

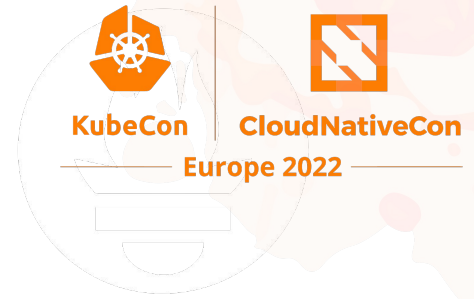
Transparent Live Migration of Services Between Kubernetes Cluster Across Multi-Cloud

Adam Janikowski & Jörg Schad, ArangoDB



tl;dr

- Moving a stateful Service between different K8s cluster, regions, cloud providers is challenging
 - But very useful from an operational perspective
 - Ideally 'transparent' for user
- Spoiled for Choice...
 - Cooperation of K8s, operator, networking, and database
- Demo



PromCon
North America 2021





KubeCon



CloudNativeCon

Europe 2022



Adam Janikowski

Cloud Architect in ArangoDB

Gdańsk, Pomorskie, Poland · [Contact info](#)



ArangoDB



Politechnika Gdańska

Additionally

- **Operator Architect & Developer**
- **Kubernetes Expert**
- **GoLang Developer**
- **GitHub: ajanikow**



Jörg Schad

CTO at ArangoDB

San Francisco, California, United States · [Contact info](#)



ArangoDB

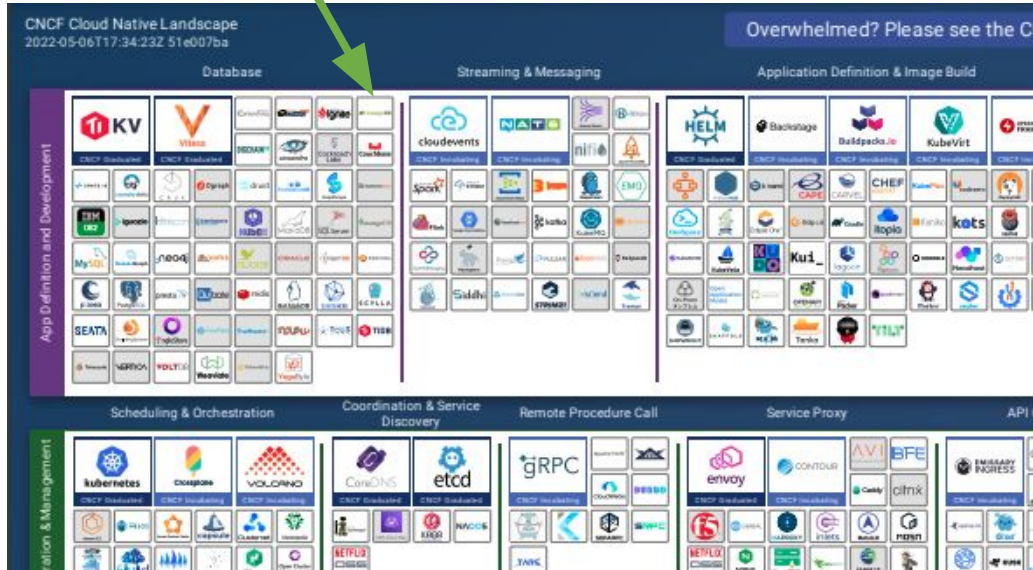
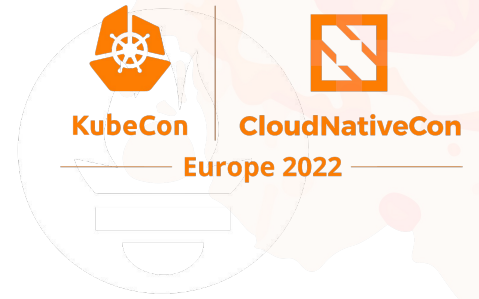


Universität des Saarlandes

Prev

- **Suki.ai**
- **Mesosphere**
- **Architect @SAP Hana**
- **PhD Distributed DB Systems**
- **Twitter: @joerg_schad**

