

Databases on Kubernetes:

A Storage Story

Databases



Kubernetes

Chris Engelbert

Devrel @ simplyblock

Previous fun companies:

- Ubisoft / Blue Byte
- Hazelcast
- Instana
- clevabit
- Timescale

Interests:

- Developer Relations
- Anything Performance Engineering
- Backend Technologies
- Fairy Tales (AMD, Intel, Nvidia)

 @noctarius2k

 @noctarius2k@mastodon.online

 @noctarius.com



Why Databases in Kubernetes?

Why not to run a database in Kubernetes?



Christoph Engelbert / Noctarius ツ / エンゲルベルト クリス 
@noctarius2k

Why you ***SHOULD NOT*** run a database in Kubernetes?
What do you think? Please help me, I need you! 🙏❤️

via: [#postgresql](#) [#mysql](#) [#mariadb](#) [#kafka](#) [#kubernetes](#)

Why not to run a database in Kubernetes?



Christoph Engelbert / Noctarius ツ / エンゲルベルト クリス 
@noctarius2k

Why you ***SHOULD NOT*** run a database in Kubernetes?
What do you think? Please help me, I need you! 🙏❤️

PS: Asking for a friend! 😅

via: [#postgresql](#) [#mysql](#) [#mariadb](#) [#kafka](#) [#kubernetes](#)

Why not to run a database in Kubernetes?

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

Nobody understands Kubernetes!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

Nobody understands Kubernetes!

What's the benefit; databases don't need autoscaling!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

Nobody understands Kubernetes!

What's the benefit; databases don't need autoscaling!

Databases and applications should be separated!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

Nobody understands Kubernetes!

What's the benefit; databases don't need autoscaling!

Databases and applications should be separated!

Not another layer of indirection / abstraction!

Why not to run a database in Kubernetes?

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

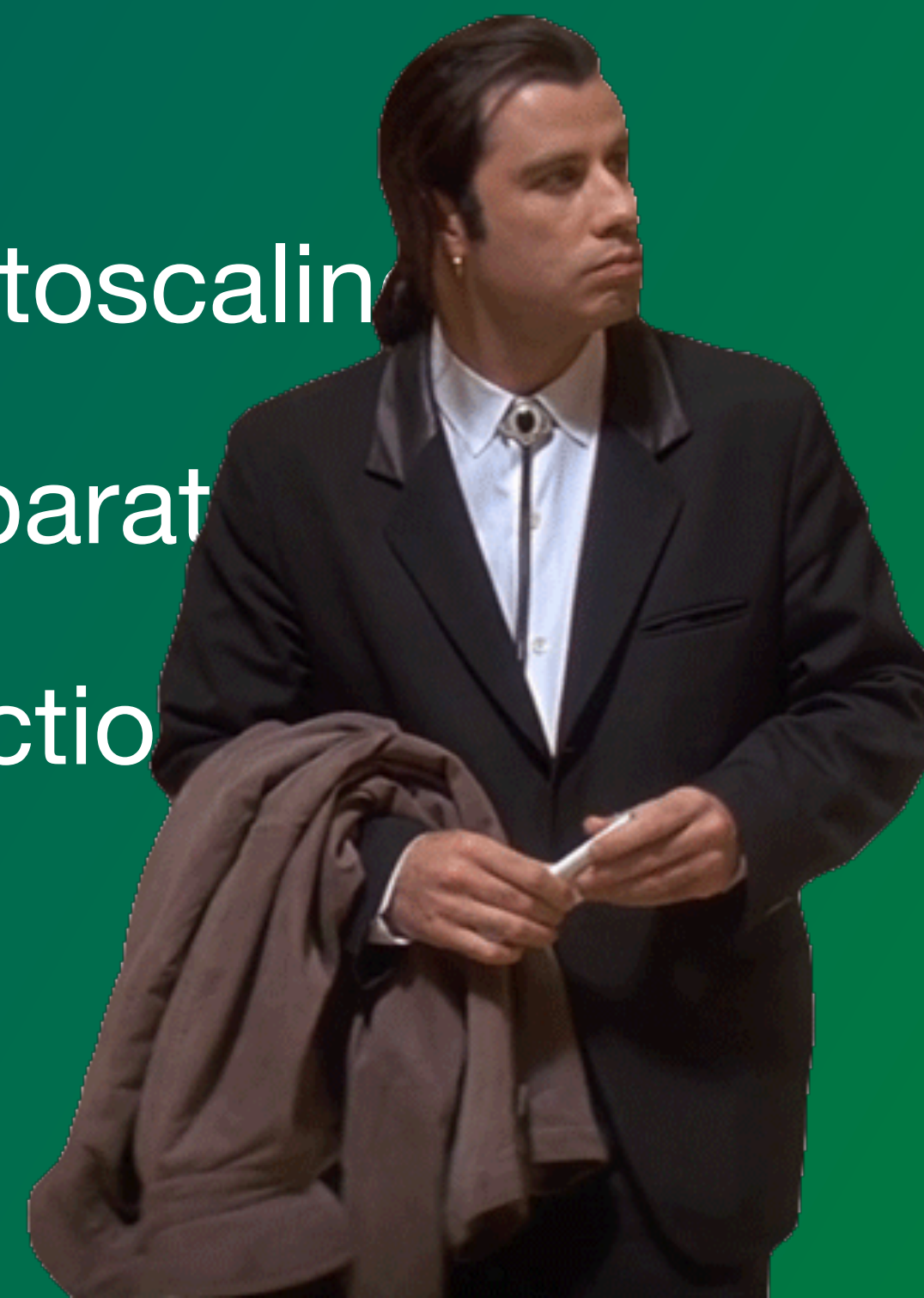
Persistent Data will kill you! Too slow!

