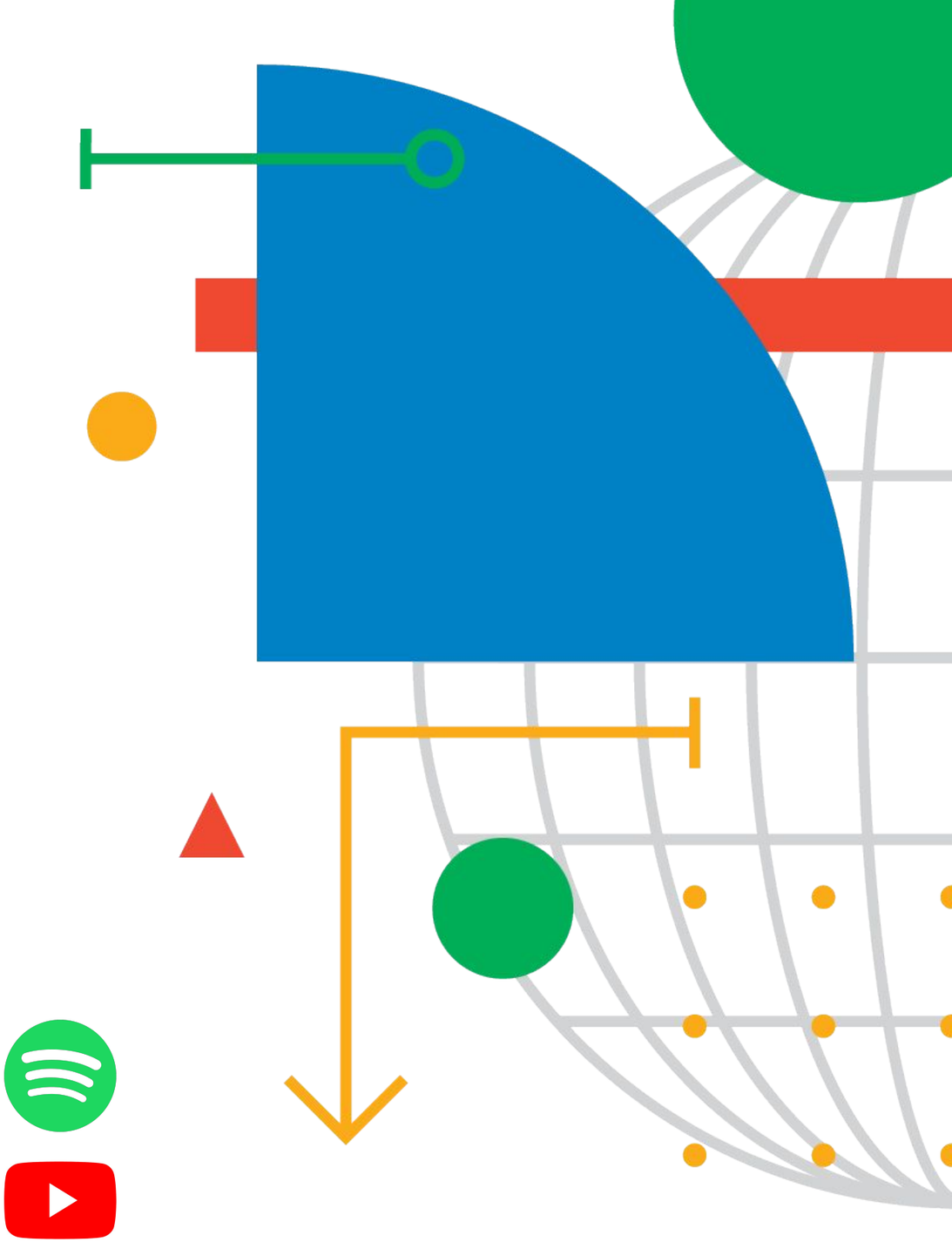
Kubernetes Statefulset

Imre Nagi *@imrenagi*

Google Developer Expert Cloud
Podcast Ngobrolin Startup & Teknologi



Deploying Stateless App is fairly easy..