Databases in 2022



PromCon

ArangoDB



kubernetes

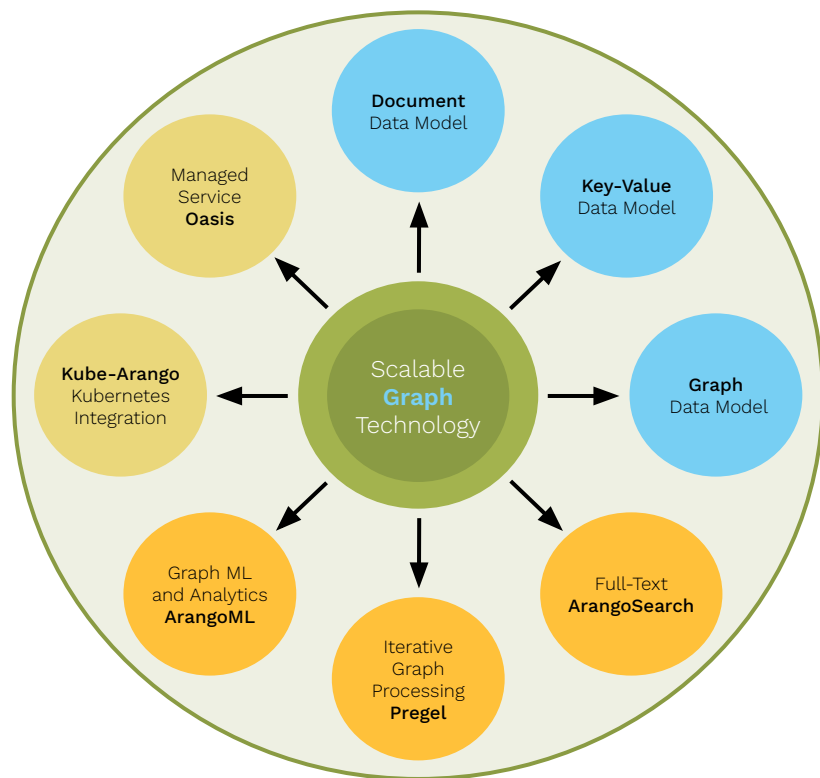


Azure

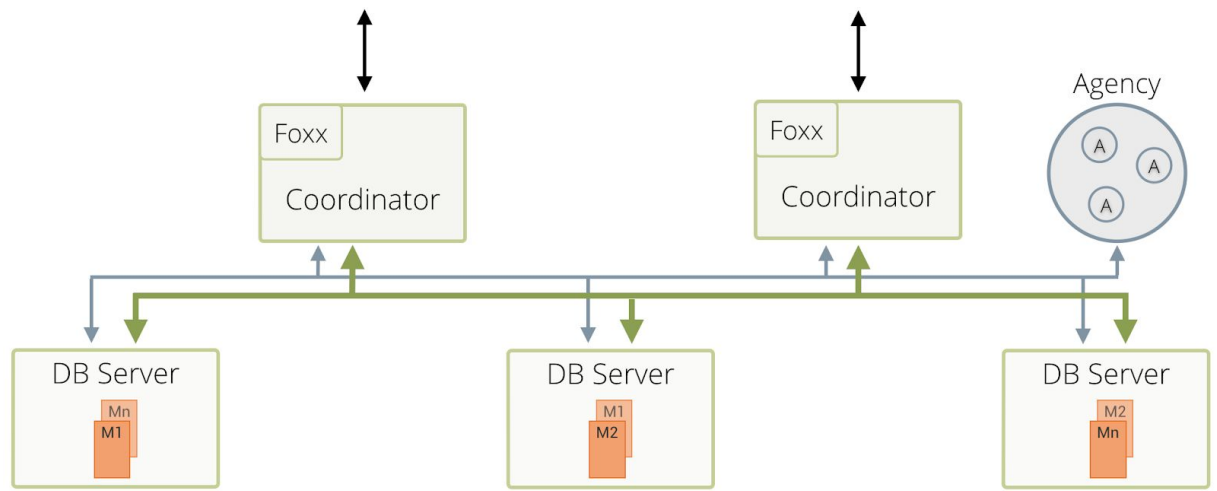


Distributed Stateful Systems on Dynamic Infrastructure!