Nobody understands Kubernetes!

What's the benefit; databases don't need autoscaling

Databases and applications should be separated

Not another layer of indirection / abstraction



You Already Run A Database In Kubernetes

🟢 You already run a database in Kubernetes



Why?



Why?

No Cloud-Vendor Lock-In



Why?

No Cloud-Vendor Lock-In
Faster Time To Market



Why?

No Cloud-Vendor Lock-In

Faster Time To Market

Decreasing cost



Why?

No Cloud-Vendor Lock-In

Faster Time To Market

Decreasing cost

Automation



Why?

No Cloud-Vendor Lock-In

Faster Time To Market

Decreasing cost

Automation

Unified deployment architecture



Why?

No Cloud-Vendor Lock-In

Faster Time To Market

Decreasing cost

Automation

Unified deployment architecture

Need read-only replicas



Typical Challenges

Challenges with Databases on Kubernetes

Challenges with Databases on Kubernetes

Selecting The Right Database

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Backups and Restores

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Backups and Restores

Access Control and Network Policies

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Backups and Restores

Access Control and Network Policies

Container Images and Extensions

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Backups and Restores

Access Control and Network Policies

Container Images and Extensions

Monitoring and Observability

Challenges with Databases on Kubernetes

Selecting The Right Database

Security Requirements (Data-At-Rest Encryption)

Backups and Restores

Access Control and Network Policies

Container Images and Extensions

Monitoring and Observability

Storage

Storage 🙄

The Biggest Issue: Storage

K8s is not designed with Databases in mind!

Never run Stateful Workloads in k8s!

Persistent Data will kill you! Too slow!