Stateful App / Database ????

StatefulSets

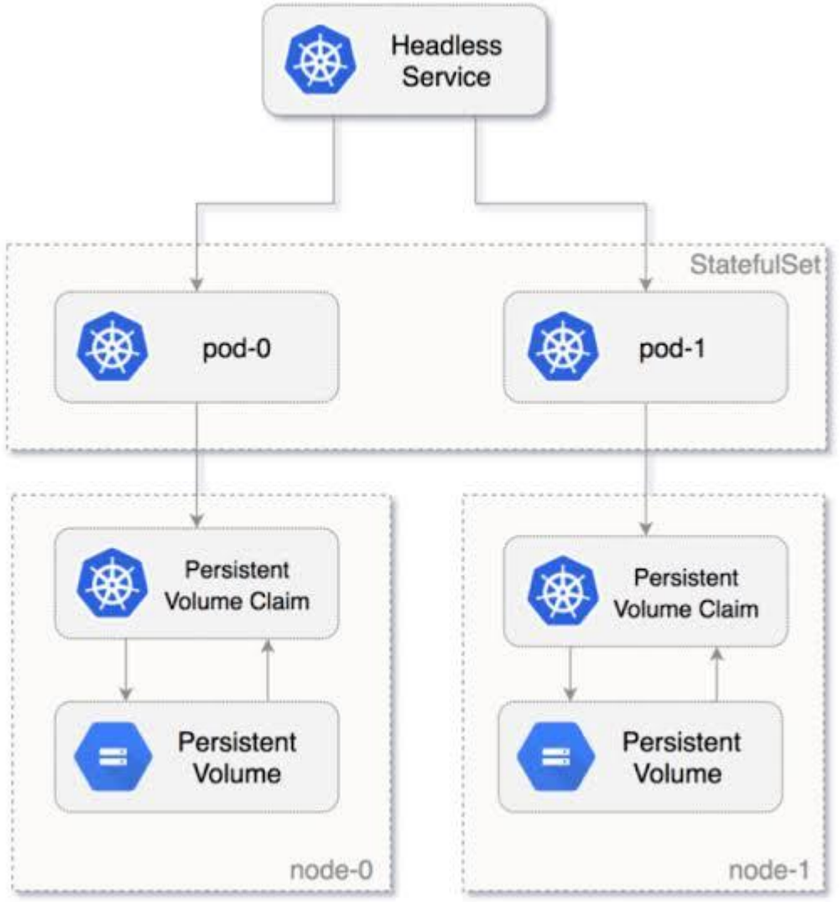
- StatefulSet is the workload API object used to manage stateful applications.
- Manages the deployment and scaling of a set of Pods *and provides guarantees about the ordering and uniqueness* of these Pods.

StatefulSets are valuable for applications that require one or more of the following.

1. Stable, unique network identifiers.
2. Stable, persistent storage.
3. Ordered, graceful deployment and scaling.
4. Ordered, automated rolling updates.

	Deployment	StatefulSet
Application Type	Stateless	Stateful
Pod Management	Manage Replicaset	Unique & Ordered Pod
Volume Sharing	Might Share Its PVC	Each Pod Has its own PVC
Failure Mitigation	Rollback	Rollback not available
Pod upgrade	Rolling Update	Rolling Update, On Delete

