

sodaCON 2020

DATA CONNECTED

#sodacon2020

Connecting Data for Future! - SODA Roadmap

Sanil Kumar D

TOC, AWG Lead, SODA Foundation
Huawei
@sanilkumard

Kiran Mova

TOC, AWG, SODA Foundation
MayaData, OpenEBS
@kiranmova





IMMERSED IN DATA...

DATA EXPLOSION IS NEW NORMAL

1 ZB = 1000^7 bytes = 10^{21} bytes = 1000000000000000000000000 bytes = 1000 exabytes

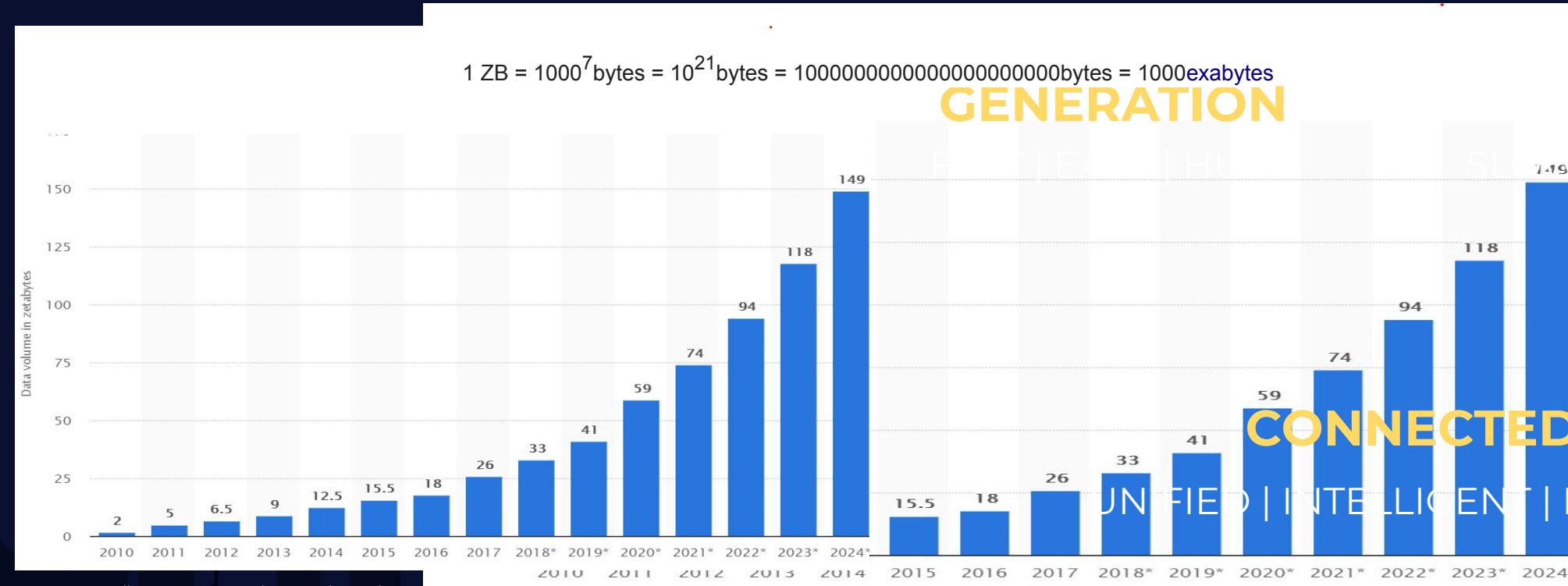
GENERATION

VALUE

| TOUGH | LESS

CONNECTED

| INTELLIGENT | EFFICIENT



Source : <https://www.statista.com/statistics/871513/worldwide-data-created/>

NARROWING BOUNDARIES | CORE, CLOUD & EDGE

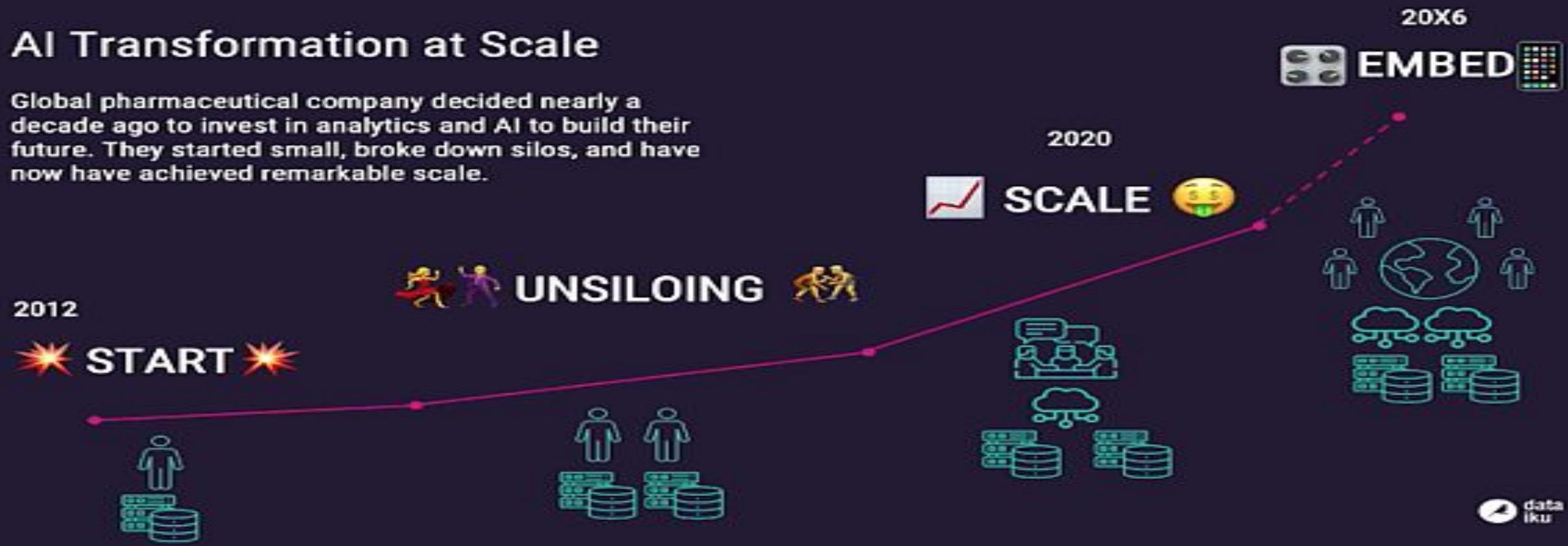


SEAMLESS CONNECTED DATA

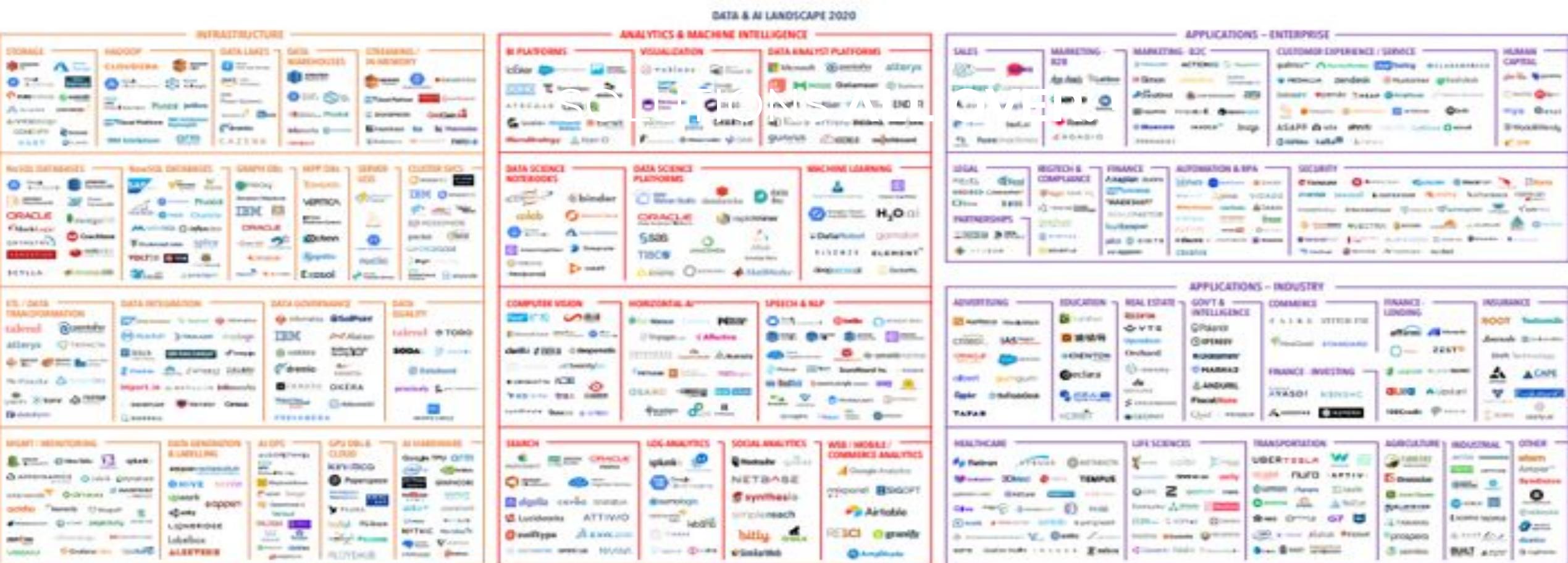
UNPREDICTABLE SCALE

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 have achieved remarkable scale.