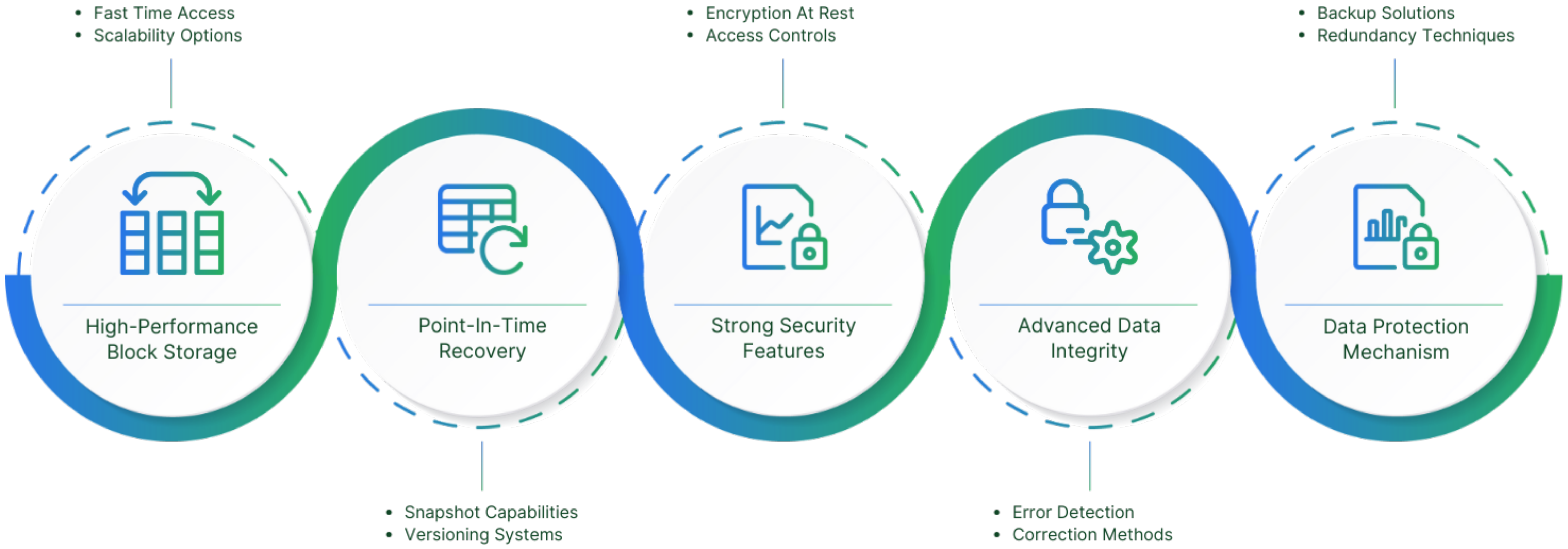
Nobody understands Kubernetes!

What's the benefit; databases don't need autoscaling!

Databases and applications should be separated!

Not another layer of indirection / abstraction!

The Biggest Issue: Storage



Database Storage Requirements





Should be dynamically provisioned



Should be dynamically provisioned
CSI provider enables encryption at rest



Should be dynamically provisioned
CSI provider enables encryption at rest
High IOPS (SSD or NVMe)



Should be dynamically provisioned
CSI provider enables encryption at rest
High IOPS (SSD or NVMe)
Low Latency



Should be dynamically provisioned

CSI provider enables encryption at rest

High IOPS (SSD or NVMe)

Low Latency

Database performance is as fast as your storage



Should be dynamically provisioned

CSI provider enables encryption at rest

High IOPS (SSD or NVMe)

Low Latency

Database performance is as fast as your storage

I'd recommend a disaggregated storage!

Kubernetes CSI

CSI = Container Storage Interface

CSI = Container Storage Interface

```
interface CSIDriver {  
  
    ListVolumes ()  
    ProvisionVolume ()  
    DeleteVolume ()  
    AttachVolume ()  
    ExpandVolume ()  
    SnapshotVolume ()  
    DeleteSnapshot ()  
    ...  
  
}
```

CSI = Container Storage Interface

CSI = Container Storage Interface

Avoids in-tree volume drivers

CSI = Container Storage Interface

Avoids in-tree volume drivers

Independent release cycles

CSI = Container Storage Interface

Avoids in-tree volume drivers

Independent release cycles

Works across Kubernetes versions

CSI = Container Storage Interface

Avoids in-tree volume drivers

Independent release cycles

Works across Kubernetes versions

Enables third-party volume drivers

CSI Driver Listing

Searchable listing of Kubernetes volume providers (Kubernetes CSI). Find your perfect implementation based on features, access modes, and more.

The following table contains a list of available Kubernetes container storage interface (CSI) implementations. All CSI drivers are listed with their respective capabilities. The list is updated bit by bit according to the documentation on a best effort basis. The list may not be complete and properties may be incorrect or incomplete. Vendors of the CSI drivers are welcome to validate and update their respective properties.