Headless Service

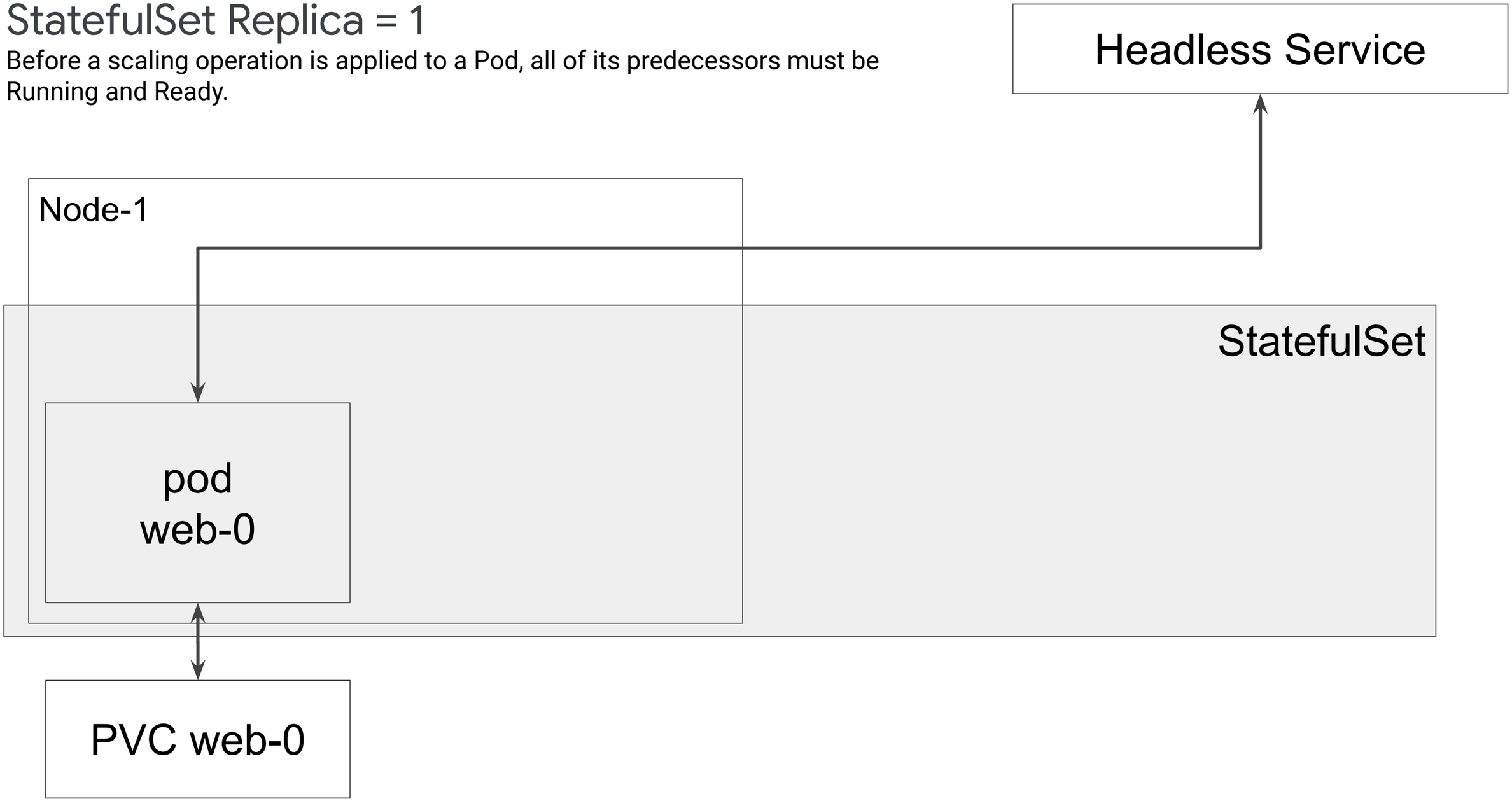


StatefulSets currently require a **Headless Service** to be responsible for the network identity of the Pods. You are responsible for creating this Service.