ArangoDB



ArangoDB



Fully Managed

- ▶ Automated deployment
- ▶ Monitoring and Backup policies
- ▶ Scale-up / scale-out

Fully Programmable

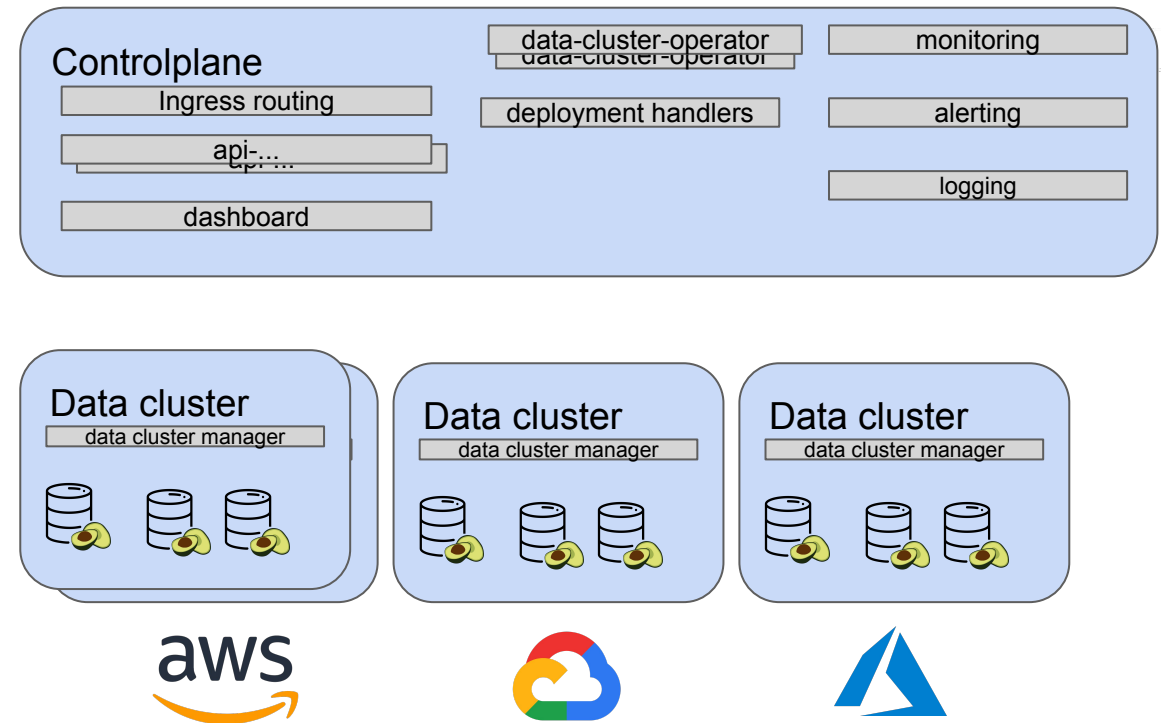
- ▶ Open APIs with 100% Coverage

Public Clouds

- ▶ AWS, Google and Azure

Flexible

- ▶ Start with a free-to-try offering
- ▶ Enterprise features in the cloud



Operational Challenges *

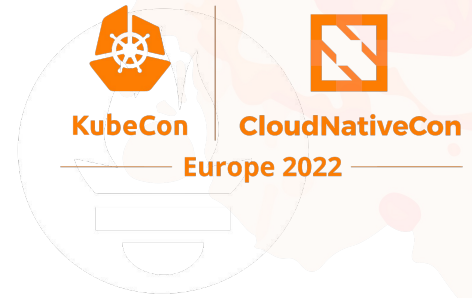
- (Managed) Kubernetes Upgrades
 - Upgraded Cluster \neq New Cluster
- Migrate between Regions
 - Cloud (Region) Outages
 - Reduce Latency between Client/Cluster
- Migrate between Cloud Provider
- Migrate from Self-Deployed to ArangoDB Cloud

Migration between different K8s cluster (and regions, providers, ...)