SearchTerm:


Lifecycle Modes:

Capabilities:

Access Modes:

Filter Deprecated:

Filter result: 136/150

Driver Name	Driver Class	Description	Lifecycle Modes	Supported Capabilities	Access Modes
 AlibabaCloud Disk	diskplugin.csi.alibabacloud.com	CSI Driver for an AlibabaCloud Disk	<ul style="list-style-type: none">• Persistent: ✓• Ephemeral: ✗	<ul style="list-style-type: none">• Block Storage: ✓• File Storage: ✗• Object Storage: ✗• Dynamic: ✓• Snapshot: ✓• Clones: ✗• Expansion: ✓• Topology: ✓• Tracking: ✗	<ul style="list-style-type: none">• Read Only Many: ✓• Read Write Once: ✓• Read Write Many: ✗• Read Write Once Pod: ✓

www.storageclass.info/csdrivers

Demo

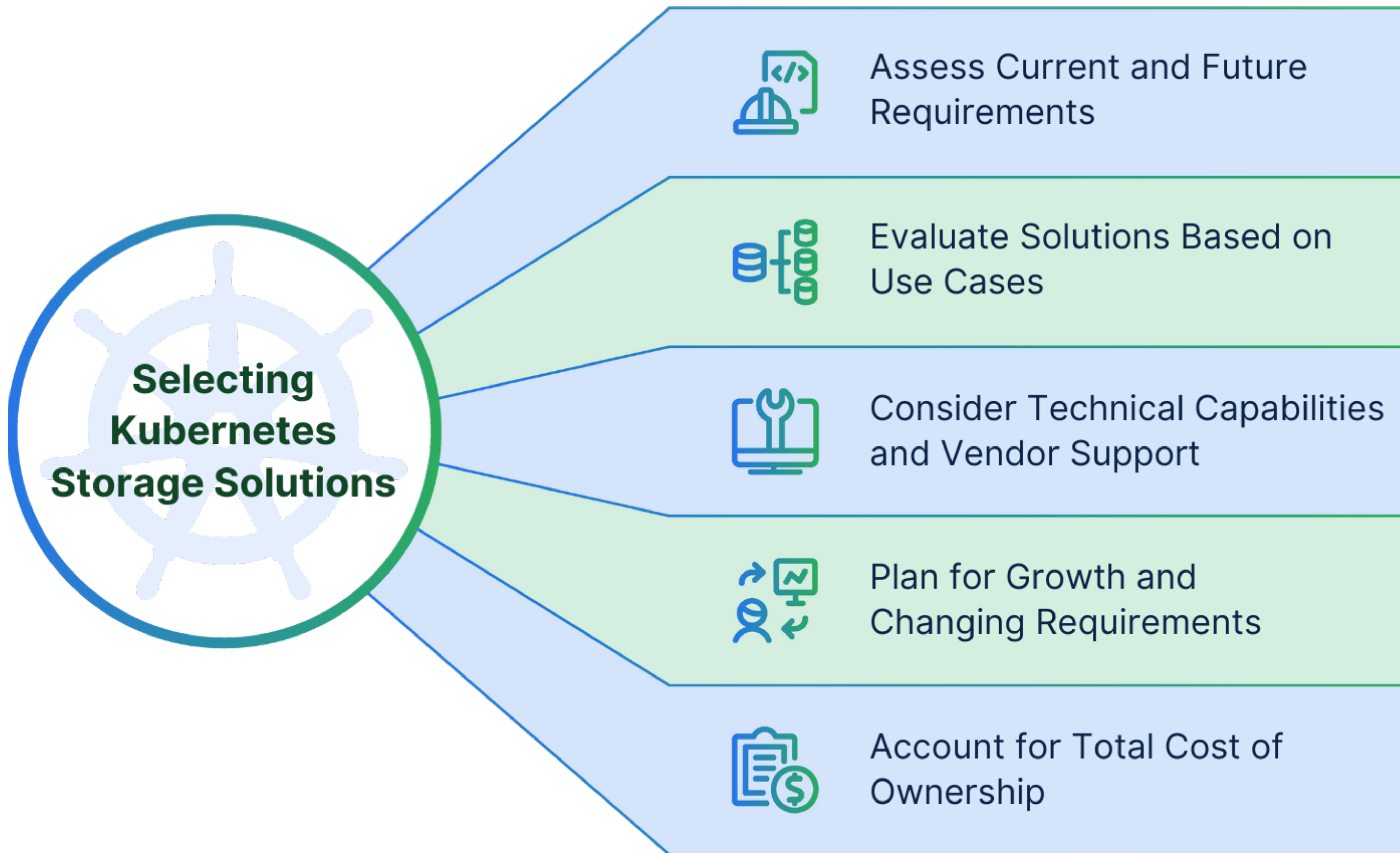
Intelligent Cloud-Native Storage Platform

