Unified CSI plugin for All?

Experimenting Heterogeneous Storage with Single CSI Plugin for Kubernetes

Kiran Mova, Chief Architect, Co Founder, MayaData
Mohammad Asif Siddiqui, Senior System Architect, Huawei

Speakers

Kiran Mova

Chief Architect, Co-Founder
MayaData Inc
TSC, SODA Foundation

Mohammad Asif Siddiqui

Senior Systems Architect
Huawei
Member, ASF

@kiranmova

kiranmova

@asifdxtreme

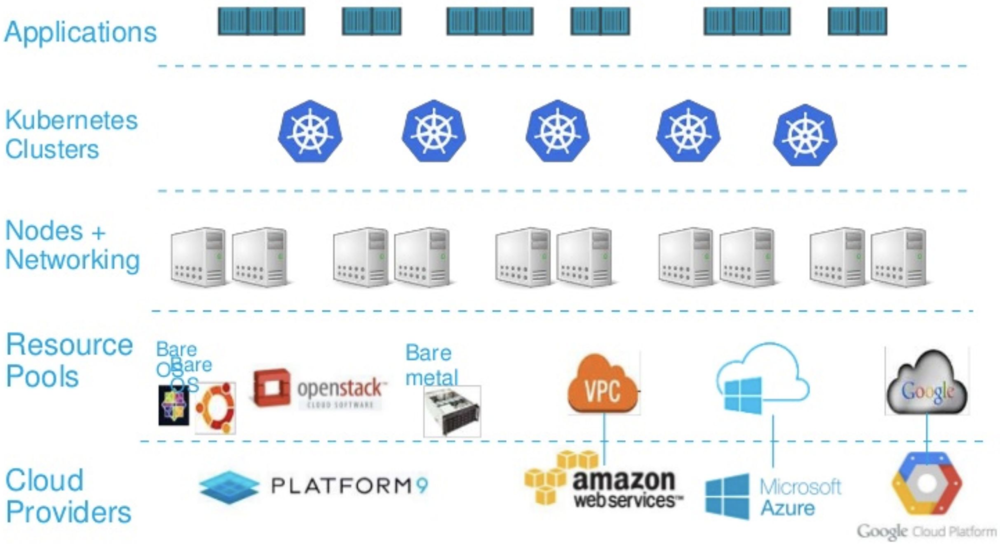
Mohammad Asif Siddiqui

OpenEBS Litmus CLOUD NATIVE COMPUTING FOUNDATION soda foundation THE APACHE SOFTWARE FOUNDATION VOLCANO KubeEdge

Agenda

- Overview on Kubernetes and CSI
- Understanding SODA CSI Plugin
- One CSI Plugin for all - An Approach

Why Kubernetes matters?



#1: Fault-tolerant by Design. Self-healing
#2: Modular App Design w/ Containers
#3: Automated Rolling Upgrades
#4: Auto-scaling. Reproducibility.
#5: Write Once, Run Anywhere. Portability.

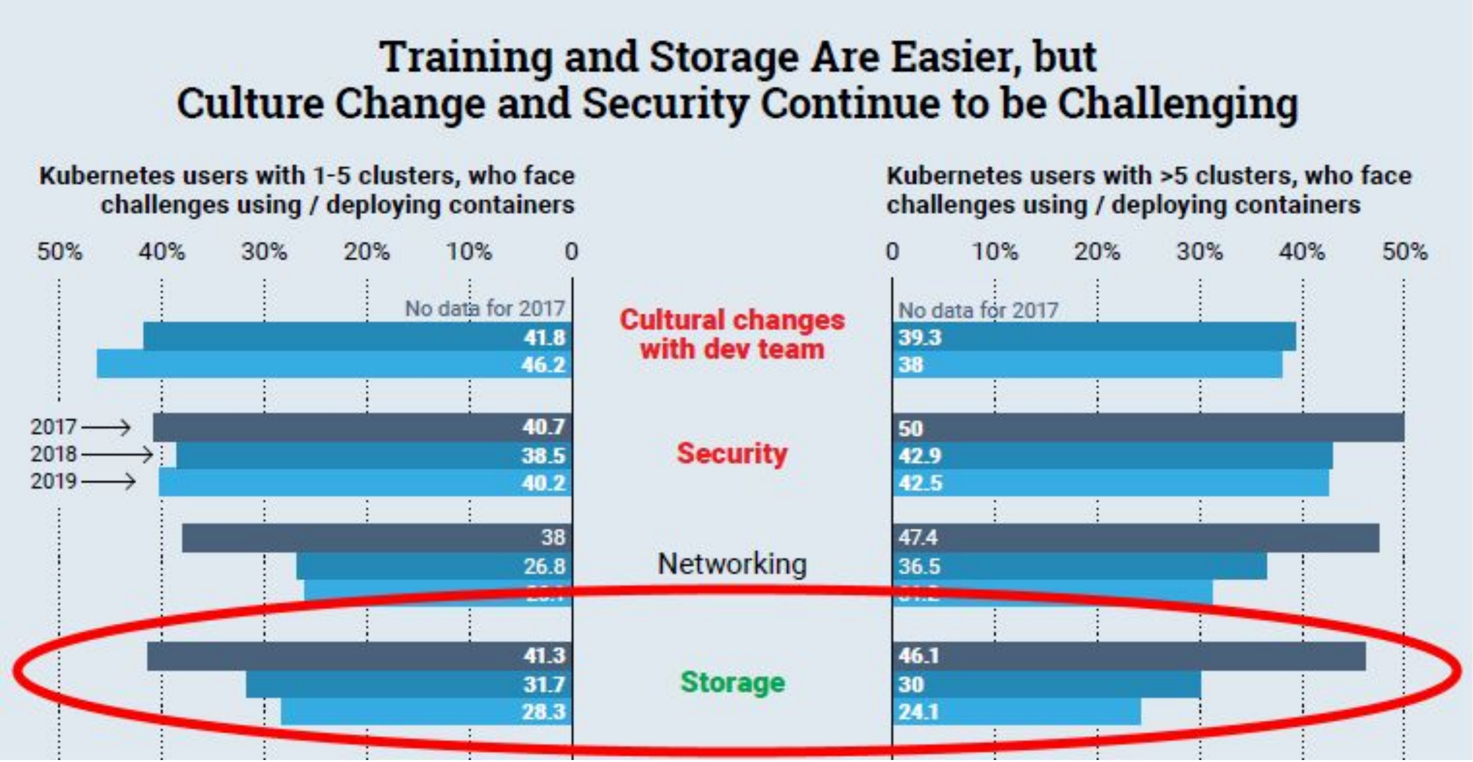
(K8s is the new Linux)

hybrid and multi-cloud environments
eases the deployment and management
manages app health, replication, load balancing
hardware resource allocation