* small Subset :-)



Reliable Cloud



AWS went down hard, yet again - here's what happened

AWS suffered yet another major outage

By [Mike Moore](#) , [Joel Khalili](#) last updated December 22,



Google Cloud Service Health > Incidents

Azure status history

This page contains all root cause analyses (RCAs) for incidents that occurred on November 20, 2019 or later. Each RCA will be retained on this page for 5 years. RCAs before November 20, 2019 aren't available.

Product:

All

Region:

All

Date:

Most recent

April 2022

4/8

Service Management Operation Errors Across Azure Services in East US 2 (Tracking ID Y_5-9C0)

Summary of Impact: Between 12:25 UTC on 08 Apr 2022 and 14:40 UTC on 09 Apr 2022, customers running services in the East US 2 region may have experienced service management errors, delays, and/or timeouts. Customers may have experienced issues that caused GET and PUT errors impacting the Azure portal itself, as well as services including Azure Virtual Machines (VMs), Virtual Machine Scale Sets (VMSS), Azure Data Factory (ADF), Azure Databricks, Azure Synapse, Azure Backup, Azure Site Recovery (ASR), and Azure Virtual Desktop (AVD). Customers may have seen errors including "The network connectivity issue encountered for Microsoft.Compute cannot fulfill the request. For some downstream services that have auto-scale enabled, this service management issue may have caused data plane impact."

us-central1-c: GKE experiencing issues with some cluster and nodepool operations. Mitigation Underway.

20 Oct 2021

✓ 7 hours, 30 minutes

us-central1-c: GKE experiencing issues with some cluster and nodepool operations. Issue Mitigated.