ClusterIP : None

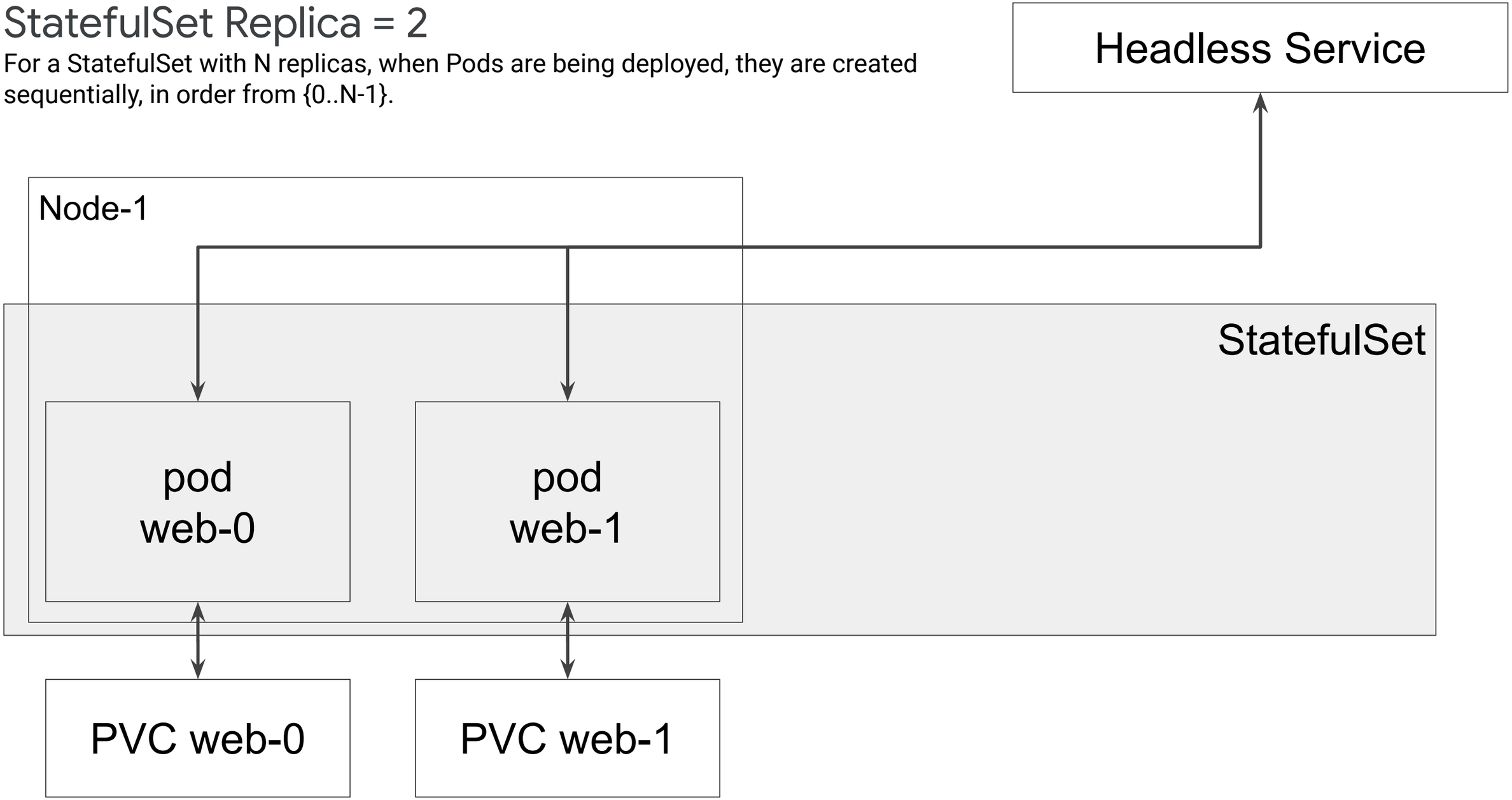
StatefulSet Replica = 1

Before a scaling operation is applied to a Pod, all of its predecessors must be Running and Ready.