CLOUD NATIVE
COMPUTING FOUNDATION

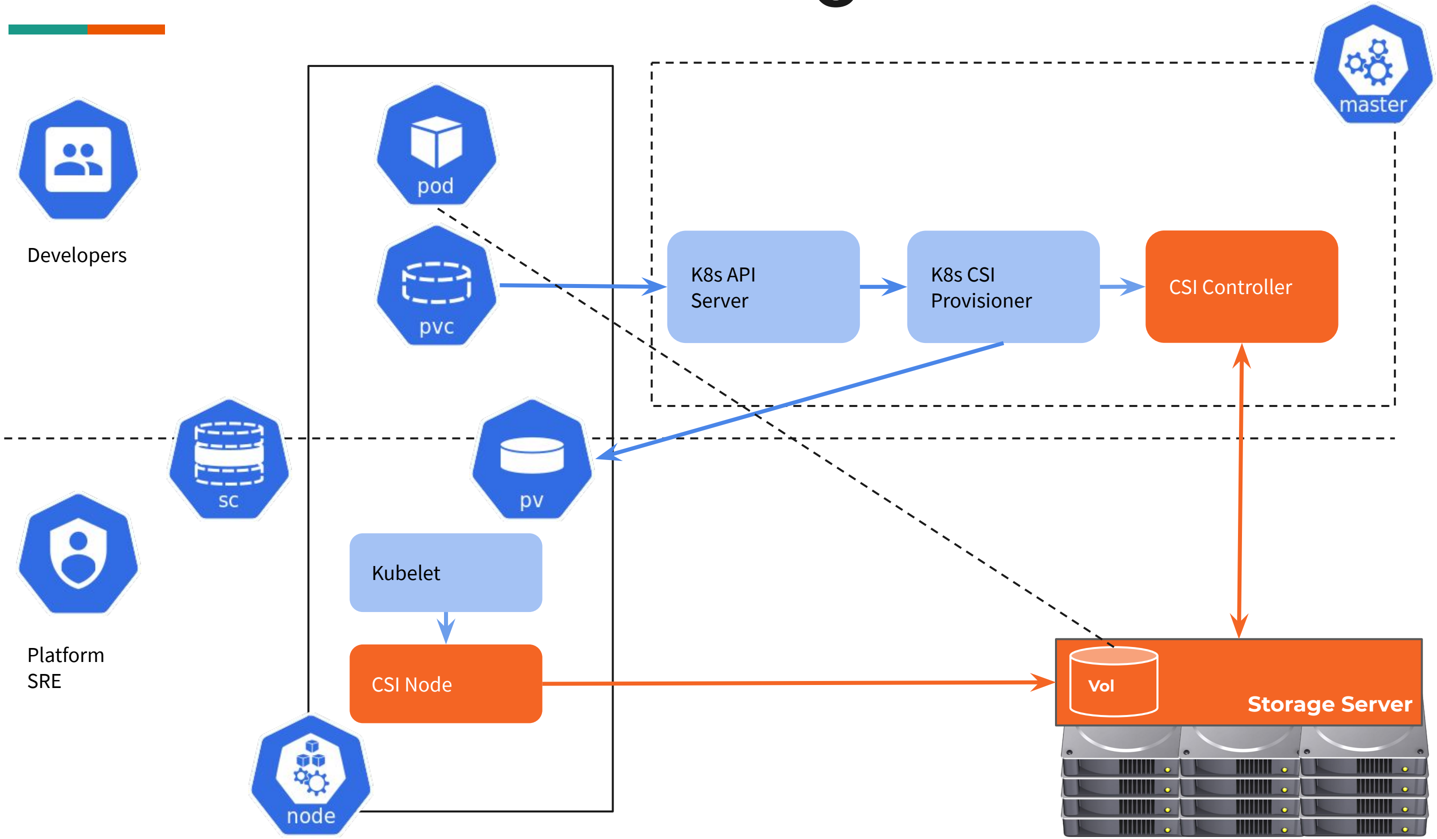
https://thenewstack.io/ebooks/kubernetes/state-of-kubernetes-ecosystem-second-edition-2020/

Data on Kubernetes

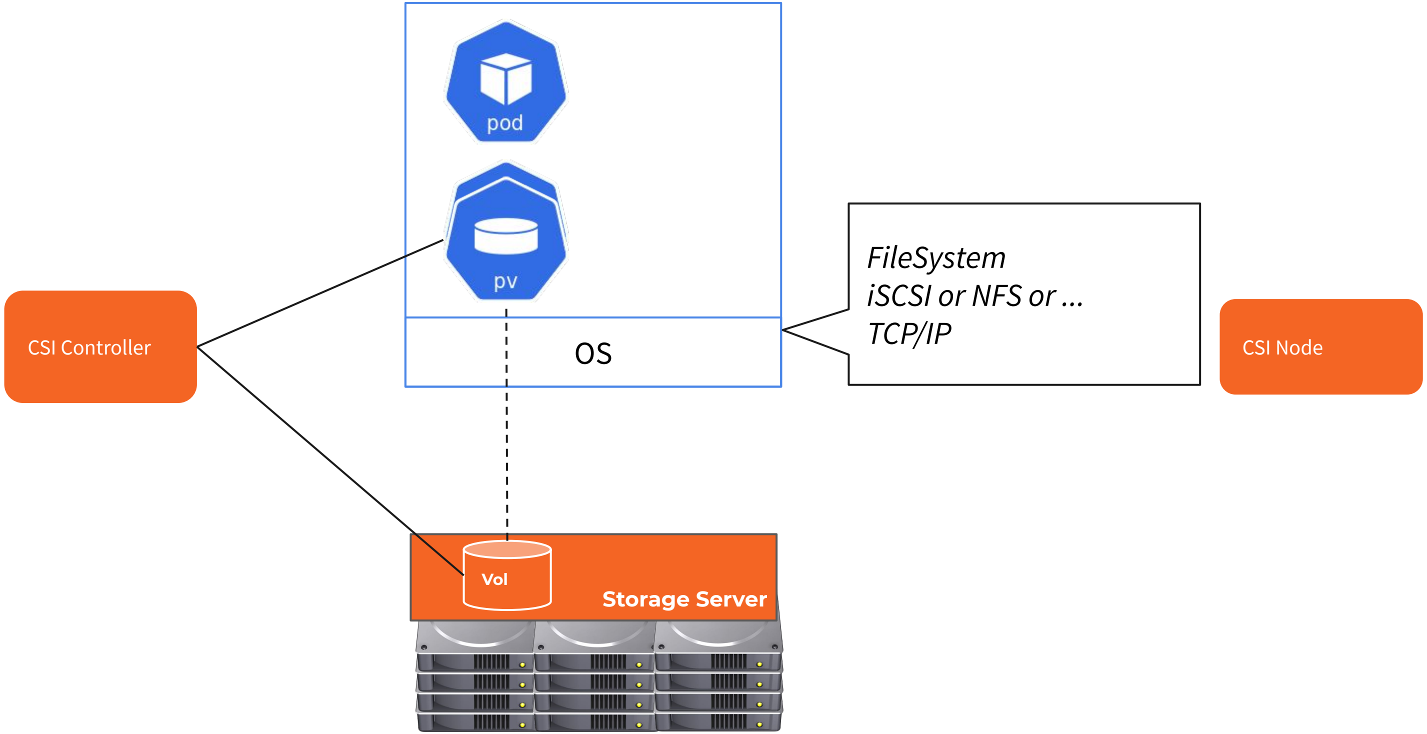


https://thenewstack.io/ebooks/kubernetes/state-of-kubernetes-ecosystem-second-edition-2020/

Data on Kubernetes using CSI



Data on Kubernetes using CSI



https://thenewstack.io/kubernetes-storage-dynamic-volumes-and-the-container-storage-interface/

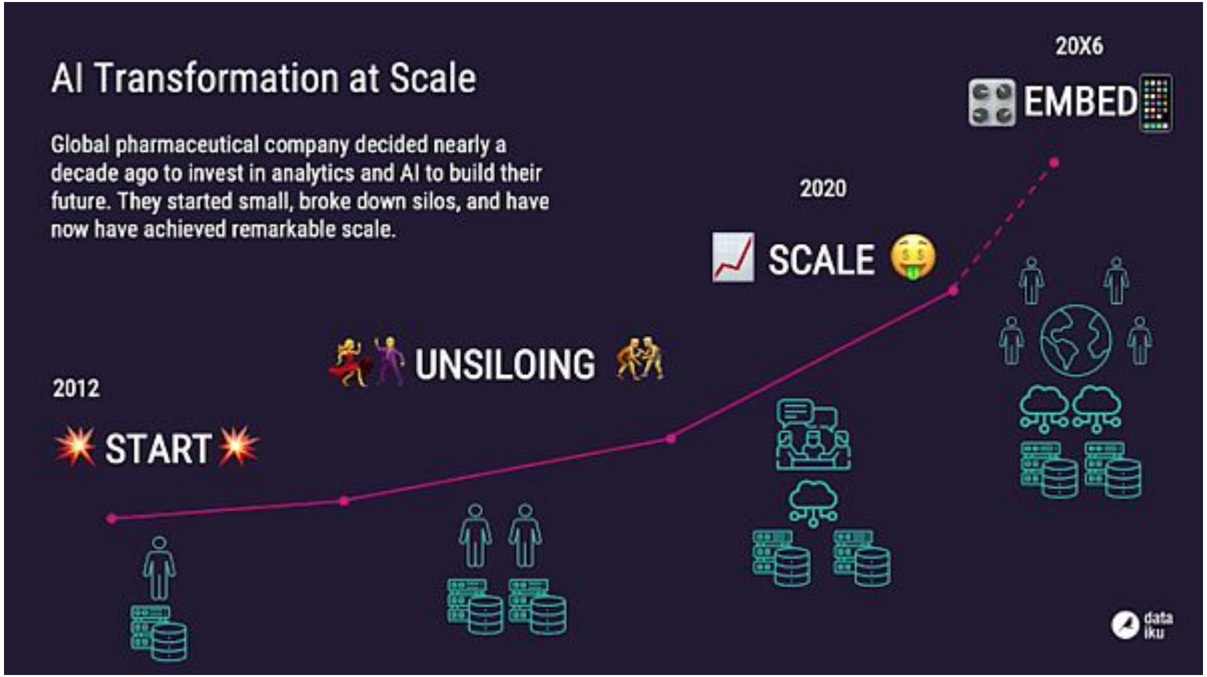
How does CSI help end-users?