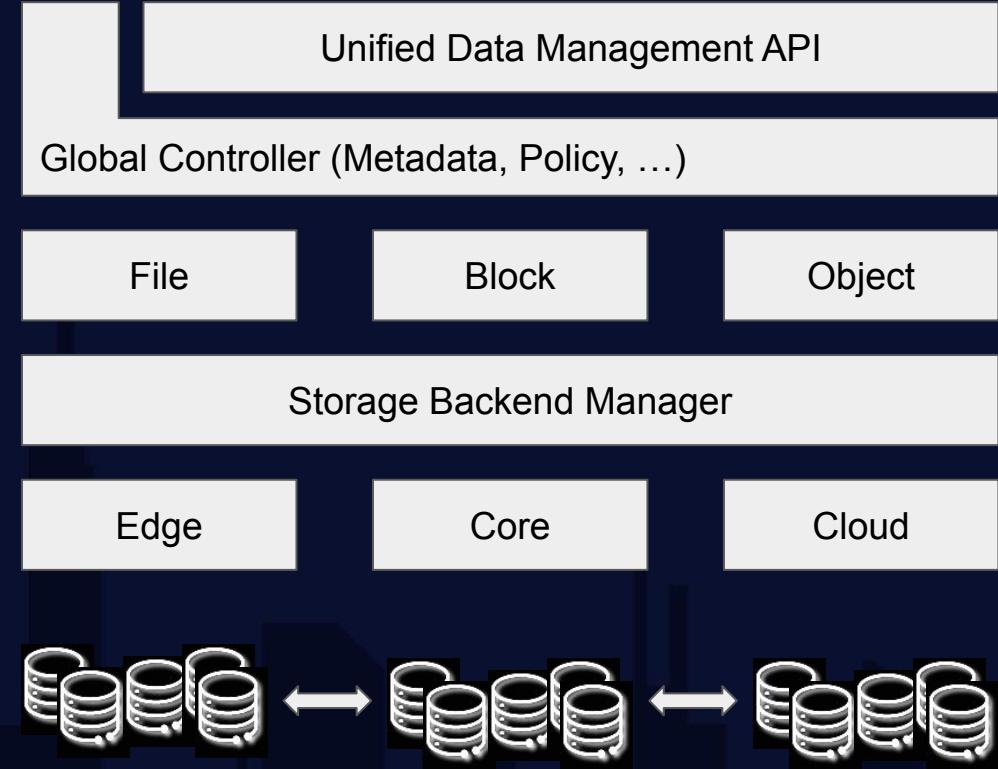


<https://mattturck.com/data2020/>

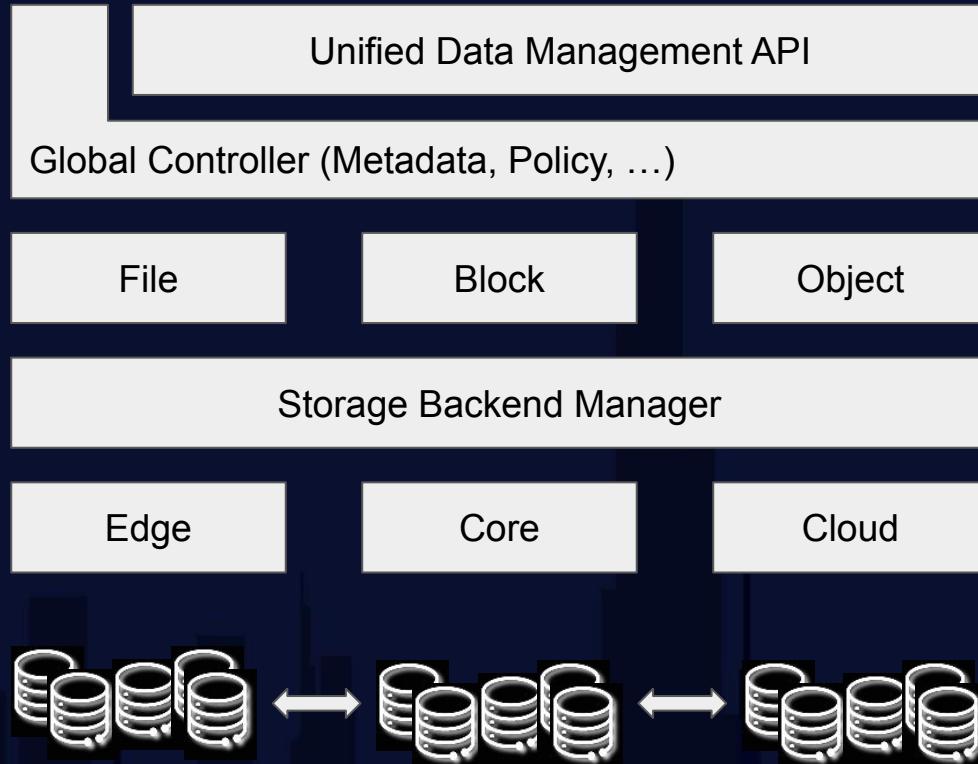


CONNECTING DATA : ONE DATA STORE

**Unified
Safe
Efficient
Intelligent
Traceable**



ONE DATA STORE : CHALLENGES



UNIFY ACROSS VENDORS, PLATFORMS & DOMAINS

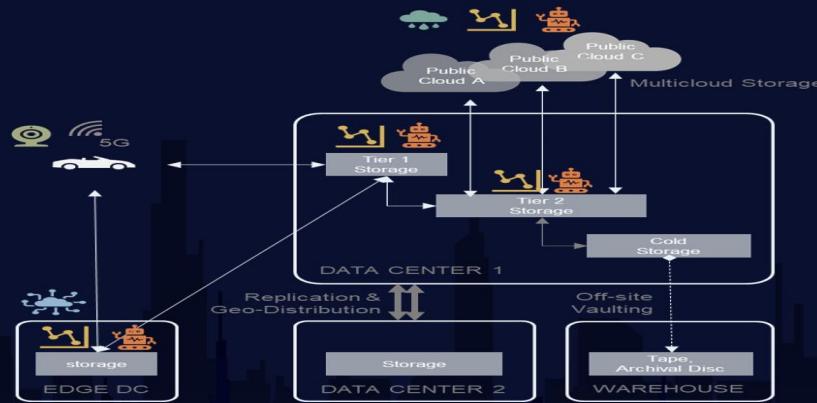
UNSETLED TYPES OF STORAGE

UNPRECEDENTED DATA MOBILITY & SCALE

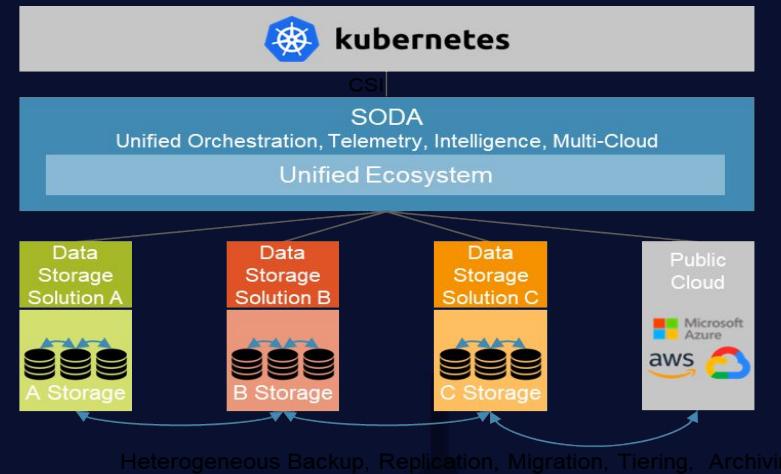
“REALLY” DISTRIBUTED HETEROGENEOUS

SO MANY USE CASES

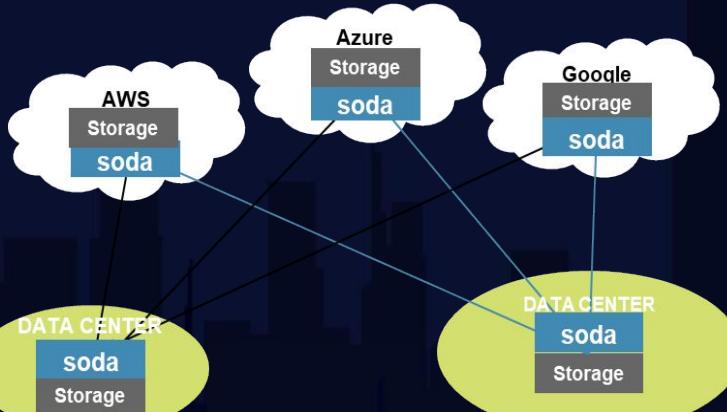
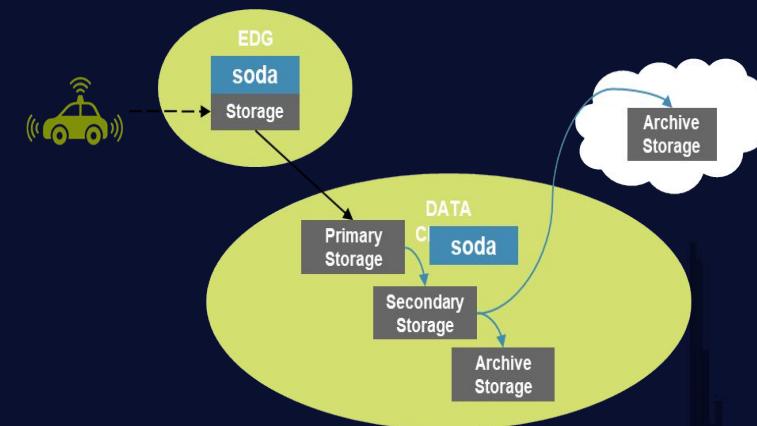
Connected Car (Edge, Core , Multicloud)