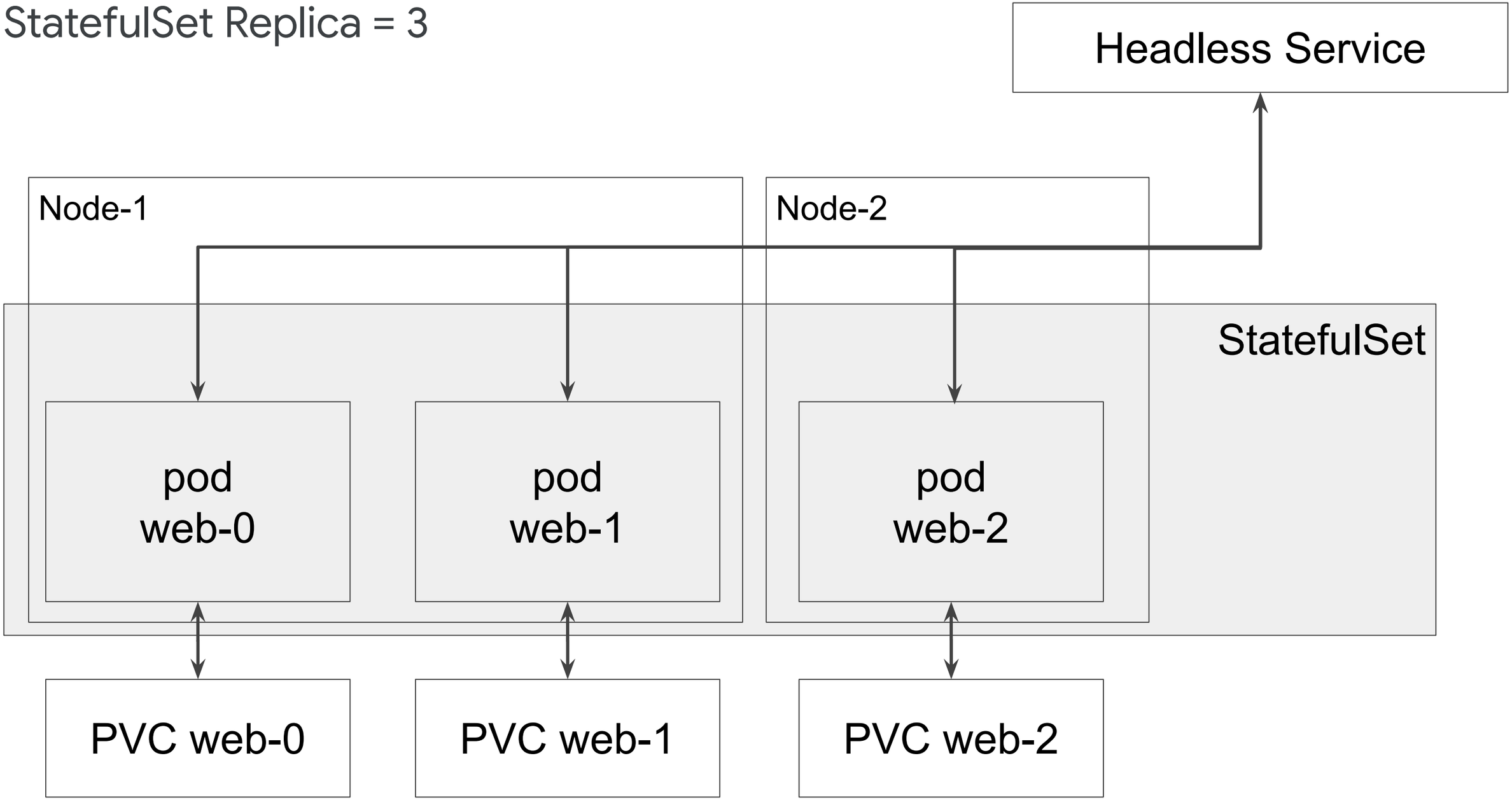


StatefulSet Replica = 2

For a StatefulSet with N replicas, when Pods are being deployed, they are created sequentially, in order from {0..N-1}.