So, although ***CSI doesn’t have major effects for end-users***, it correlates to a much *richer and better product offerings (90+)* from a wide range of storage vendors in the cloud native space.

- CSI solves the problem of upgrading storage functionality in Kubernetes without needing to upgrade Kubernetes itself.
- CSI thus requires that you have independent units running in a cluster which are privileged and able to speak the CSI language for attaching storage to a mount point.
- Container native storage does not require CSI in any way — you can run storage in a container and create volumes just as easily without CSI as with it. The only difference is that CSI allows vendors to have complete control over their release cycle.

https://thenewstack.io/kubernetes-storage-dynamic-volumes-and-the-container-storage-interface/

End-users: Evolving Data Infrastructure

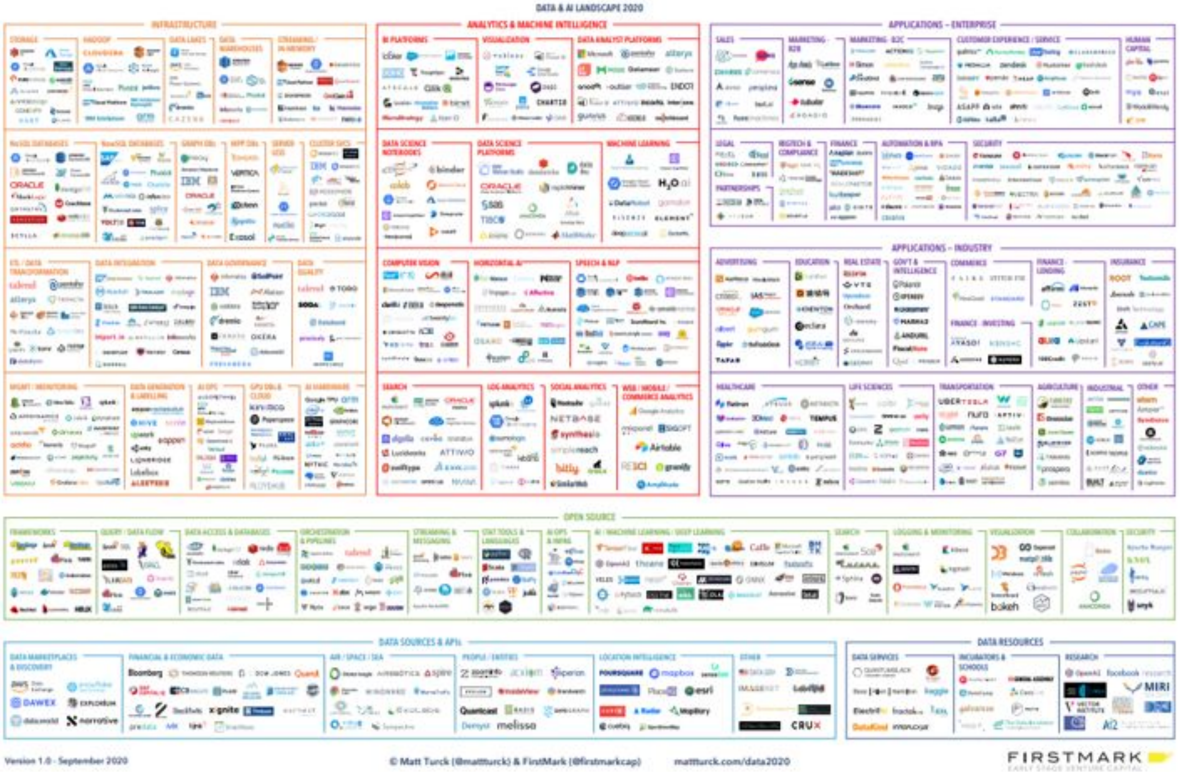


AI Transformation at Scale

Global pharmaceutical company decided nearly a decade ago to invest in analytics and AI to build their future. They started small, broke down silos, and have now achieved remarkable scale.

https://mattturck.com/data2020/

2020 - Big Data and AI Landscape



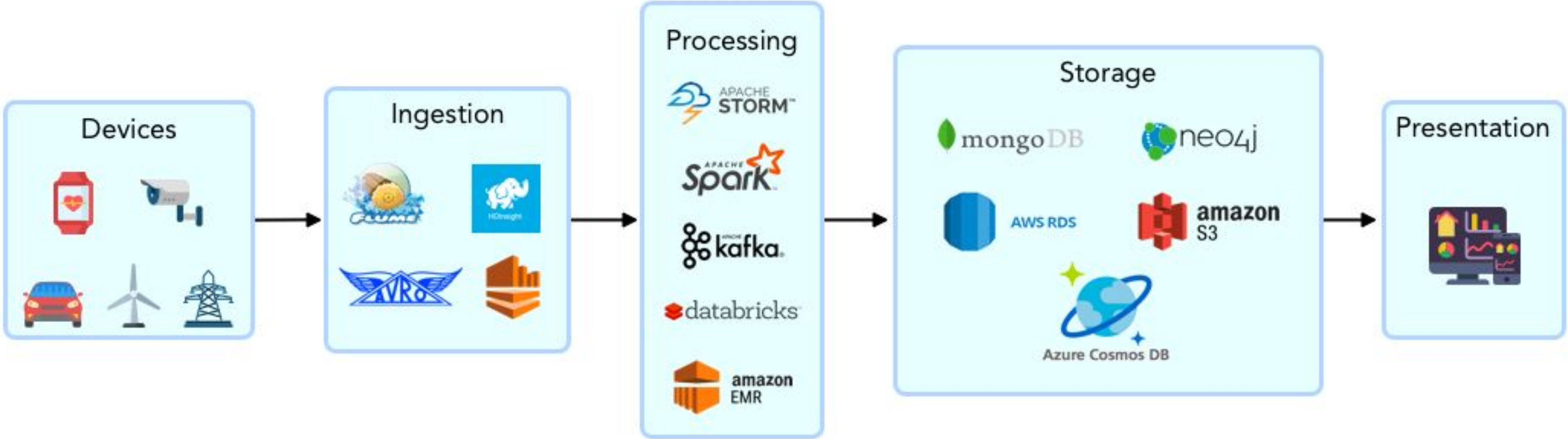
https://mattturck.com/data2020/

"to succeed, every modern company will need to be ***not just a software company, but also a data company****."*

"overall volume of data flowing through the enterprise continues to ***grow an explosive pace.*** The number of sources of data keeps increasing as well, with ever more SaaS tools. There is ***not one but many data pipelines*** operating in parallel in the enterprise."

"raises the bar on data infrastructure (and the teams building/maintaining it)"

Typical Data/Storage Stack



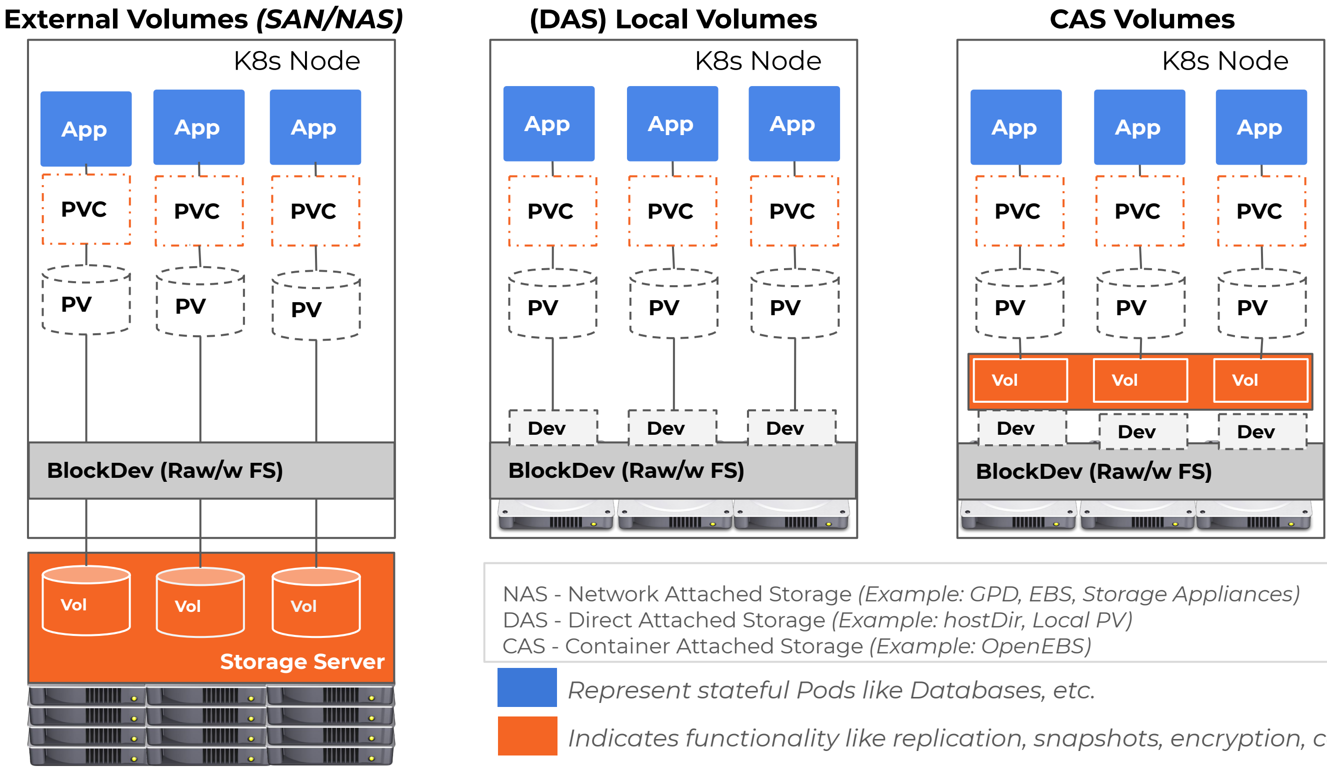
https://blog.qburst.com/2019/02/data-engineering-iot/