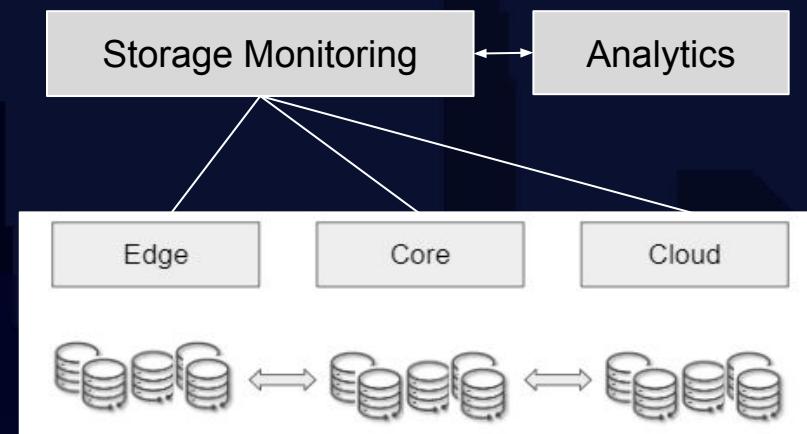
Cloud Native Storage



Data Lifecycle (Edge, Core , Multicloud)

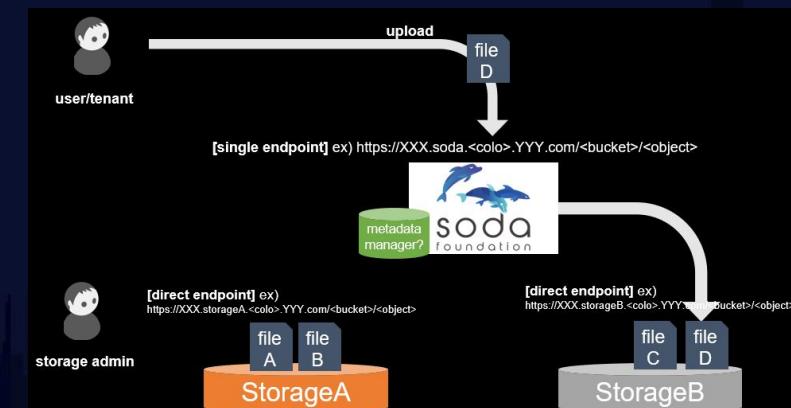


Multicloud Data Management



Heterogeneous Storage Monitoring

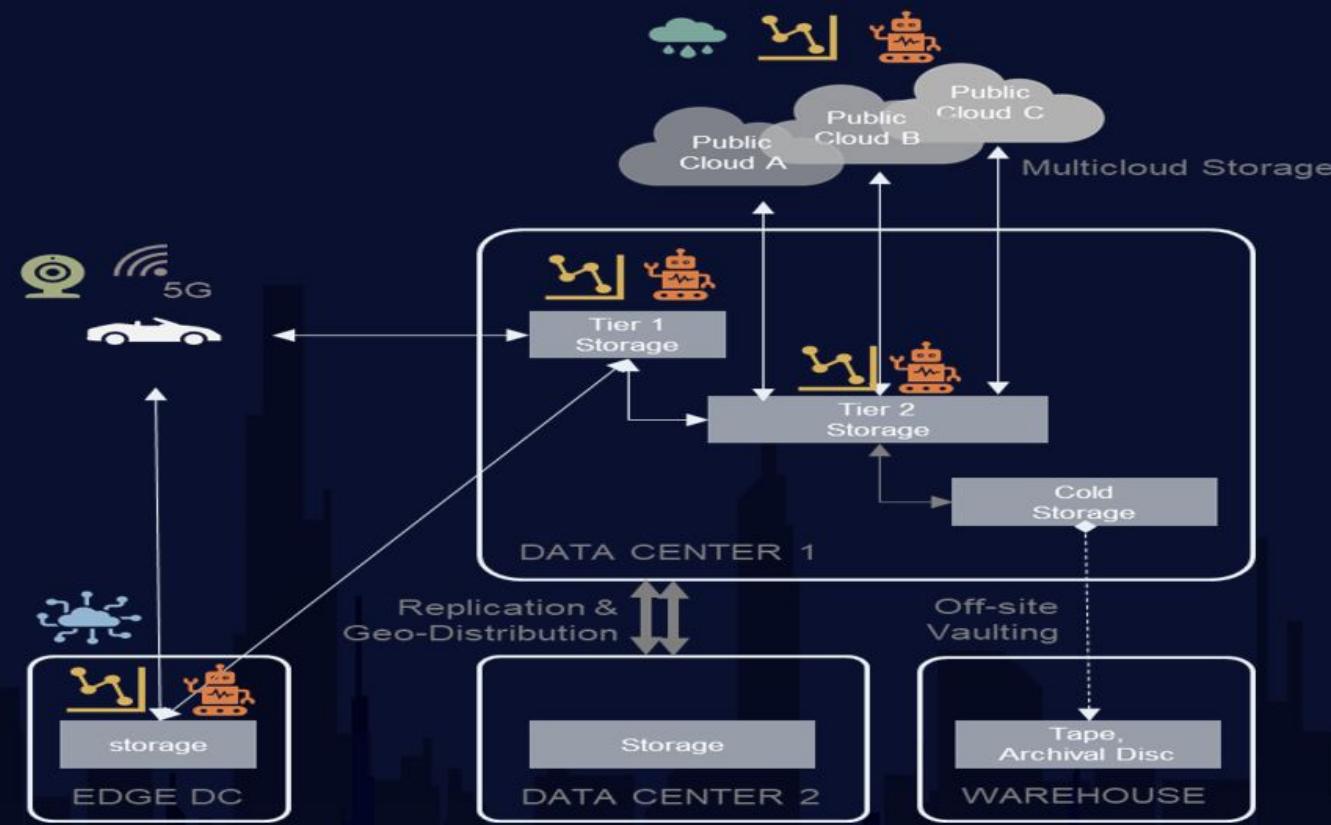
sodacon 2020



Unified Object Storage Management

#sodacon2020

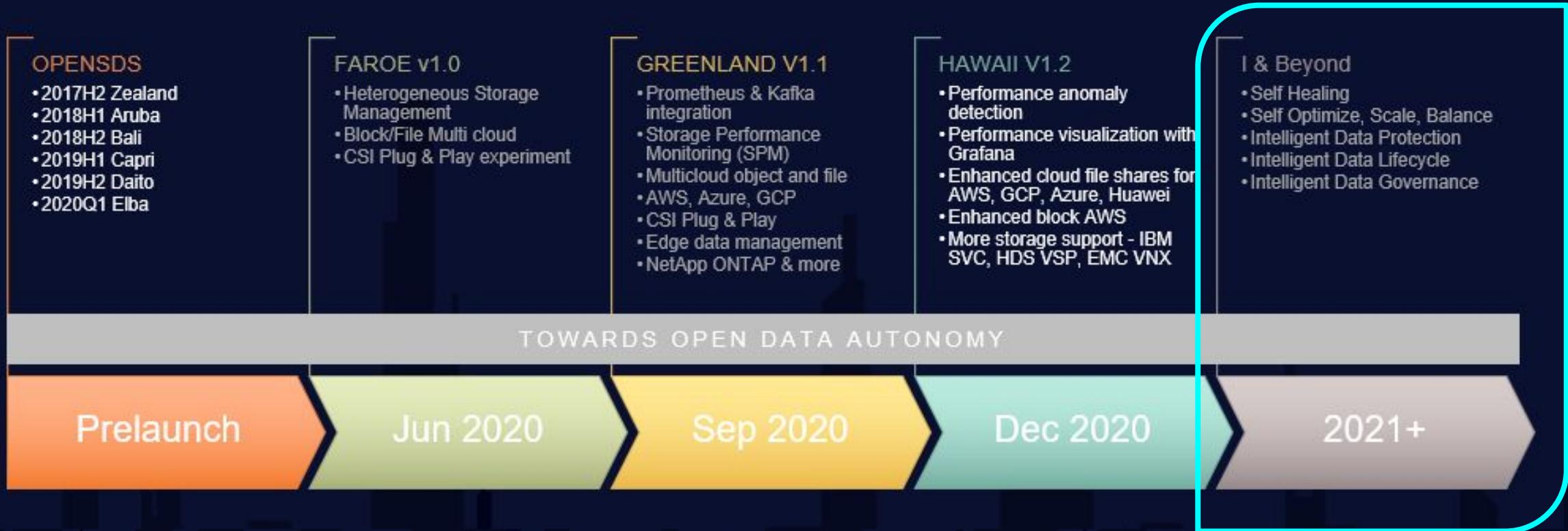
WE WANT INTELLIGENT EDGE, CORE, CLOUD DATA