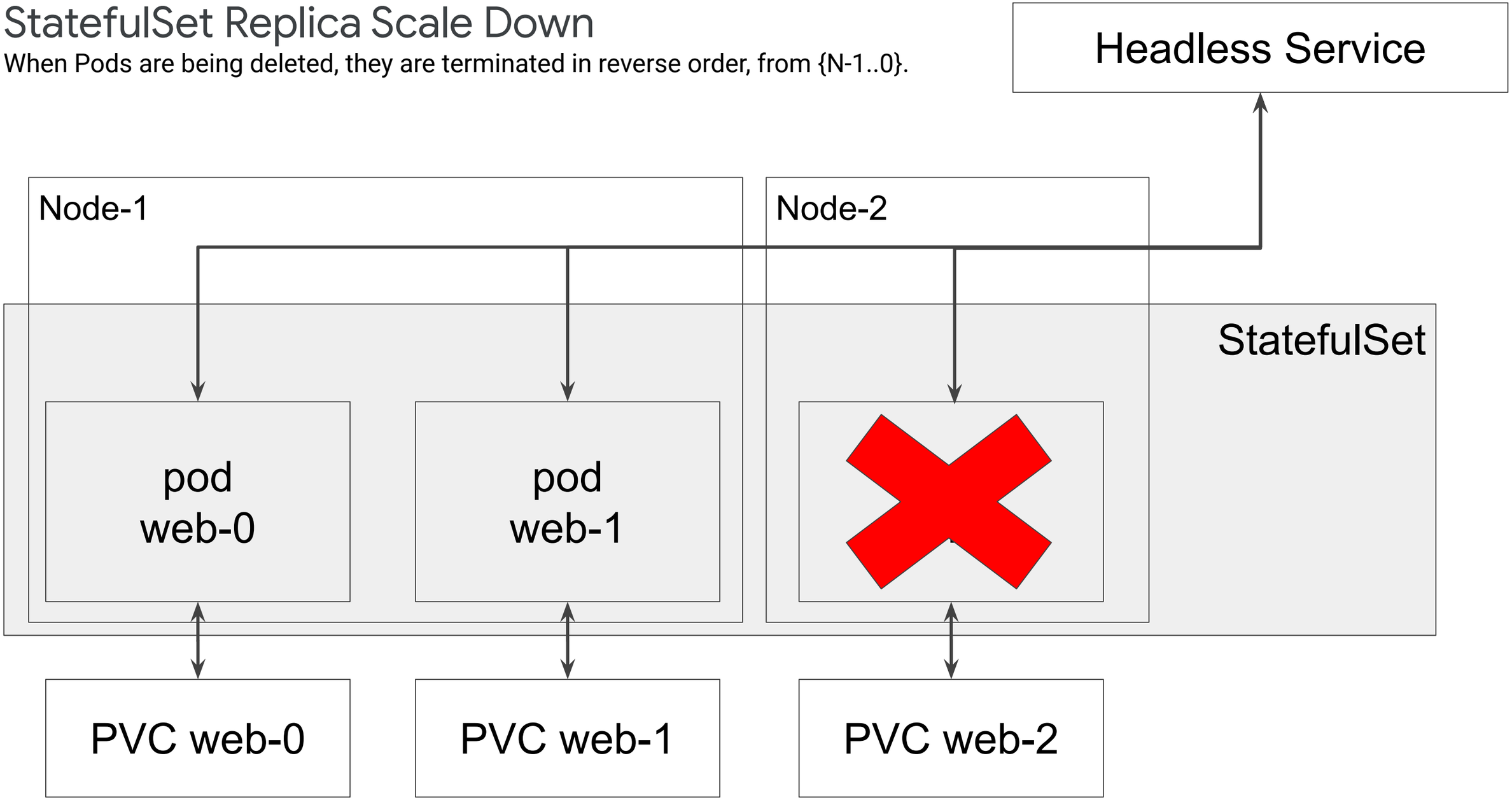


StatefulSet Replica = 3



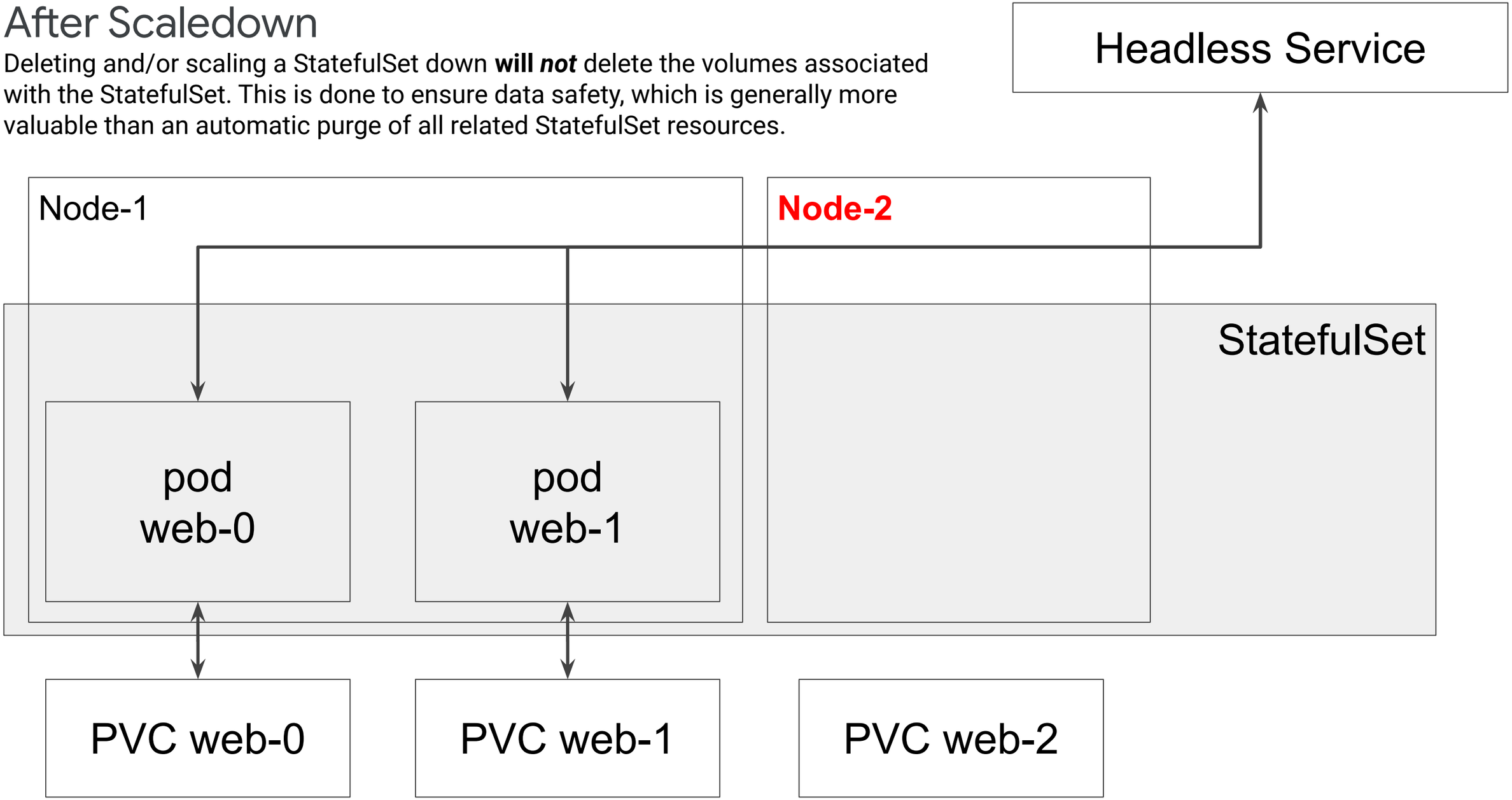
StatefulSet Replica Scale Down

When Pods are being deleted, they are terminated in reverse order, from {N-1..0}.