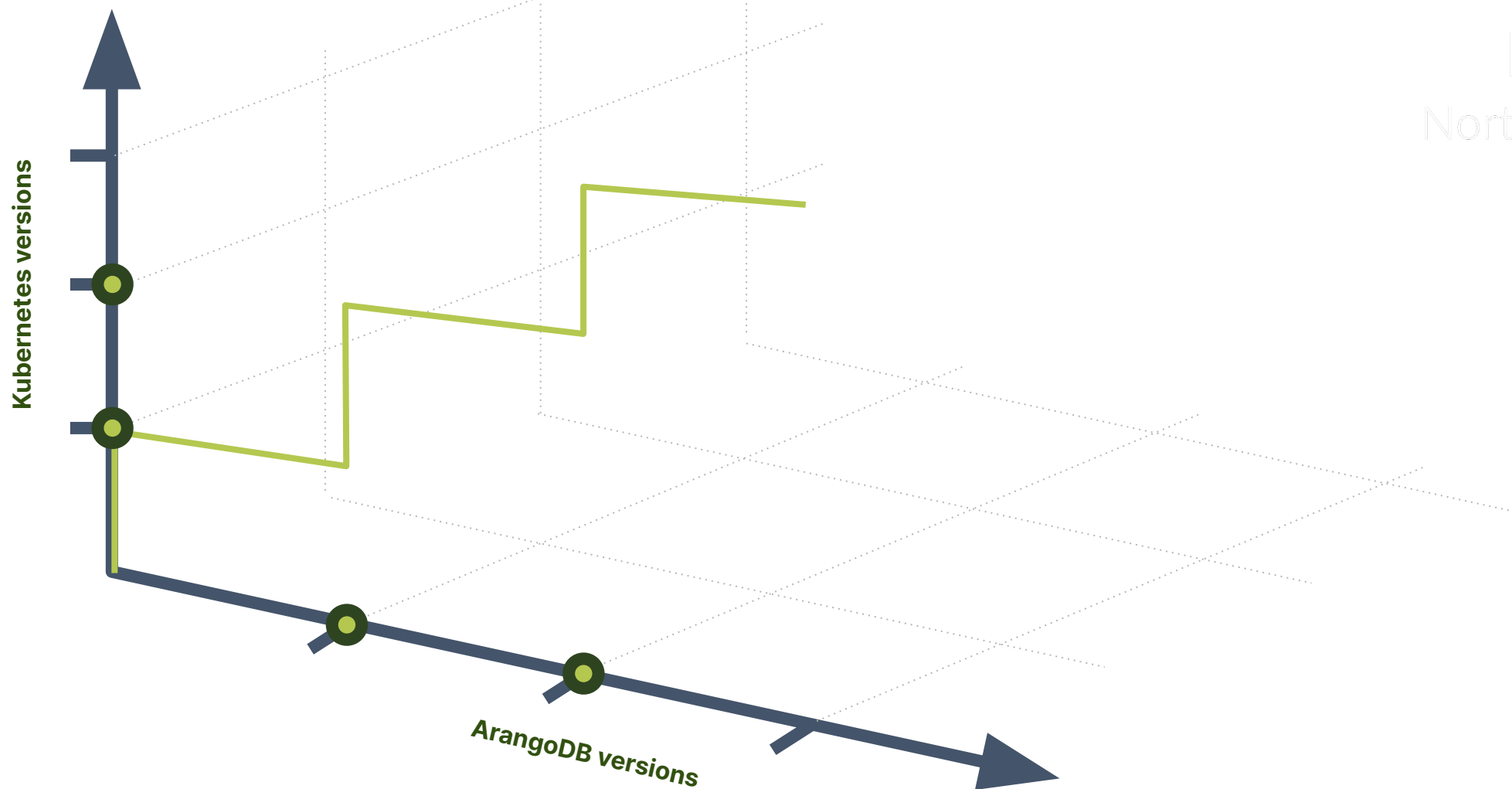
19 Oct 2021

✓ 4 hours, 41 minutes

Operations 101



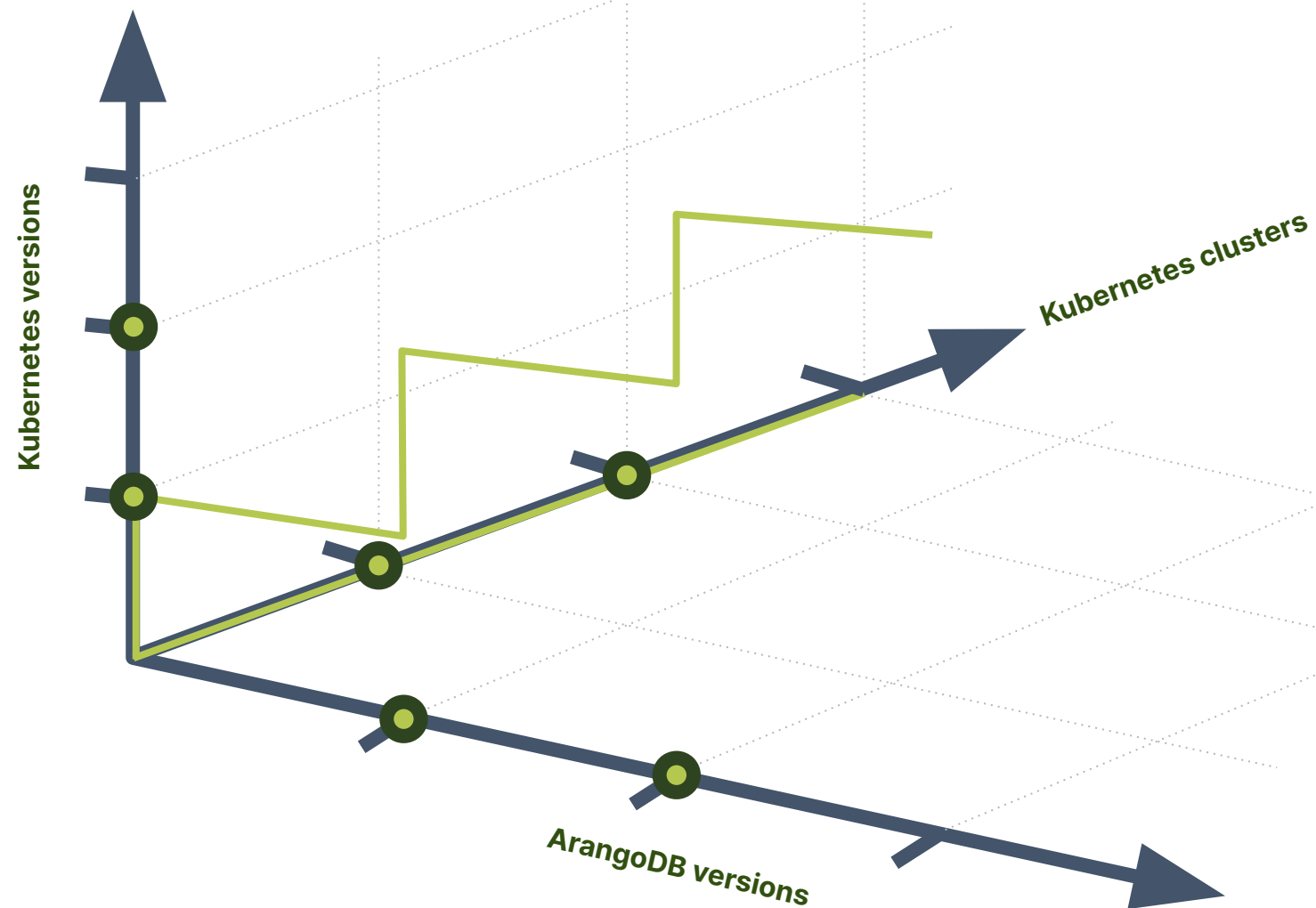
PromCon
North America 2021



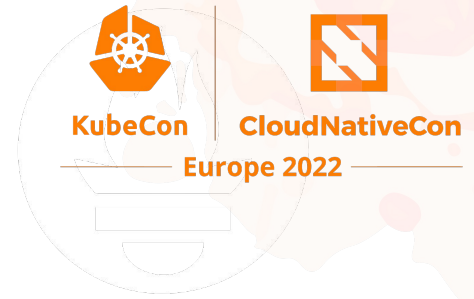
Operations 101



PromCon
North America 2021



Migration Challenge(s)



- User Transparency of Migration
 - Database and Client migration is not atomic operation!
 - within the period of DNS propagation, service needs to be also accessible from both clusters
 - latency should be as low as possible
 - database performance should not be significantly affected
- K8s and Cloud Networking
 - Reachability
 - Security
 - Reliability

PromCon
North America 2021