Workload (Stateful App) Needs

Workload	Needs Durability	Needs to be fast	Avg size per container	Dedicated disk?
AI Workloads with model or training caches	No	No	1 to 100GB	No
Postgres workloads with millions of rows	Yes	Yes	1G to 5TB	Yes
HDFS datanode backend folders	No	Yes	10TB	Yes
Cold storage	Yes	No	Infinite (assume that over time the amount of storage increases indefinitely)	No

https://thenewstack.io/kubernetes-storage-dynamic-volumes-and-the-container-storage-interface/

Persistent Volume types



OpenEBS

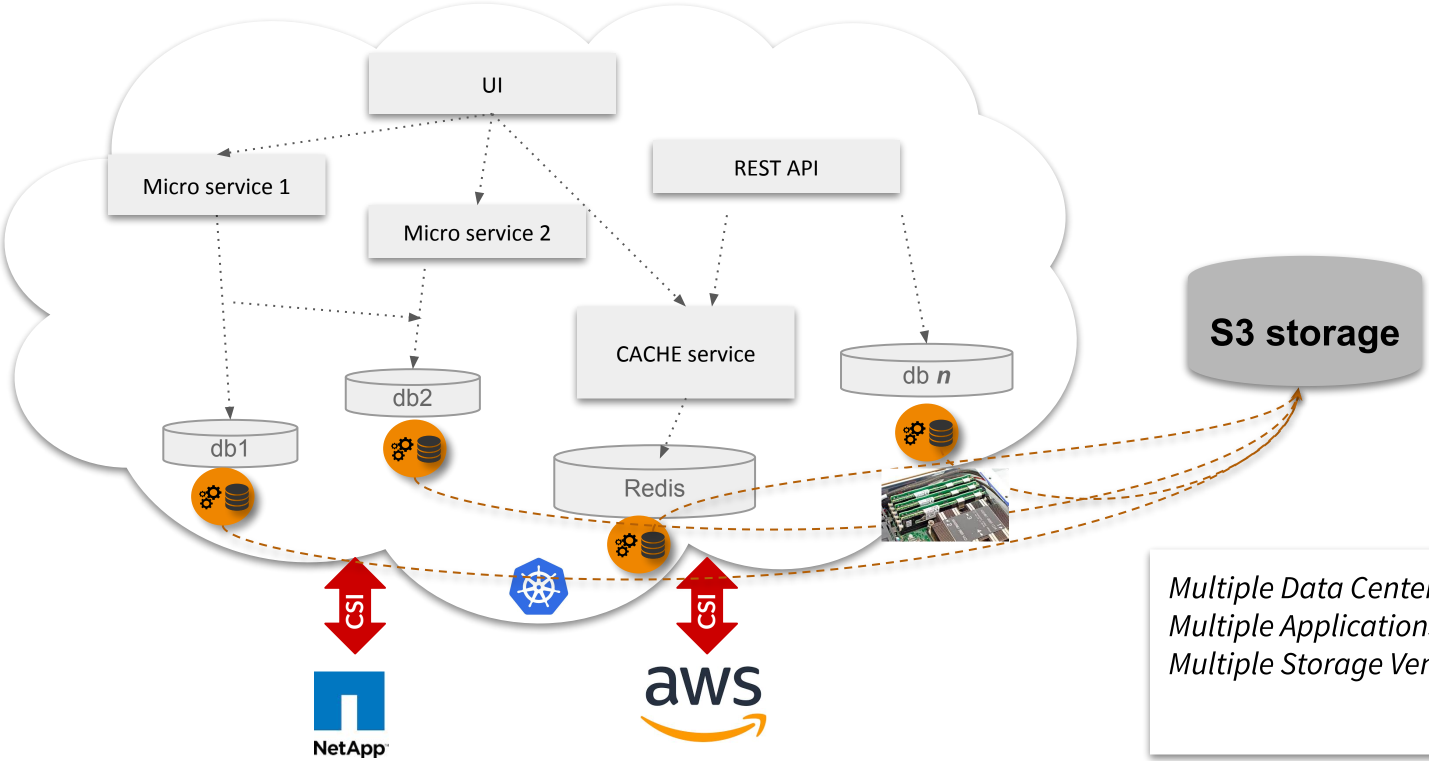
NAS - Network Attached Storage *(Example: GPD, EBS, Storage Appliances)*
DAS - Direct Attached Storage *(Example: hostDir, Local PV)*
CAS - Container Attached Storage *(Example: OpenEBS)*

🟦 *Represent stateful Pods like Databases, etc.*

🟧 *Indicates functionality like replication, snapshots, encryption, compression, etc.*

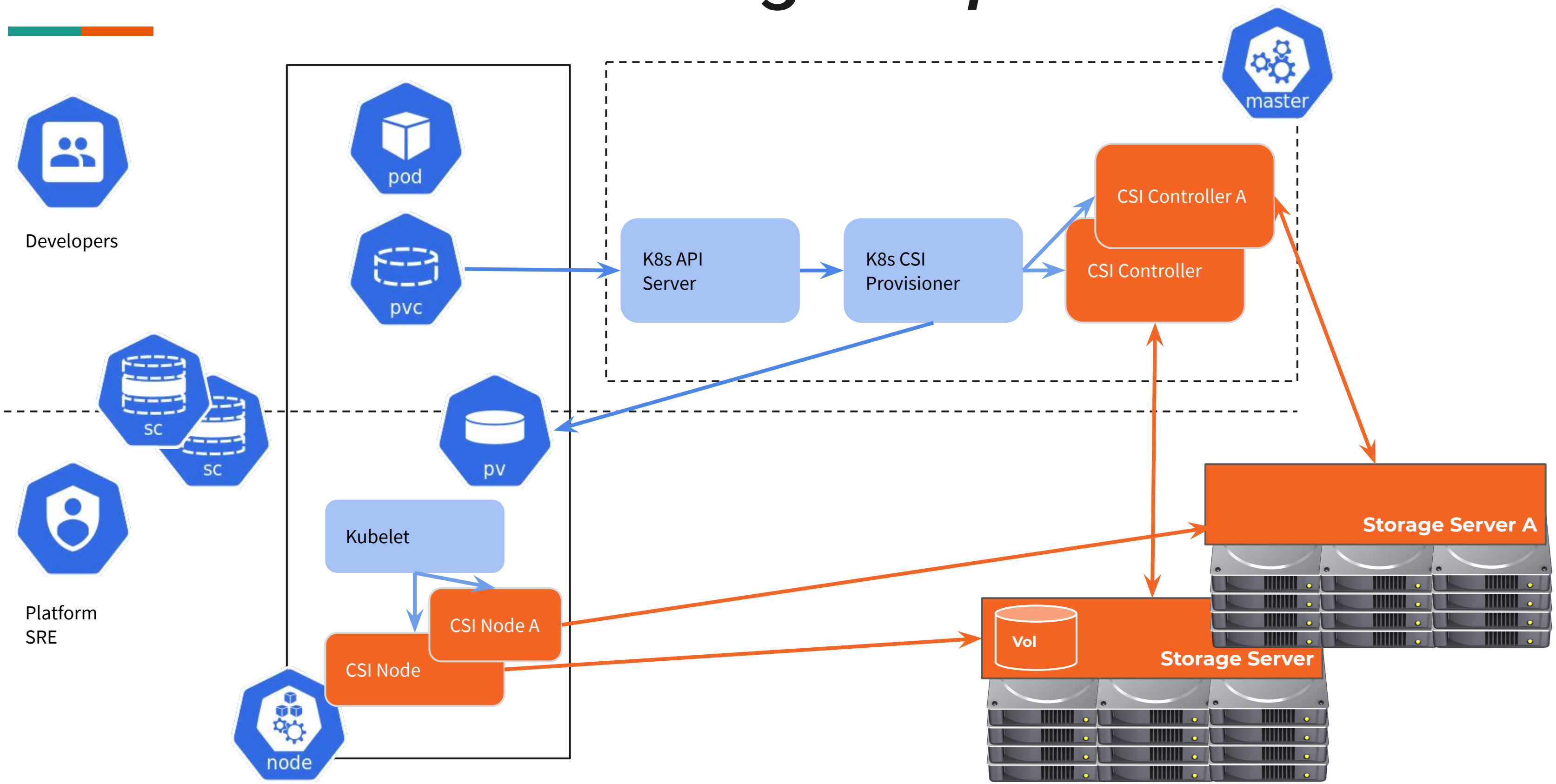
https://www.cncf.io/blog/2020/09/22/container-attached-storage-is-cloud-native-storage-cas/