SEAMLESS CONNECTED DATA ALL ACROSS

ROADMAP FOR FUTURE

OPEN SOURCE | OPEN STANDARD | OPEN DEVELOPMENT



<https://github.com/sodafoundation>

OPEN UNIFIED DATA FRAMEWORK

HETEROGENEOUS | HYBRID DATA

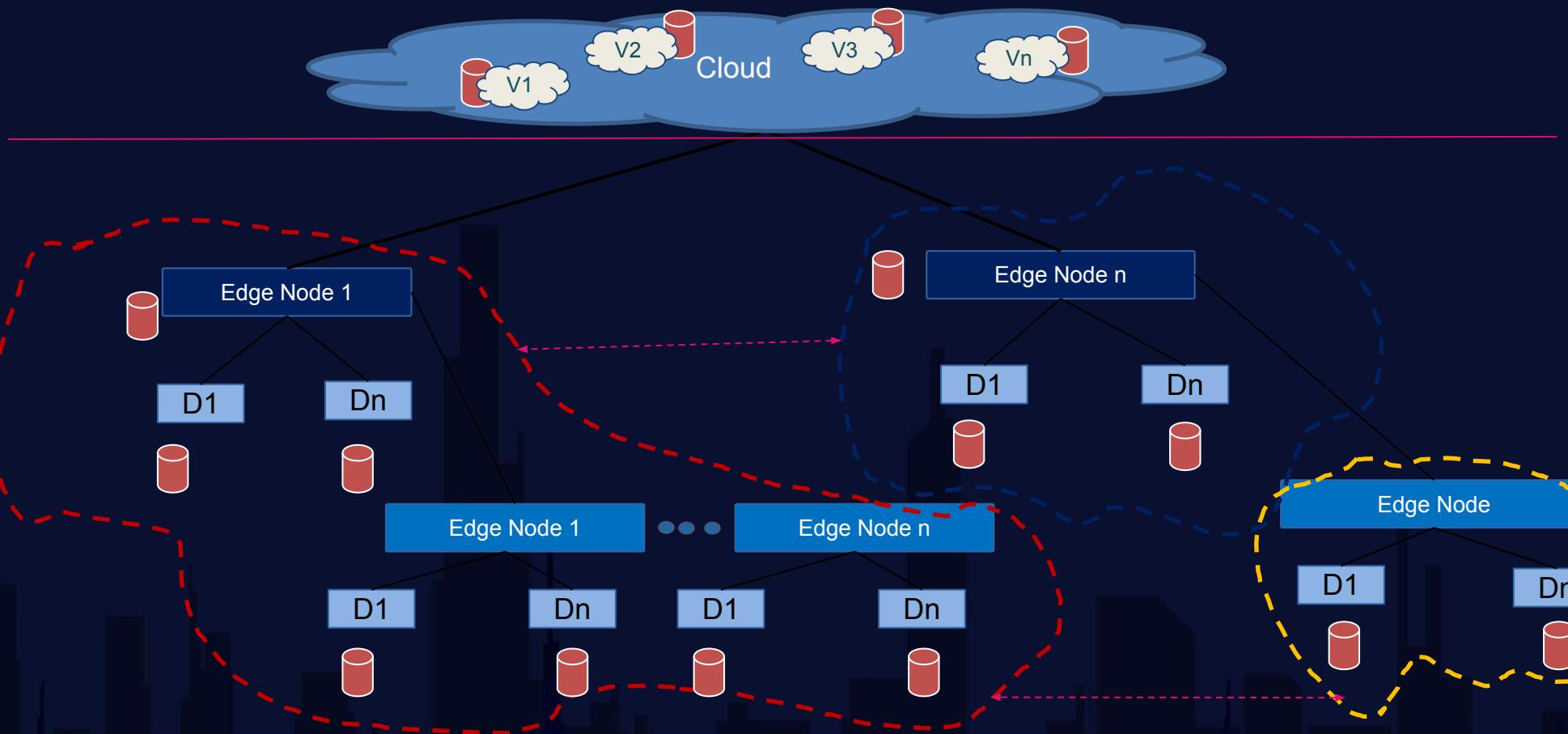


SODA EDGE

sodacon 2020

#sodacon2020

EDGE IS NOT JUST EDGE



IMPOSE NEW CHALLENGES

Multi-Source

Heterogeneous Data

Distributed...Really!

Data Consistency

Data Ownership

Low Latency

Data Mobility

Offline Scenarios

FINDING A WAY

Distributed & Heterogeneous Data Management Platform@Edge

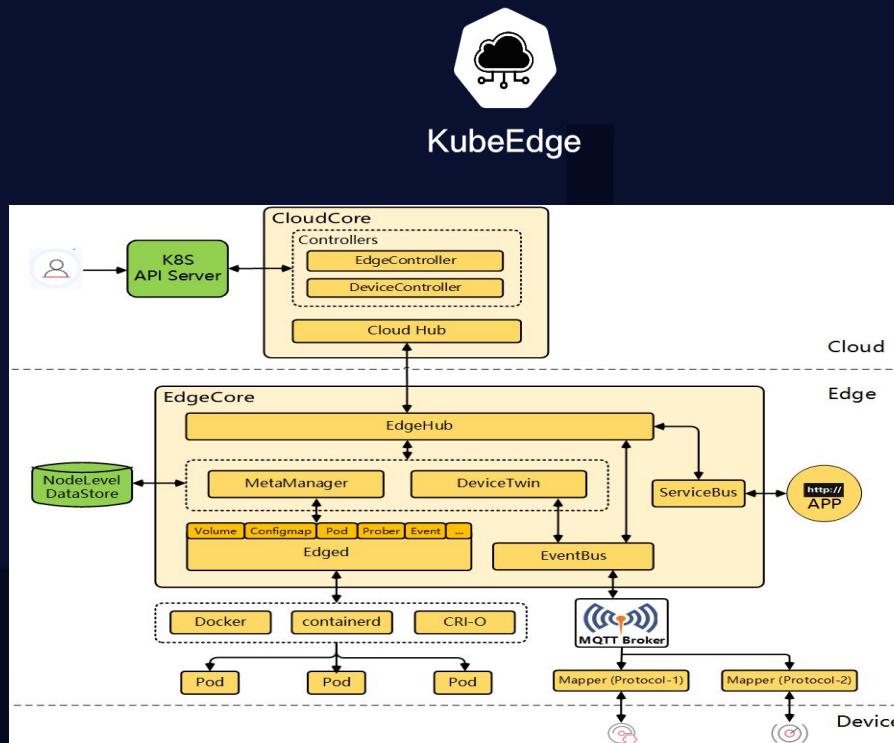
Unified Data Framework
Open
Vendor Agnostic
Platform Agnostic
Distributed
Low Resource
Extensible or Shrinkable
Standardized

Edge Computing

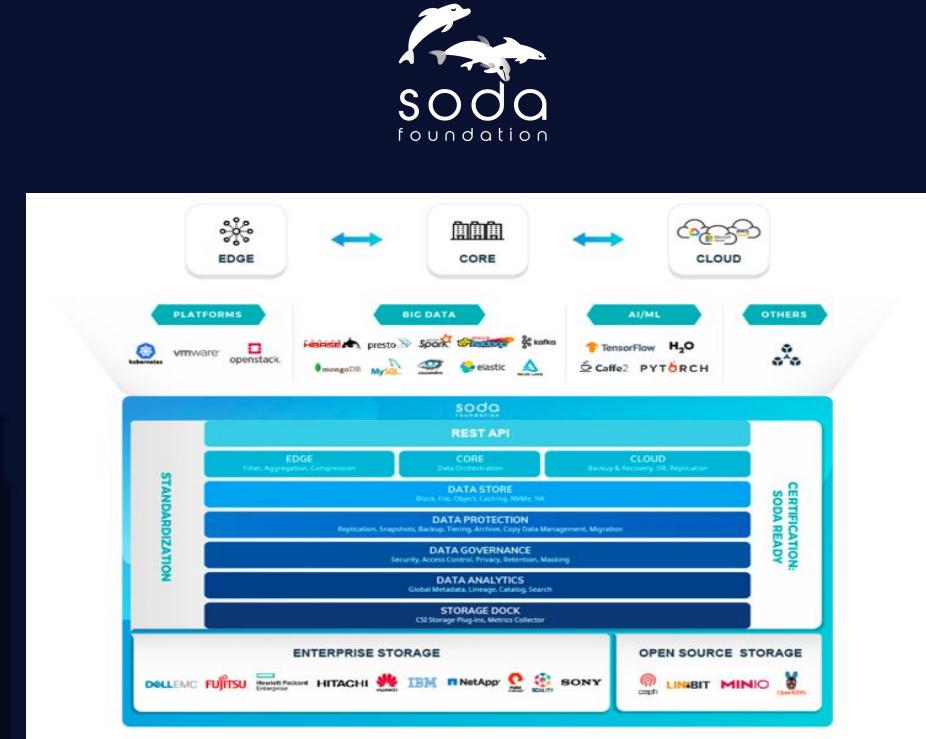


Edge Data Management

STARTING SOMEWHERE



<https://github.com/kubeedge>



<https://github.com/sodafoundation>

CSI BASED KUBE EDGE INTEGRATION



Provides lightweight edge computing platform with compute, network and storage(csi) interfaces



Container Storage Interface



Single SODA CSI plugin which can support all the devices /drivers supported in SODA



SODA Unified Heterogeneous Data Interface



Heterogeneous Storages : different vendors, models, types, cloud storages...