Required cooperation of all components!

Evaluated options

- Direct networking
 - Enabling direct Pod communication
- Proxy (kubectl port-forward)
 - Starting port-forward process instead of real pod
- Pod internet exposure
 - Exposing pod on the host level
- Service port mapping (LoadBalancer)
 - Exposing service with one-to-one port mapping to specific pods



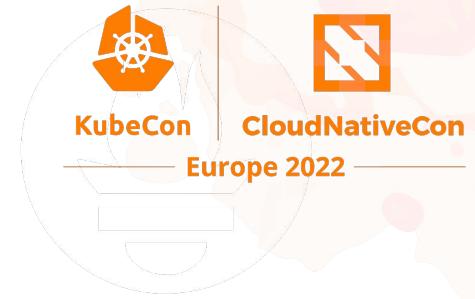
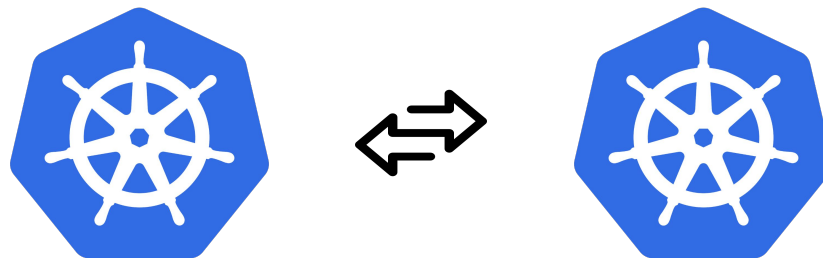
What worked for us? -1-

Direct networking (Cilium Cluster Mesh)

- Direct connect between clusters
- Endpoints needs to be managed by additional tool

Limitations:

- PodCIDR ranges must be non-conflicting
- Adjustment of old cluster can be problematic

