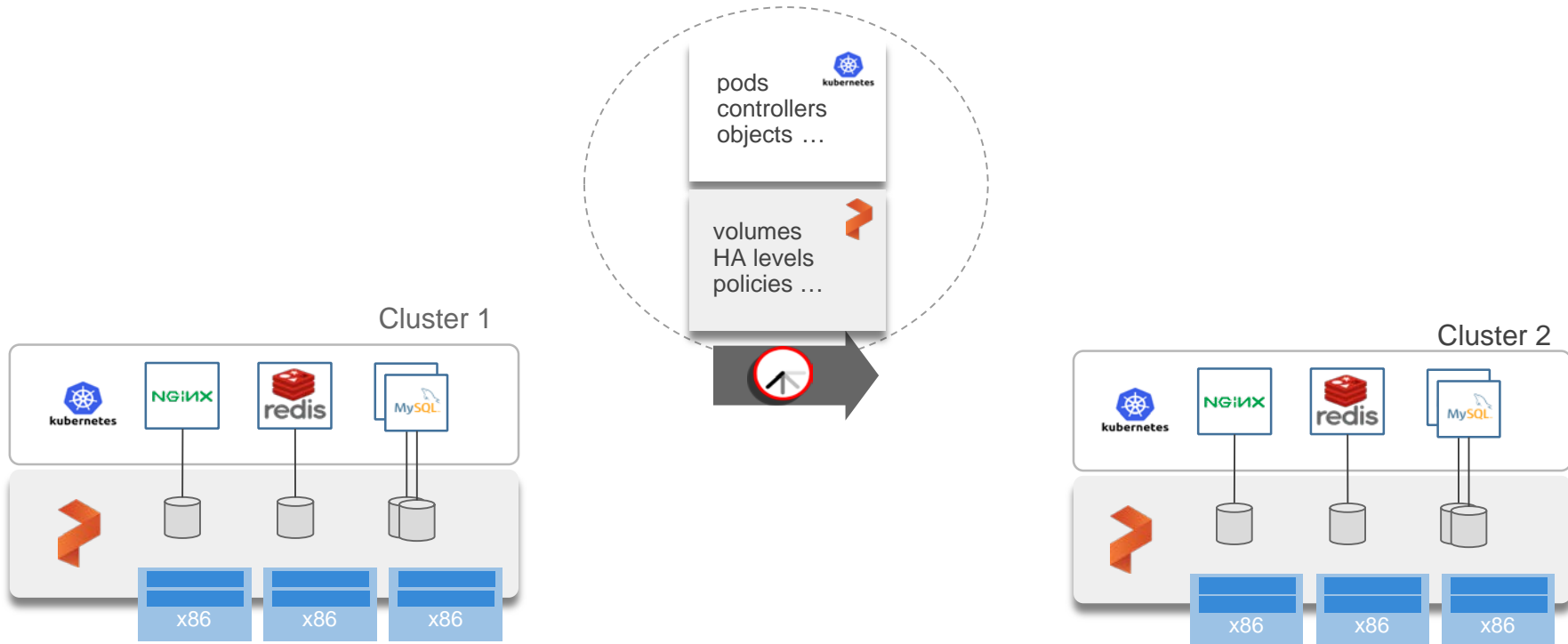
1 **Augmentation** : we are out capacity and want to move *select* applications and data to a second cluster.

2 **Blue-Green**: a new version of Portworx is released, and we want to qualify with *all* applications and data.  
(Also works for a new version of Kubernetes.)

3 **Dev/Test** : a bug in production needs to be reproduced off-cluster. We want to move *just* that app and its data.



# Cluster Migration



# Cluster Migration

- ❑ Pair two clusters
  - ❑ Creates pairing with storage nodes
  - ❑ Pairs kubernetes clusters
- ❑ Apply spec and wait for it to be “Ready”

```
apiVersion: stork.libopenstorage.org/v1alpha1
kind: ClusterPair
metadata:
  name: remotecluster
spec:
  options:
    ip: 192.168.56.75
    token: <token_from_remote_storage_cluster>
    port: "9001"
  config:
    <Remote Kubernetes Cluster config>
```

# Cluster Migration

- ❑ Migrate between clusters
  - ❑ Specify which namespaces to migrate
  - ❑ Migrate applications?
  - ❑ Start applications on remote cluster?
- ❑ First migrates all volumes from storage driver
- ❑ Then migrates resources
  - ❑ Deployments
  - ❑ PVCs, PVs
  - ❑ Secrets
  - ❑ ...

```
apiVersion: stork.libopenstorage.org/v1alpha1
kind: Migration
metadata:
  name: mysqlMigration
spec:
  clusterPair: remoteclassic
  includeResources: true
  startApplications: true
  namespaces:
  - mysql
```



# Demo Time!

# Questions?

# Learn More

- ❑ Source code: [www.github.com/libopenstorage/stork/](https://www.github.com/libopenstorage/stork/)
- ❑ Blog: <https://portworx.com/stork-storage-orchestration-kubernetes>
- ❑ Request a demo and free trial of PX-Enterprise at [info@portworx.com](mailto:info@portworx.com)

# Thank you!