SODA CSI

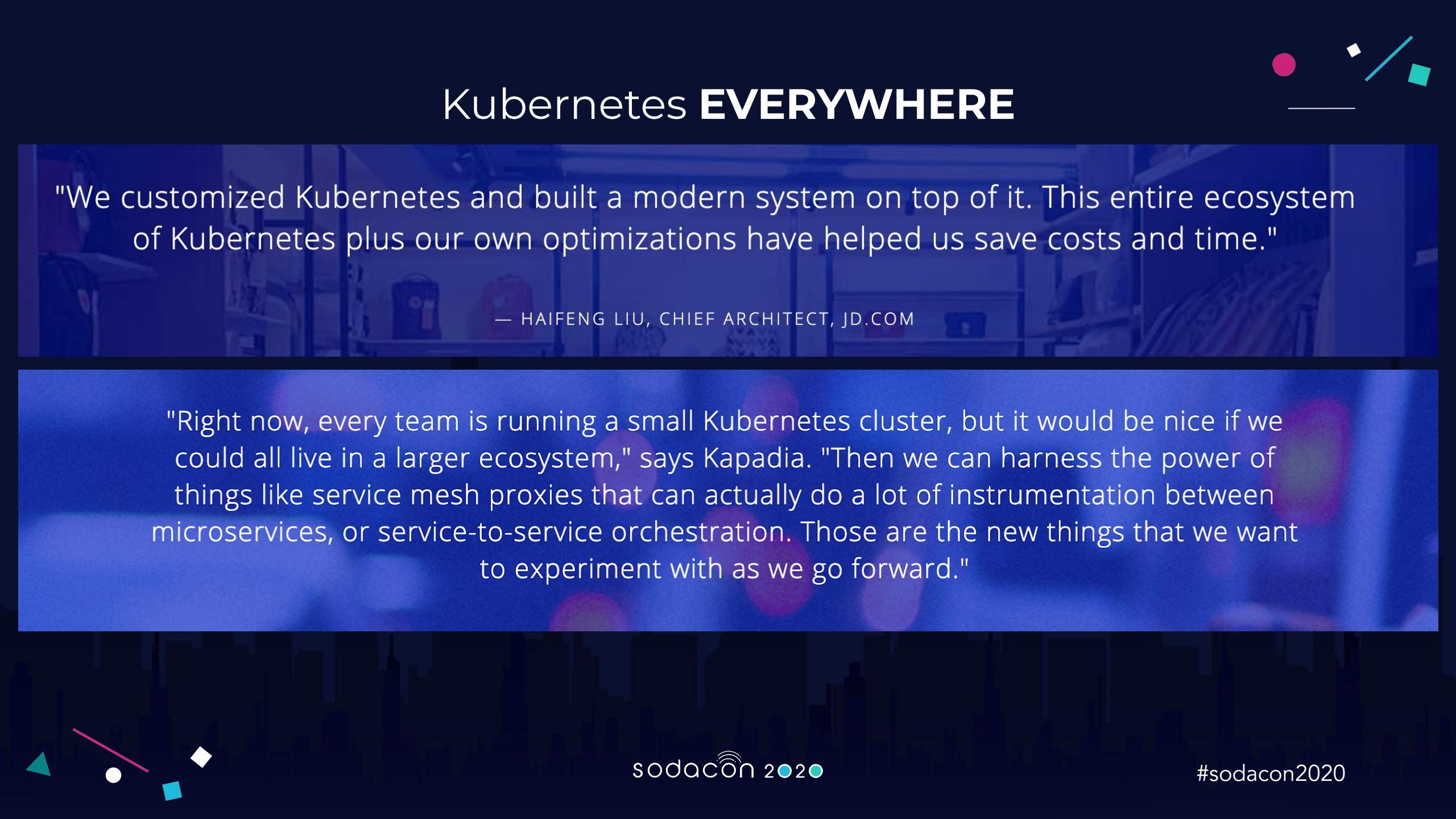
sodacon 2020

#sodacon2020

Kubernetes **EVERWHERE**



Kubernetes **EVERWHERE**



"We customized Kubernetes and built a modern system on top of it. This entire ecosystem of Kubernetes plus our own optimizations have helped us save costs and time."

— HAIFENG LIU, CHIEF ARCHITECT, JD.COM

"Right now, every team is running a small Kubernetes cluster, but it would be nice if we could all live in a larger ecosystem," says Kapadia. "Then we can harness the power of things like service mesh proxies that can actually do a lot of instrumentation between microservices, or service-to-service orchestration. Those are the new things that we want to experiment with as we go forward."

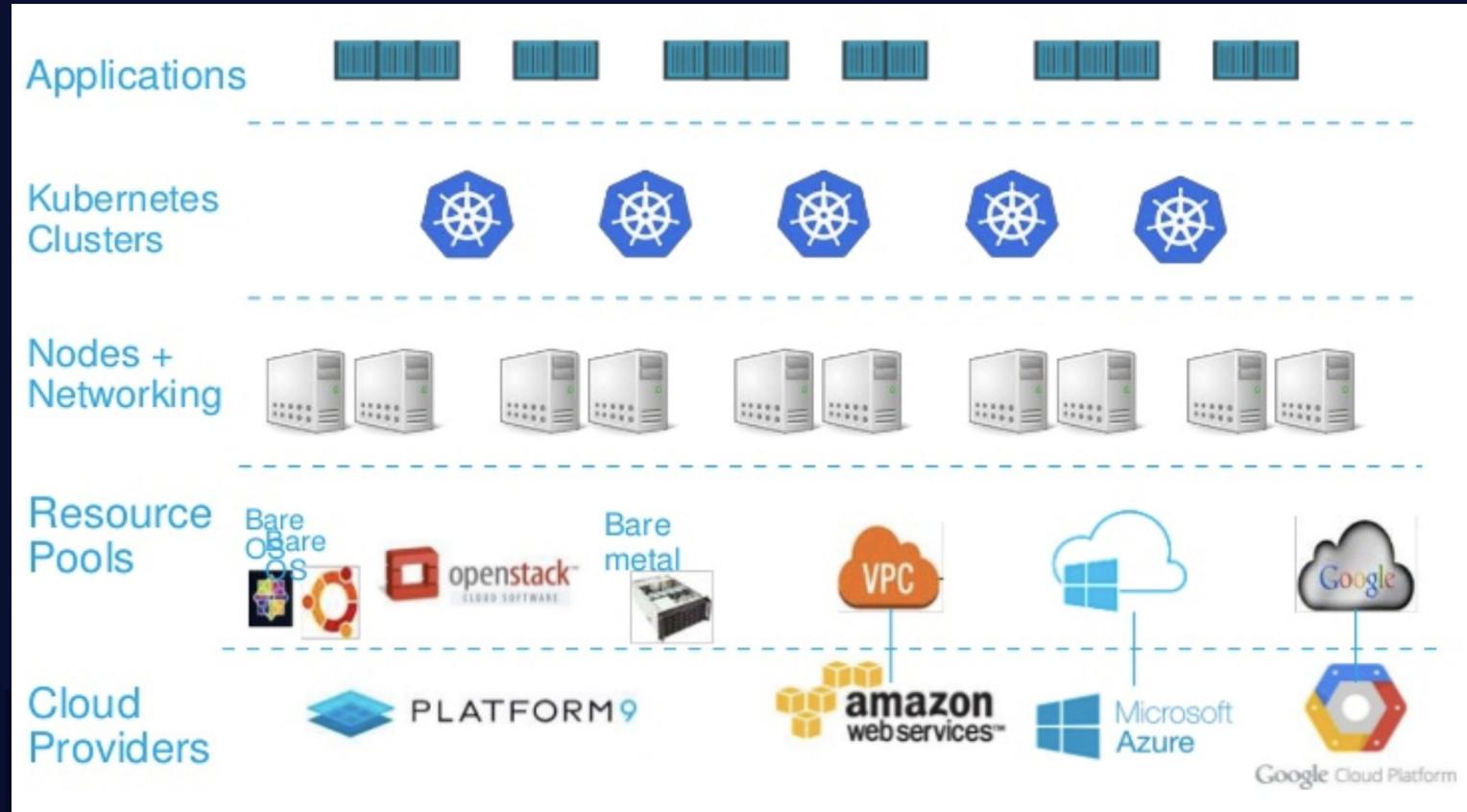
Kubernetes **EVERYWHERE**



"We want to provide the tools in the same ecosystem, in a consistent way, rather than have a large custom snowflake ecosystem where every tool needs its own custom deployment. Kubernetes gives us the ability to bring all of these together, so the richness of the open source and even the license community dealing with big data can be corralled."

With Kubernetes, "a team can come to us and we can have them up and running with a basic decisioning app in a fortnight, which before would have taken a whole quarter, if not longer. Kubernetes is a manifold productivity multiplier."