Data on Kubernetes



Multiple Data Centers
Multiple Applications
Multiple Storage Vendors

Data on Kubernetes using *multiple* CSI



CSI Challenges with Heterogeneous Storage

Developers

All I care is about PVC. Now I have to deal with selecting which Storage Class to pick and what features it provides.

Additional Kubernetes Operators to simplify this complexity?

Platform SRE

Overhead of maintaining life cycle (install and upgrades) of multiple CSI drivers. The complexity of managing the compatibility between Kubernetes versions and CSI Drivers.

More pods! Overhead on system resources.

In-tree volumes were better!

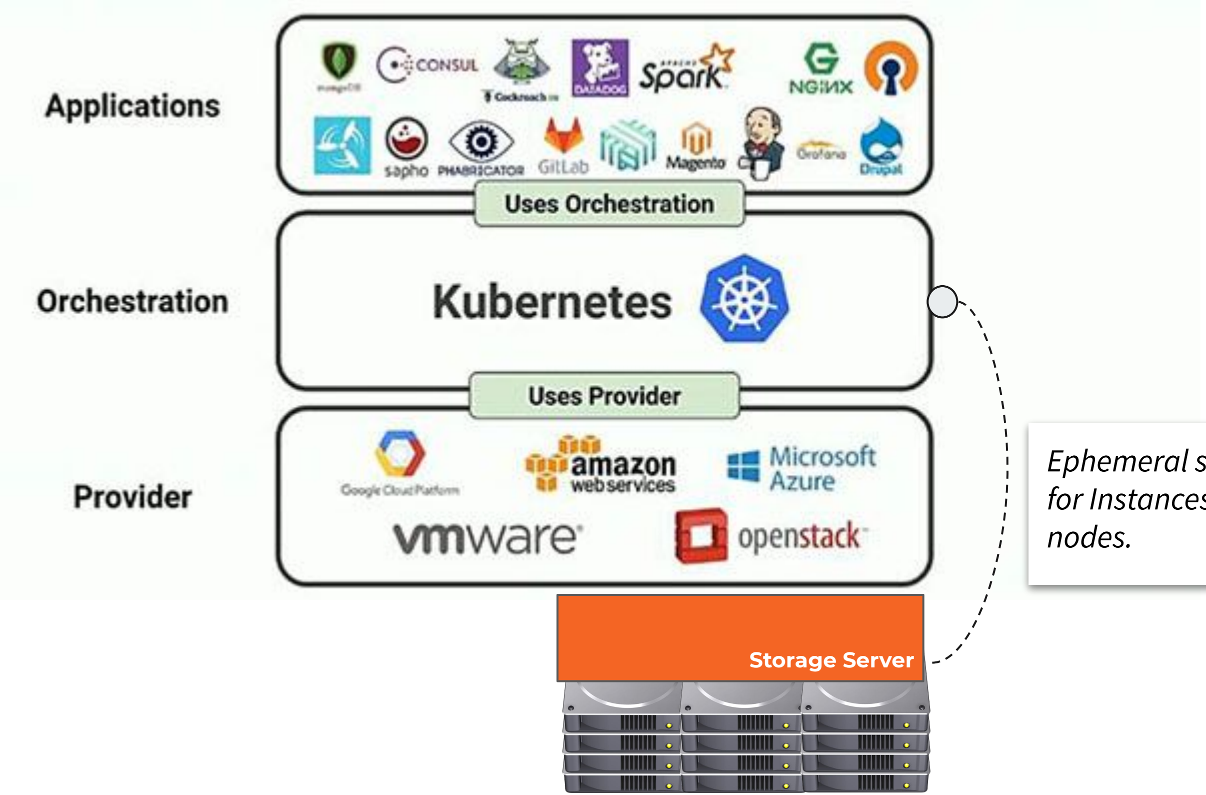
Evolving Specification. (*Data Protection!)*

Doesn't deal with Data Migration across multiple providers or moving across multiple Kubernetes Clusters.

Doesn't manage the underlying storage, which comes with its own challenges even for a single provider. *How to deal with failed drives? How to ensure availability zones?*

And VMs and Bare Metal!

Who provides storage to K8s itself?



Ephemeral storage (aka Root Disks) for Instances running Kubernetes nodes.

Data lives on, K8s clusters don’t

Cluster Lifecycle

Augment K8s capabilities that can be used by Data(Application) Operators.

DATA SCHEDULERS

DATA SECURITY

Data Operators

Manage the data beyond the life of a single K8s cluster. And beyond the application purview.

Data Lifecycle

DATA PROTECTION

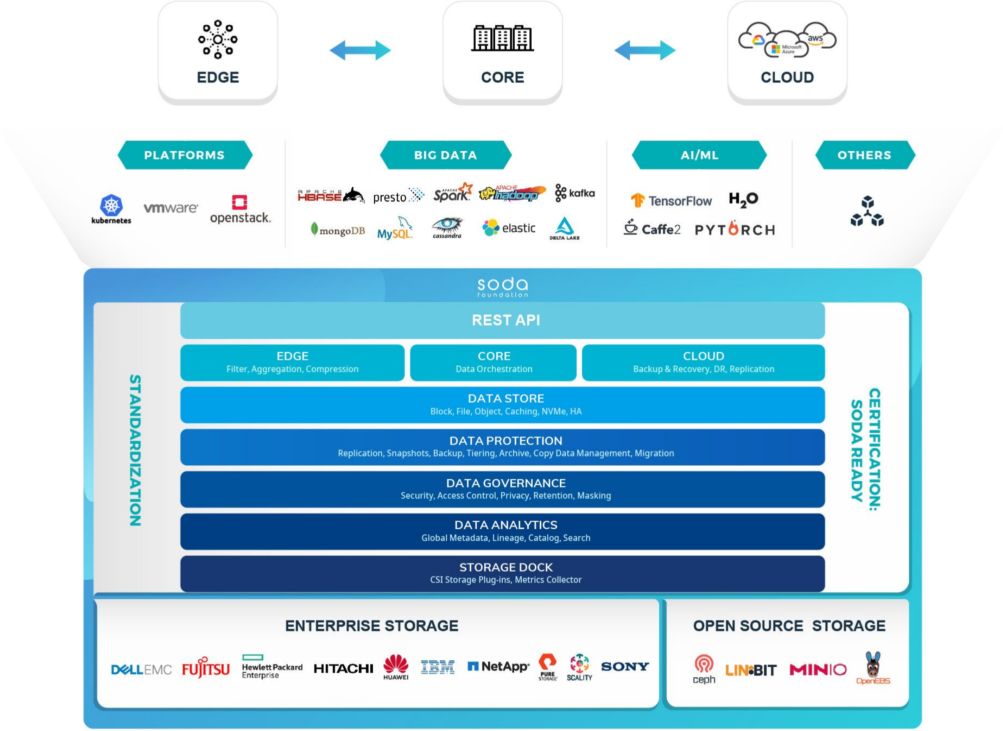
DATA MIGRATION

DATA GOVERNANCE

DATA COMPLIANCE

UNIFIED STORAGE PLATFORM

SODA: Unified CSI Driver?



