Two years ago, in 2017…

- Longhorn was built on top of Docker Swarm
- Capable of fault detection and replica rebuild
- Snapshot/backups
- UI
Now

Open Source
Distributed Block Storage Software
For Kubernetes
https://github.com/longhorn/longhorn/

Add persistent storage support to any Kubernetes cluster
kubectl apply -f longhorn.yaml
Longhorn v0.6.1

- Enterprise-grade distributed block storage software for Kubernetes
- Volume snapshots
- Volume backup and restore
- Storage Tag for node and disk selection
- Cross-cluster disaster recovery volume with defined RTO and RPO
- Live upgrade of Longhorn software without impacting running volumes
- Intuitive UI
- One click installation
- And more features are coming
  - QoS, volume resizing, real time performance monitoring, et cetera.
**Longhorn Architecture - Engine**

Pod 1
- Volume
- Engine
- Replica
- Replica
- Replica

Pod 2
- Volume
- Engine
- Replica
- Replica
- Replica

Pod 3
- Volume
- Engine
- Replica
- Replica
- Replica

Node 1
- SSD
- SSD
- SSD
- SSD
- RAM
- CPU

Node 2
- SSD
- SSD
- SSD
- SSD
- RAM
- CPU

Orchestrated by Kubernetes
Longhorn Architecture - Manager

Container Storage Interface API

Longhorn CSI Plugin

Longhorn API

Longhorn Manager (Orchestrates all the volumes)

Engine

Replica 1

Replica 2

Engine

Replica 1

Replica 2

Engine

Replica 1

Replica 2

Kubernetes Cluster

Kubernetes API Server

Longhorn UI

Longhorn API
Kubernetes: Controller Pattern

Manager

Engine

Replica1
Node 1

Replica2
Node 2

Replica3
Node 3

Replica4
Node 4

Volume:
  spec:
    numberOfReplicas: 3

Status:
  currentHealthyReplicas: 3

Engine:
  spec:
    replicaList:
      Replica1
      Replica2
      Replica3

Status:
  replicaList:
    Replica1
    Replica2
    Replica3
Demo
Kubernetes: the container orchestrator

- Start Longhorn components in the cluster
- Node status monitoring
  - Make it easier to deal with failed/pressured nodes
- Pod status monitoring
- Automatic reattach volume after node reboot
Kubernetes: the application framework

- Provide data store to the manager
- Using CRD and controller pattern
  - Reduce complexity
  - Increase resiliency
Kubernetes: the user of Longhorn

- Easy to plug in
  - Using CSI or Flexvolume driver
- Storage Class
- Any Kubernetes workload
  - Deployment, StatefulSet
- Snapshotter, raw block device, volume clone...
The challenges: driver interface

- The driver interface is keep changing
- Flexvolume
- CSI v0.3, v0.4, v1.0, v1.1
- Mostly without backward compatible with older Kubernetes version
The challenges: stateful workload and Kubernetes HA

- Workload use RWO volume cannot self-healing if the node is down
- Stateful Set uses different volumes for each Pod
  - But it will not automatically create a new pod if the node of the old pod is down
- Deployment can automatically starts a new pod on a new node if the old pod’s node failed
  - But it won’t detach the volume from the old node, which will result in error for RWO volume since the volume can only be attached to one node
Upcoming Longhorn v0.7.0

- First beta release
- Volume resizing
- Performance metrics
- Quality of Service
Thank you!