Kubernetes **EVERWHERE**



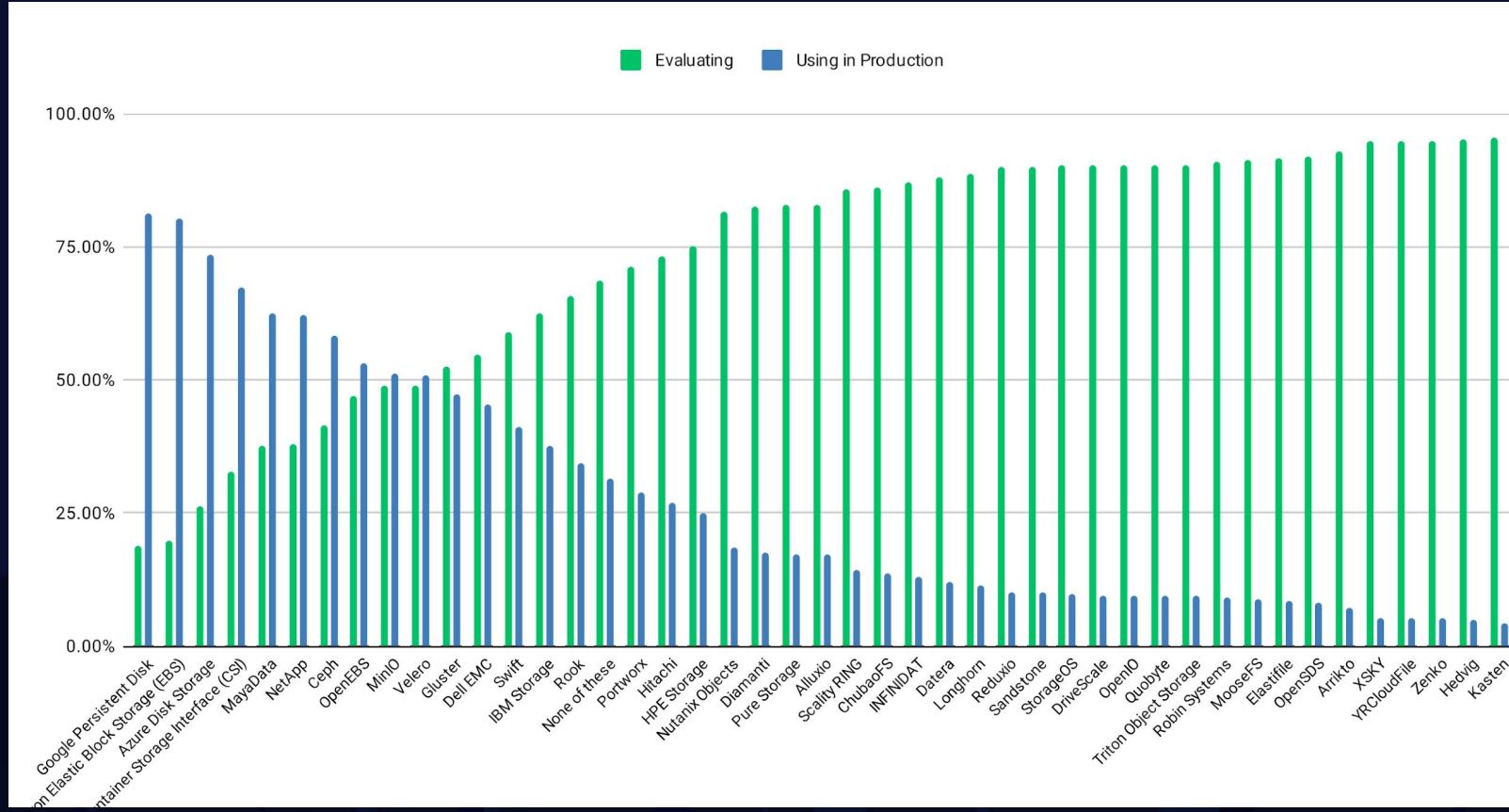
- #1: Fault-tolerant by Design. Self-healing
- #2: Modular App Design w/ Containers
- #3: Automated Rolling Upgrades
- #4: Auto-scaling. Reproducibility.
- #5: Write Once, Run Anywhere. Portability.

(K8s is the new Linux)

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

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

DATA on Kubernetes

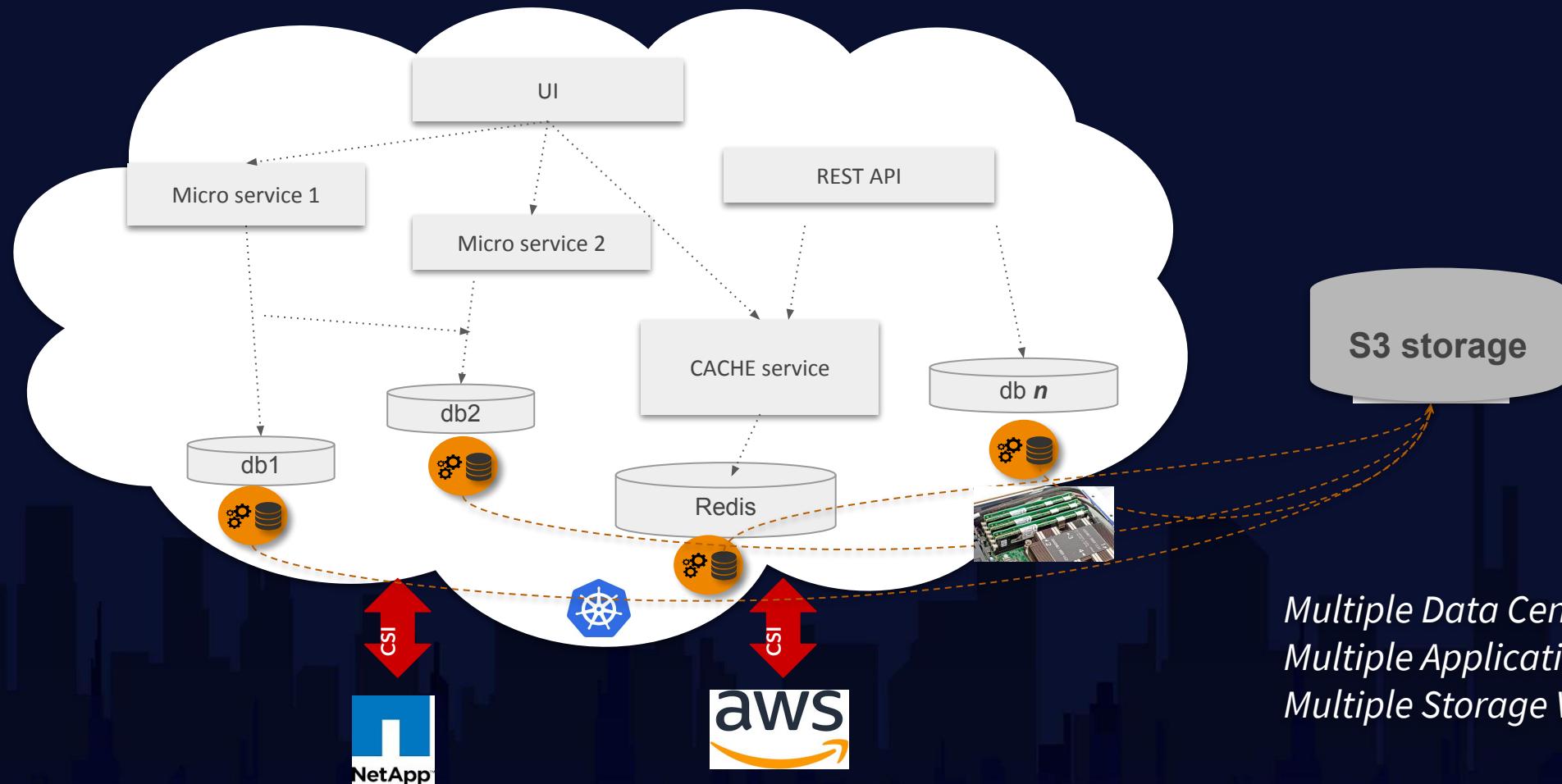


https://www.cncf.io/wp-content/uploads/2020/11/CNCF_Survey_Report_2020.pdf

Containers were originally built to be stateless in order to keep them flexible and portable. However, only 22% use only stateless applications. Demonstrating the popularity of cloud native storage, **55% use stateful applications in production**, 12% are evaluating them, and 11% plan to use them in the next 12 months.