EDGE ⟷ CORE ⟷ CLOUD

PLATFORMS | BIG DATA | AI/ML | OTHERS

REST API

EDGE
Filter, Aggregation, Compression

CORE
Data Orchestration

CLOUD
Backup & Recovery, DR, Replication

DATA STORE
Block, File, Object, Caching, NVMe, HA

DATA PROTECTION
Replication, Snapshots, Backup, Tiering, Archive, Copy Data Management, Migration

DATA GOVERNANCE
Security, Access Control, Privacy, Retention, Masking

DATA ANALYTICS
Global Metadata, Lineage, Catalog, Search

STORAGE DOCK
CSI Storage Plug-ins, Metrics Collector

STANDARDIZATION

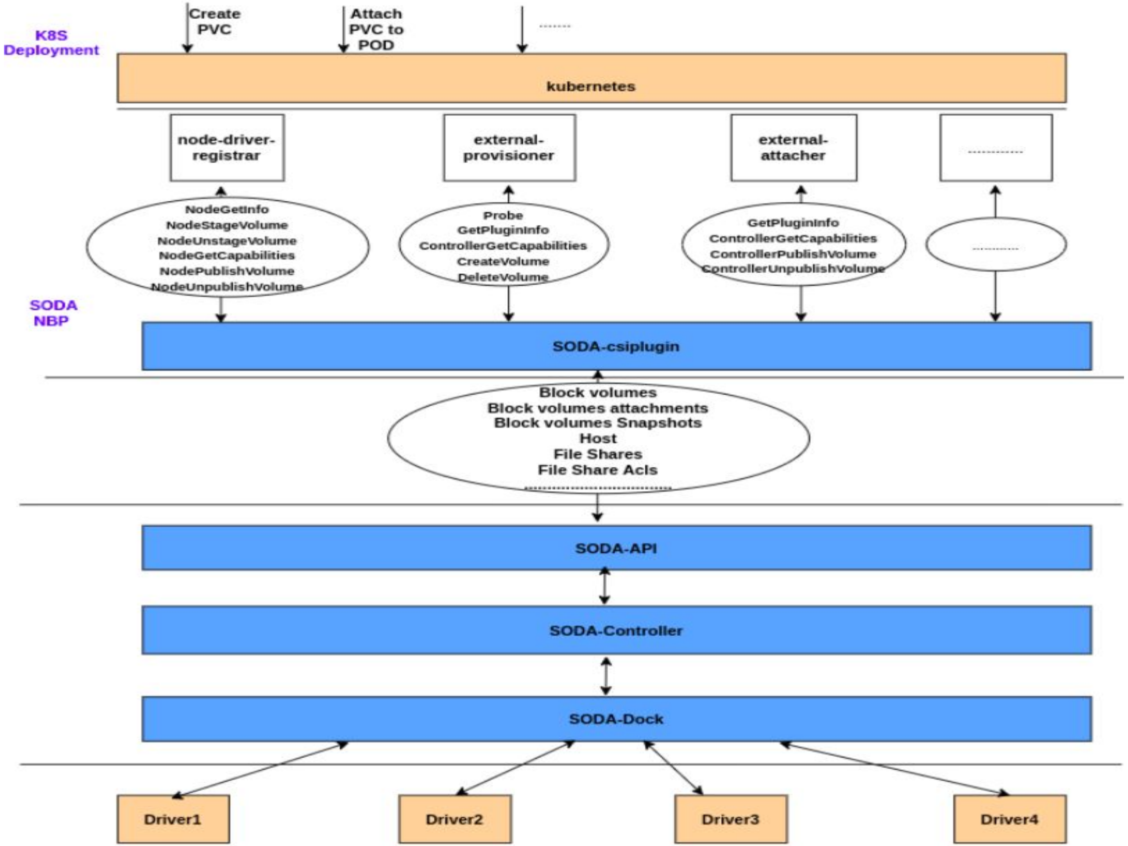
CERTIFICATION: SODA READY

ENTERPRISE STORAGE

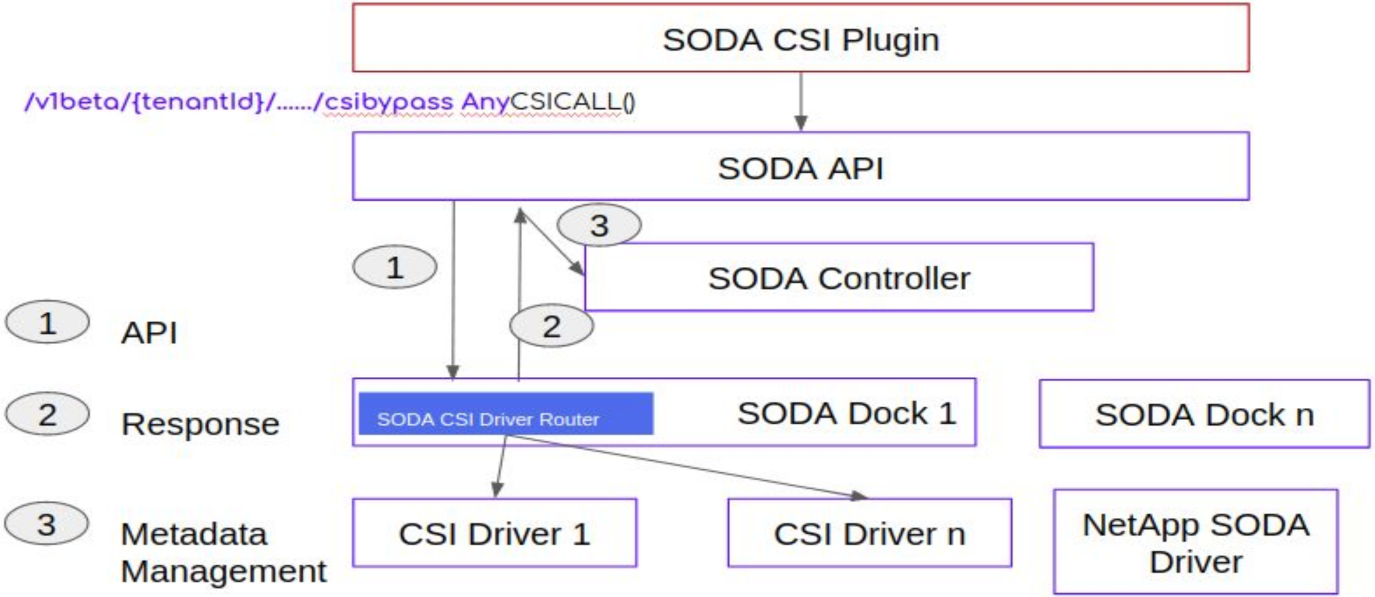
OPEN SOURCE STORAGE

So, although *CSI does help storage vendors with Kubernetes integration*, ***CSI doesn't solve the growing use cases for Data management.***

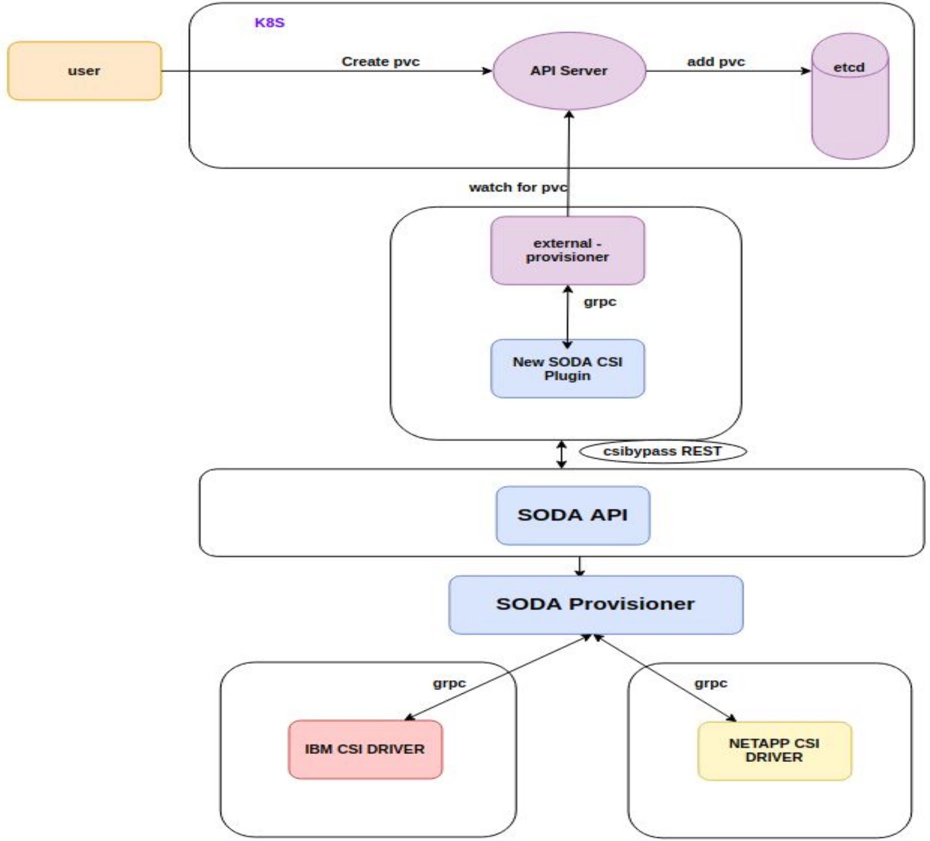
Soda ecosystem and it’s CSI Plugin



CSI PlugNPlay (One CSI plugin for all heterogenous drivers)