Intelligent Cloud-Native Storage Platform

- Copy-on-write
- Instant Snapshots
- Instant Clones / Branches / Forks (whatever you want to call it)
- NVMe over Fabrics (NVMe/TCP)
- Optimized for Database and I/O-heavy Workloads
- Disaster Recovery
- A Block Storage Device (basically a hard disk)

How to Choose?

How to Choose?



Copy-On-Write: The Hero of Storage



Nikita Shamgunov · Following

Entrepreneur (@neondatabase, @singlestore). Investor (@khoslaventures)
2d · Edited · 🌐

14.5 TB storage for the price of 2.9 TB. That's the power of copy-on-write.

(screenshot is of a logistics/IoT customer with a very common **Neon** Branching setup)

Branches (5)

Branch id, name: Project id: Timeline id: Show deleted

Id	Name	Project	Parent	Logical size	Created at	Updated at	
<input type="text" value=""/>	dev	<input type="text" value=""/>	<input type="text" value=""/>	2.9 TB	2025-02-11T16:23:31.390Z	2025-03-04T01:30:36.424Z	operations
<input type="text" value=""/>	dev	<input type="text" value=""/>	<input type="text" value=""/>	2.9 TB	2025-02-12T18:20:50.668Z	2025-03-06T17:19:40.078Z	operations
<input type="text" value=""/>	dev	<input type="text" value=""/>	<input type="text" value=""/>	2.9 TB	2025-03-06T17:06:56.131Z	2025-03-11T03:15:58.496Z	operations
<input type="text" value=""/>	dev	<input type="text" value=""/>	<input type="text" value=""/>	2.9 TB	2025-02-12T17:53:31.784Z	2025-03-08T18:29:40.092Z	operations
<input type="text" value=""/>	main	<input type="text" value=""/>	-	2.9 TB	2025-01-29T19:49:57.981Z	2025-03-11T09:12:15.826Z	operations

Trust me, I'm Kelsey!



Kelsey Hightower

@kelseyhightower



Soham Dasgupta @thesobercoder · Feb 10, 2023

@kelseyhightower Bust a myth for us please - running any sort of database on a Kubernetes instance is bad idea. I've heard this enough times to actually start believing it. #kubernetes #mythbuster

5:21 PM · Feb 10, 2023 · **326.7K** Views



<https://x.com/kelseyhightower/status/1624081136073994240>

Trust me, I'm Kelsey!



Kelsey Hightower

@kelseyhightower



You can run databases on Kubernetes because it's fundamentally the same as running a database on a VM. The biggest challenge is understanding that rubbing Kubernetes on Postgres won't turn it into Cloud SQL. 🧵



Soham Dasgupta @thesobercoder · Feb 10, 2023

@kelseyhightower Bust a myth for us please - running any sort of database on a Kubernetes instance is bad idea. I've heard this enough times to actually start believing it. #kubernetes #mythbuster

5:21 PM · Feb 10, 2023 · **326.7K** Views



43



149



839



224



<https://x.com/kelseyhightower/status/1624081136073994240>

Data on Kubernetes Community: <https://dok.community>

[Data on Kubernetes Whitepaper](#)

 @noctarius2k

 @noctarius2k@mastodon.online

 @noctarius.com

Thank you very much!
Questions?