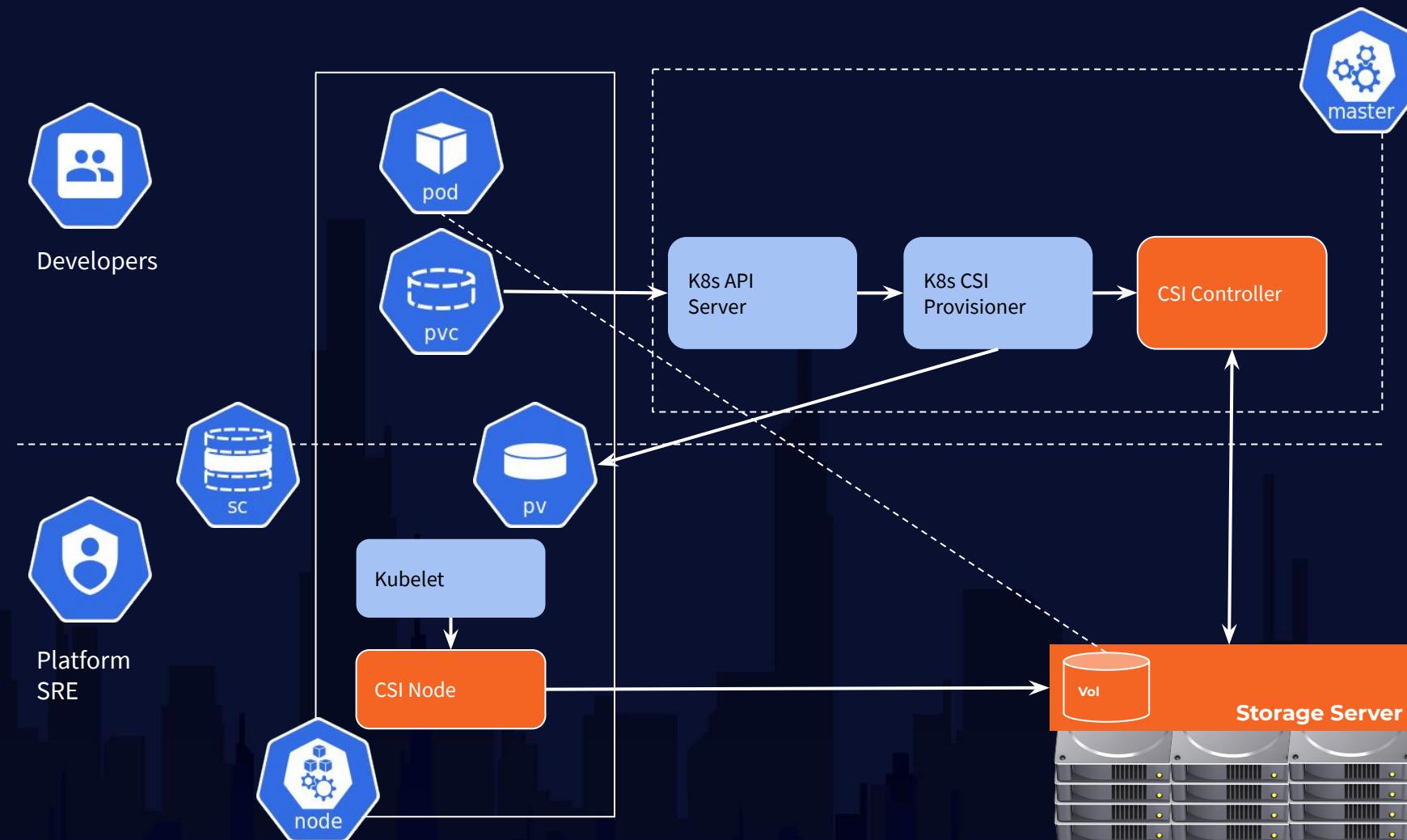


DATA on Kubernetes with CSI

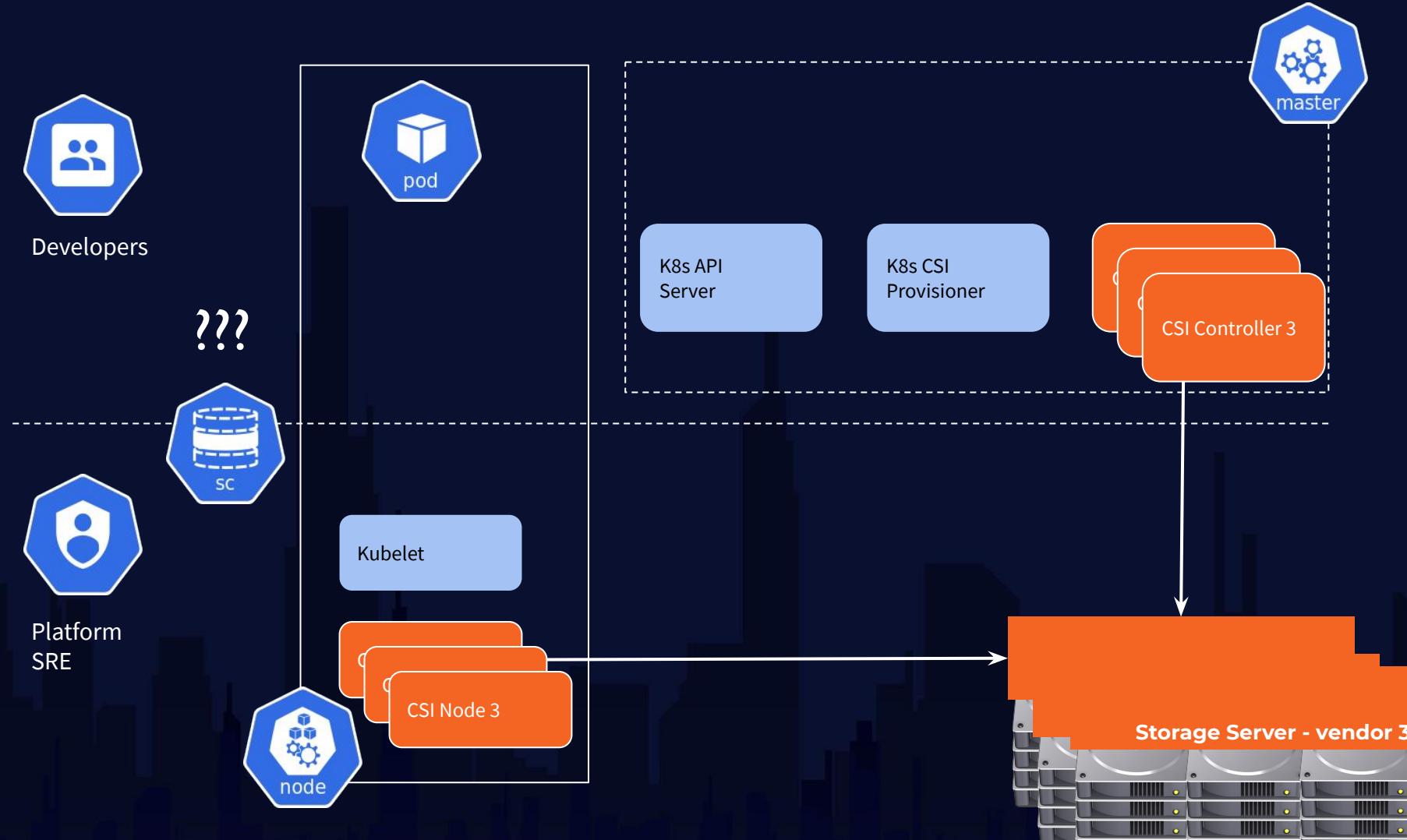


*Multiple Data Centers
Multiple Applications
Multiple Storage Vendors*

CSI : How does it work?