Volume Provisioning for PlugNPlay



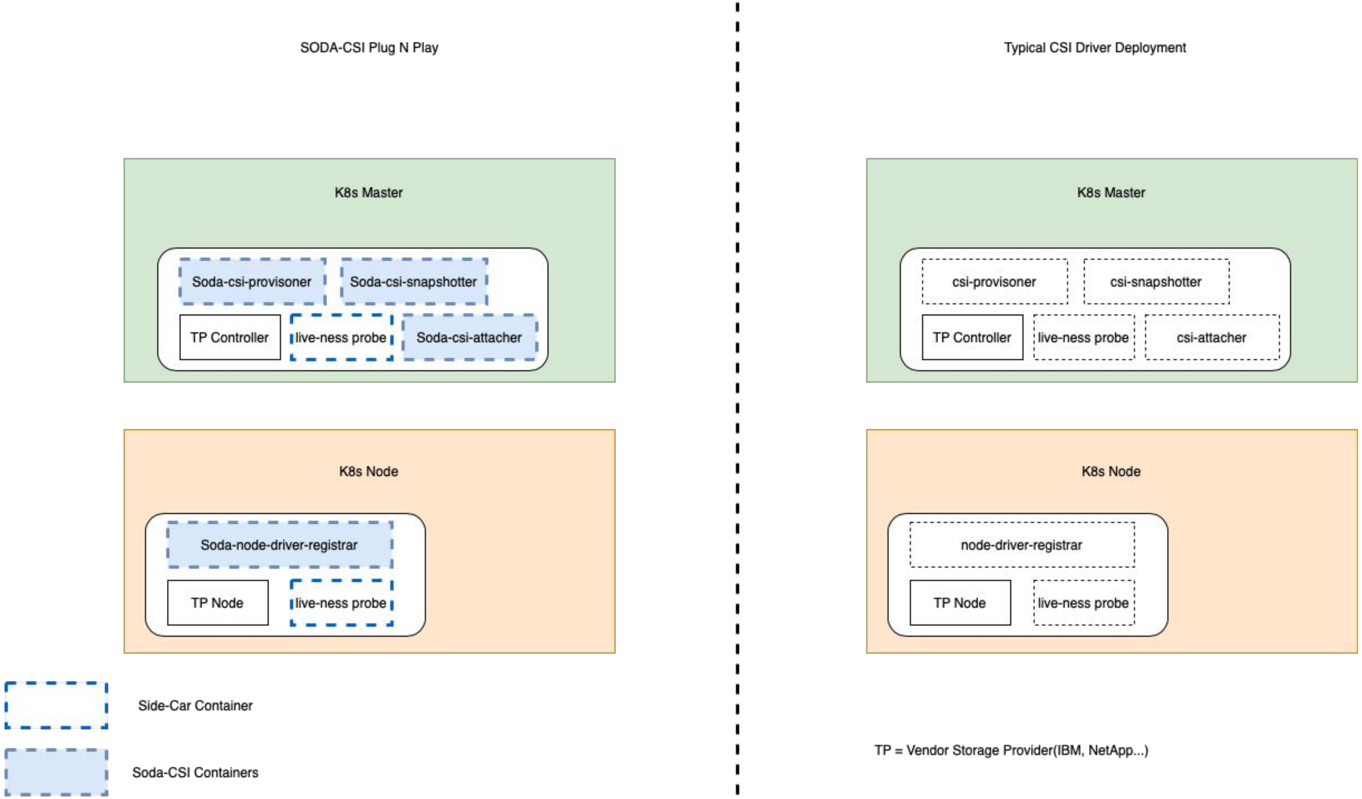
Soda CSI PlugNPlay (Alternative Approach)

Goals

Make a single CSI Plug-N-Play mechanism which helps to:

- Provision and manage the Storage of heterogeneous Storage providers.
- Third party CSI drivers should be used directly from Vendors so that SODA/Users need not worry about it's maintenance.
- Use Soda Profile ID to determine which storage vendor drivers need to be used to provision the storage.
- Experience all the features set of SODA.

Deployment View



SODA-CSI Plug N Play

K8s Master

Soda-csi-provisoner

Soda-csi-snapshotter

TP Controller

live-ness probe

Soda-csi-attacher

K8s Node

Soda-node-driver-registrar

TP Node

live-ness probe

Side-Car Container

Soda-CSI Containers

Typical CSI Driver Deployment

K8s Master

csi-provisoner

csi-snapshotter

TP Controller

live-ness probe

csi-attacher

K8s Node

node-driver-registrar

TP Node

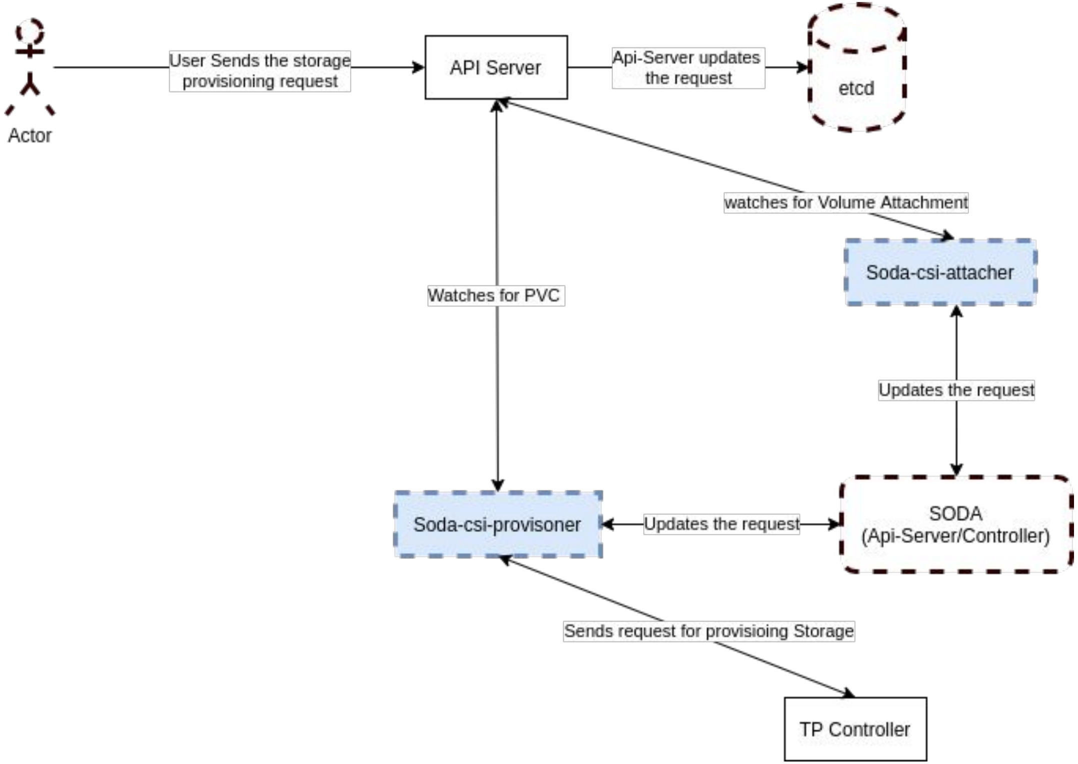
live-ness probe

TP = Vendor Storage Provider(IBM, NetApp...)