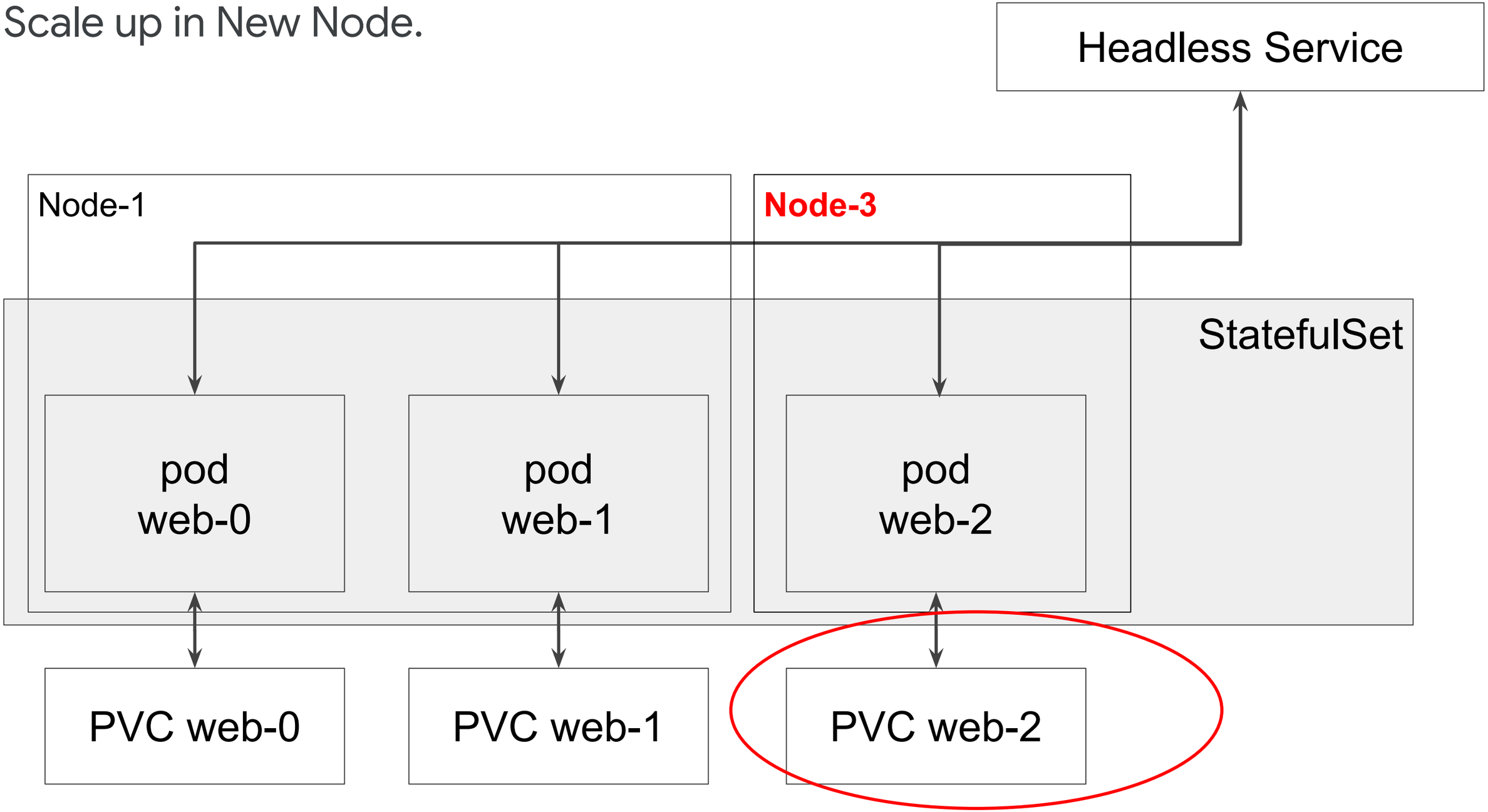


After Scaledown

Deleting and/or scaling a StatefulSet down **will *not*** delete the volumes associated with the StatefulSet. This is done to ensure data safety, which is generally more valuable than an automatic purge of all related StatefulSet resources.



Scale up in New Node.



Updating a StatefulSet

- The **RollingUpdate** update strategy implements automated, rolling update for the Pods in a StatefulSet.
- StatefulSet controller will delete and recreate each Pod in the StatefulSet in the same order as Pod termination (from the largest ordinal to the smallest), updating each Pod one at a time.
- It will wait until an updated Pod is Running and Ready prior to updating its predecessor.

It’s the DEMO time

Thank you!