CSI w/ Heterogeneous Storage



SODA CSI: Beyond CSI

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 PROTECTION

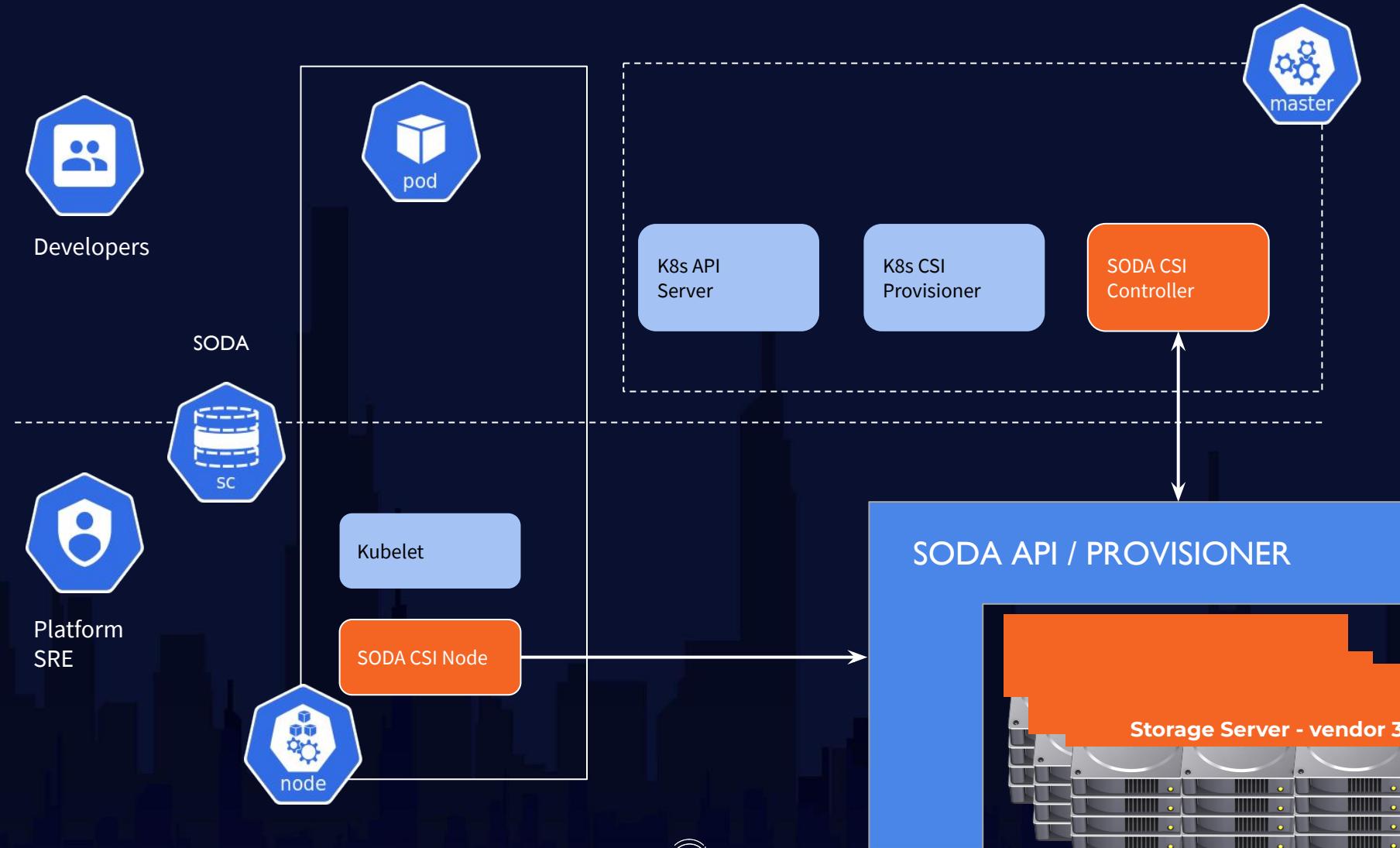
DATA MIGRATION

DATA GOVERNANCE

DATA COMPLIANCE

UNIFIED STORAGE PLATFORM

SODA CSI Plug n Play: An approach

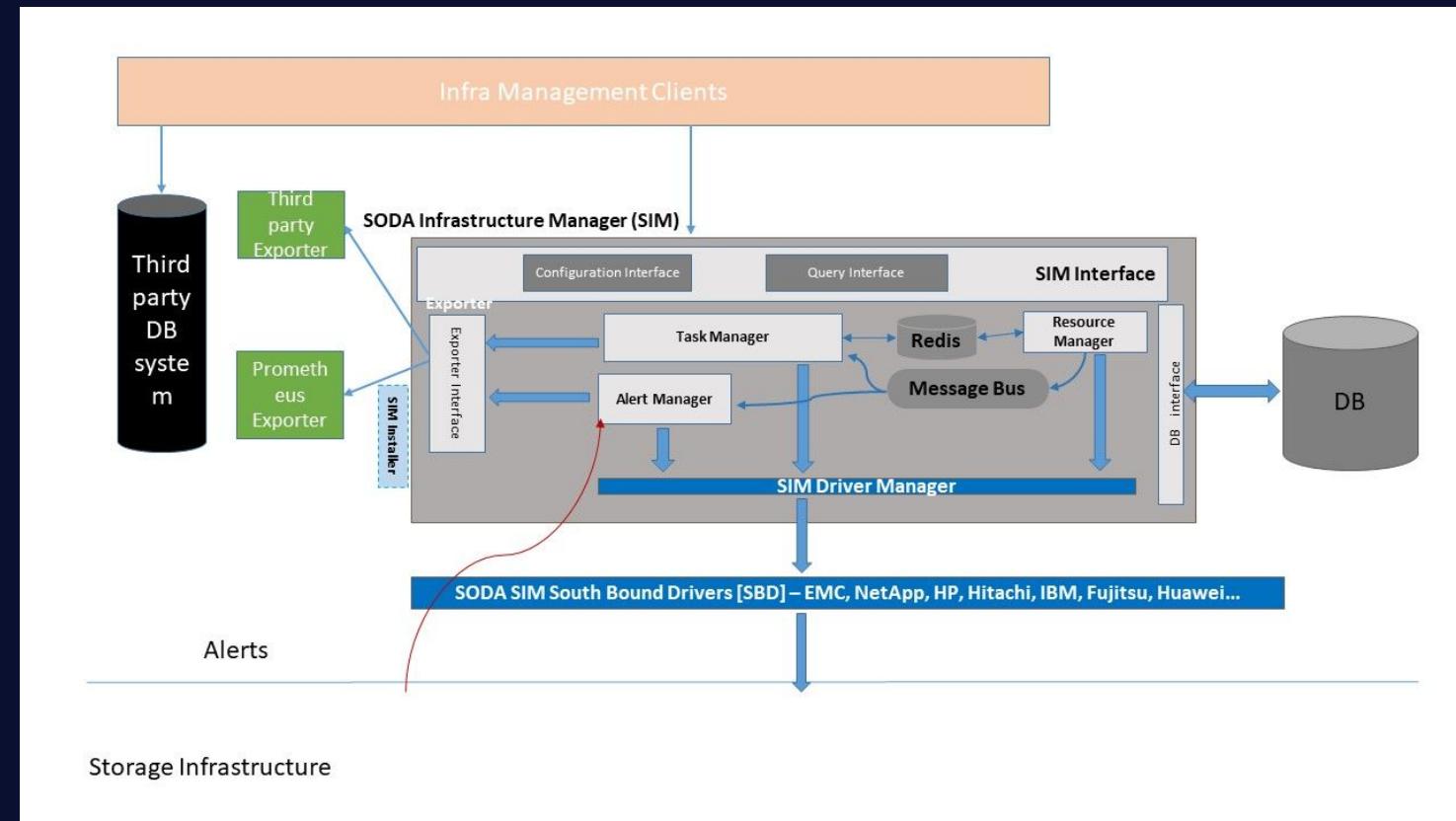
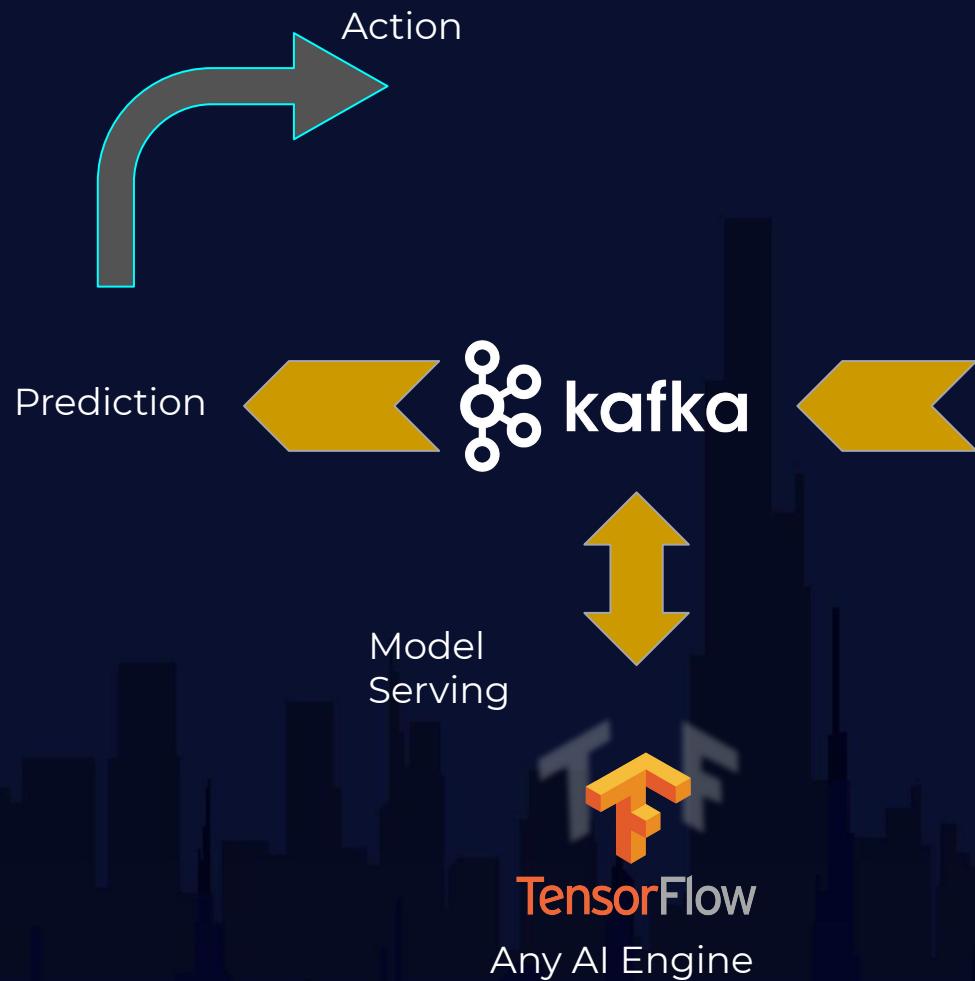


SODA SPM (DELFIN)

sodacon 2020

#sodacon2020

INTELLIGENT HETEROGENEOUS STORAGE MONITORING



OUR TECHNICAL EFFORTS

- SODA EDGE
- SODA MULTICLOUD
- DRIVER DOCK
- SODA CSI
- UNIFIED API & MODELS
- HYBRID DR
- SODA SPM
- DATA PROJECTS
- SODA EXPERIENCE

~ADD YOURS....

..and **Join Us**



SODA **Github:**

<https://github.com/sodafoundation>

SODA **Slack:**

<https://sodafoundation.io/slack/>

THANK YOU



soda foundation

<https://sodafoundation.io/>

SODA **Source Code**: <https://github.com/sodafoundation>

SODA **Docs**: <https://docs.sodafoundation.io/>

Join SODA **Slack**: <https://sodafoundation.io/slack/>

Follow SODA **Twitter**: <https://twitter.com/sodafoundation>

Join Us: <https://sodafoundation.io/join/>

sodacon 2020

#sodacon2020

Disclaimer : There are some information/images used from internet. Any copyrights or such ownership, belong to the original authors/orgs. The contents are used in this presentation is only for sharing and educational purpose. If you want to use it otherwise, please ensure the original license of the contents.