Call flows

A typical call flow for SODA CSI Plug-N-Play mechanism will be:

- Once the PVC Object is created by user then soda-csi-provisioner will get a request to provision the Storage.
- soda-csi-provisioner will interact with SODA API's to get the profile details and other parameters to determine if the required driver is same as the current driver in a pod.
- soda-csi-provisioner updates the SODA API server with the volume provisioning request and gather the required intelligence from SODA about the current provisioning.
- soda-csi-provisioner will forward the request to CSI driver in the same pod to do the actual provisioning.


Actor
User Sends the storage provisioning request
API Server
Api-Server updates the request
etcd
watches for Volume Attachment
Soda-csi-attacher
Watches for PVC
Updates the request
Soda-csi-provisoner
Updates the request
SODA (Api-Server/Controller)
Sends request for provisioing Storage
TP Controller

Advantages of the SODA CSI PlugNPlay

- A simple design which let's user to patch the existing Third party CSI drivers deployment without any changes to actual CSI driver containers provided by TP vendors.
- All the sidecar containers which will be used by TP CSI drivers will be provided by SODA.
- Robust and easy to maintain design which can be used by heterogeneous csi storage providers with minimal changes to side cars.
- Easy to upgrade along with TP CSI driver.
- Easy to maintain the side car code as the code-base will be used from kubernetes-csi org and the SODA features will be added to it as a plugin, so it's easy to upgrade side-car as k8s csi spec evolves.

Demo Video

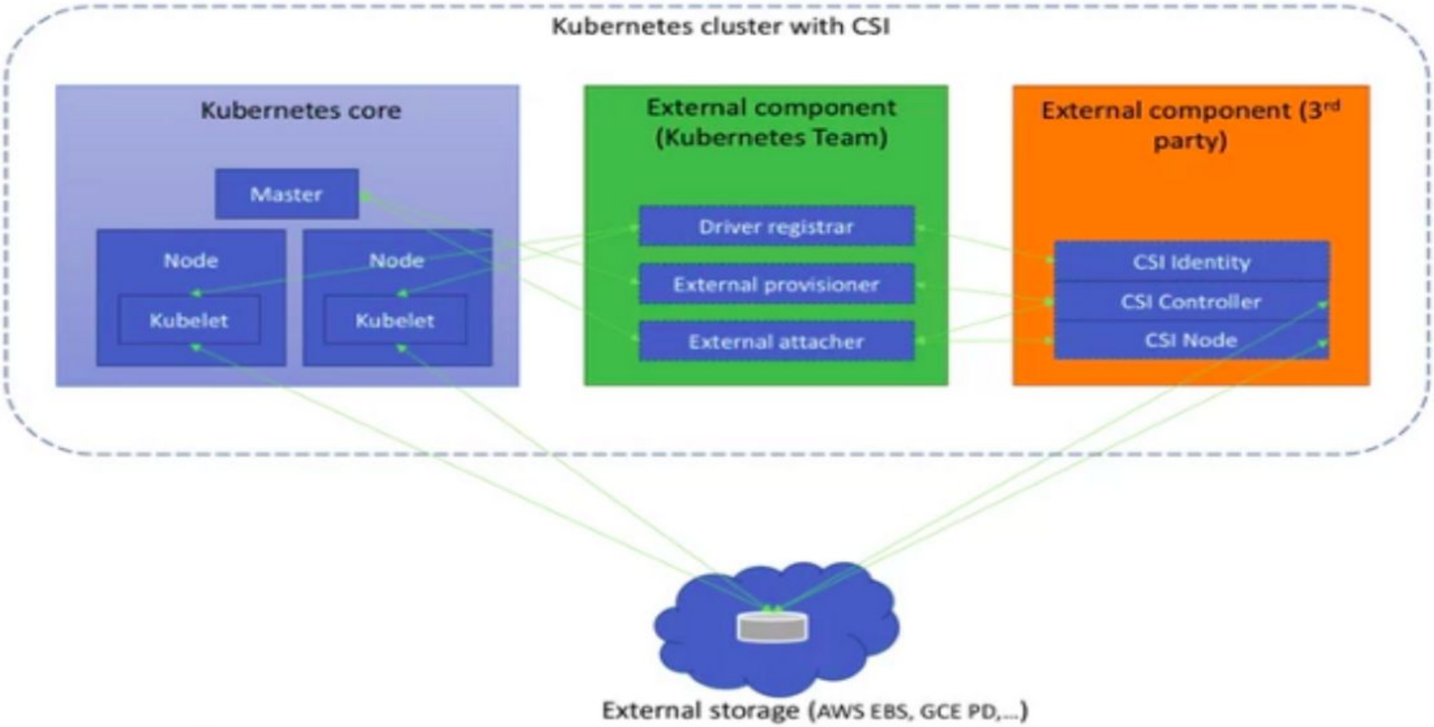
Q and A

Additional Information

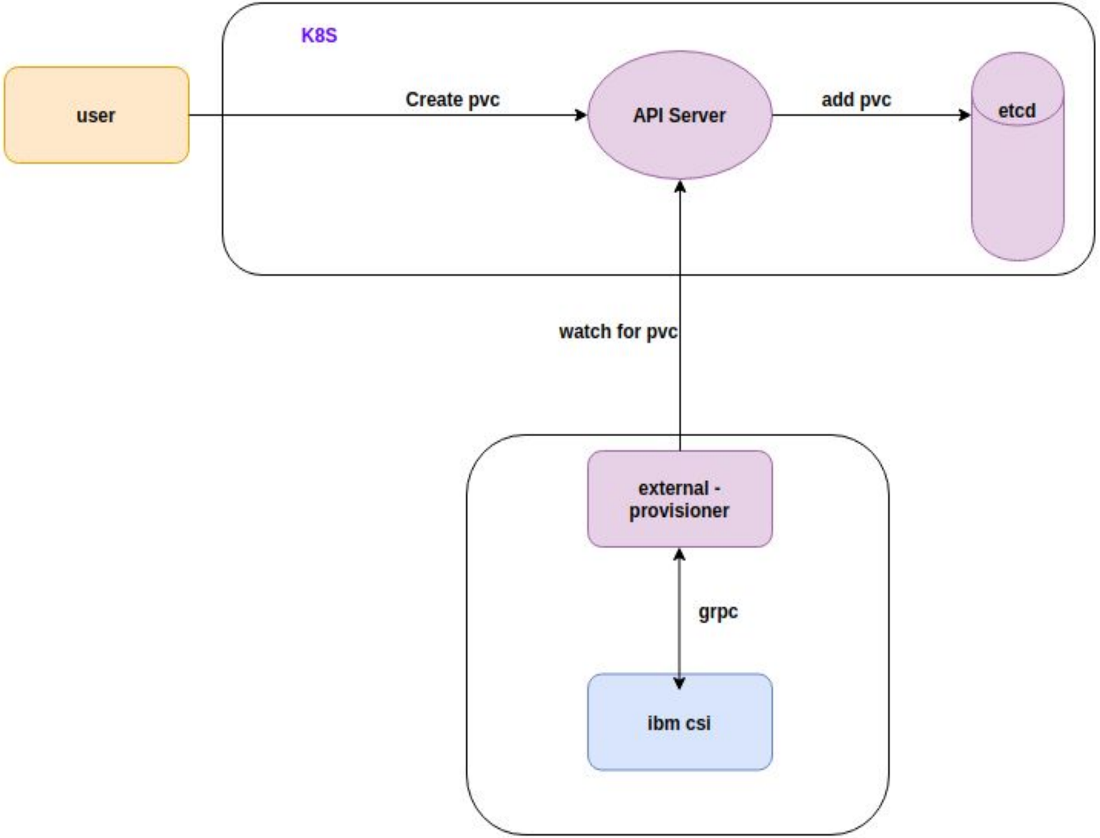
- SODA slack. https://sodafoundation.io/slack/
- https://github.com/sodafoundation/design-specs
- https://www.meetup.com/OpenDataAutonomy/
- https://www.meetup.com/Bangalore-CNCF-Meetup/
- https://www.meetup.com/Cloud-Native-Data-Management/
- https://www.meetup.com/Data-on-Kubernetes-community/
- Demo Video : https://youtu.be/ytXY_dKQCYg

@kiranmova	@asifdxtreme
kiranmova	Mohammad Asif Siddiqui

Kubernetes and CSI Interaction


Kubernetes cluster with CSI
Kubernetes core
Master
Node
Kubelet
Node
Kubelet
External component (Kubernetes Team)
Driver registrar
External provisioner
External attacher
External component (3rd party)
CSI Identity
CSI Controller
CSI Node
External storage (AWS EBS, GCE PD,...)

Normal Provisioning for CSI Drivers


K8S
user
Create pvc
API Server
add pvc
etcd
watch for pvc
external - provisioner
grpc
ibm csi