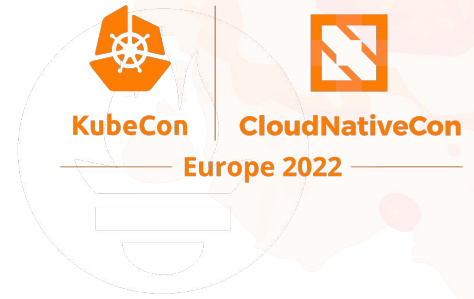


PromCon



cilium

What worked for us? -2-



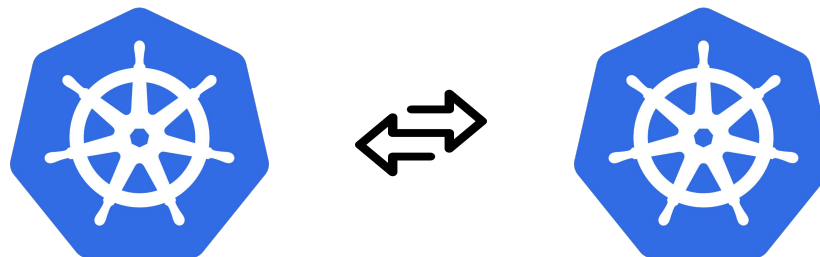
PromCon
North America 2021

Service (LoadBalancer)

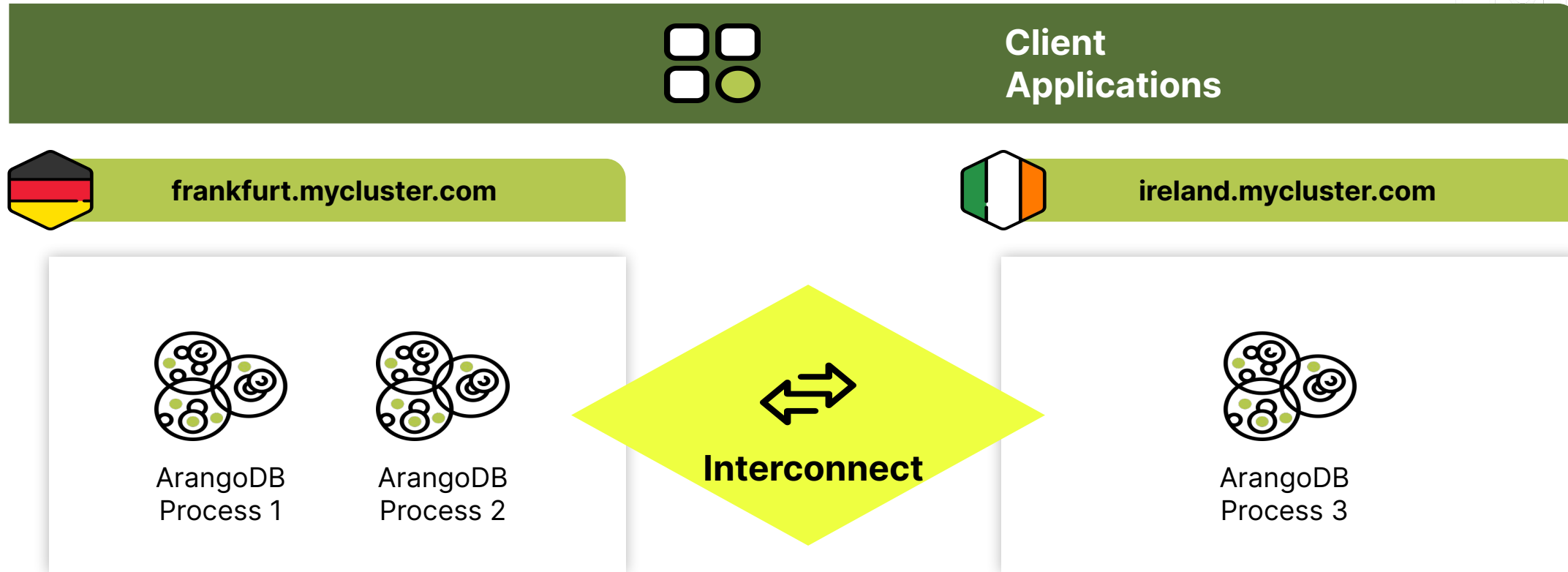
- Can be used in any PodCIDR ranges
- Works on all Cloud Providers
- One LoadBalancer per namespace

Limitations:

- LoadBalancer implementation required
- Additional latency

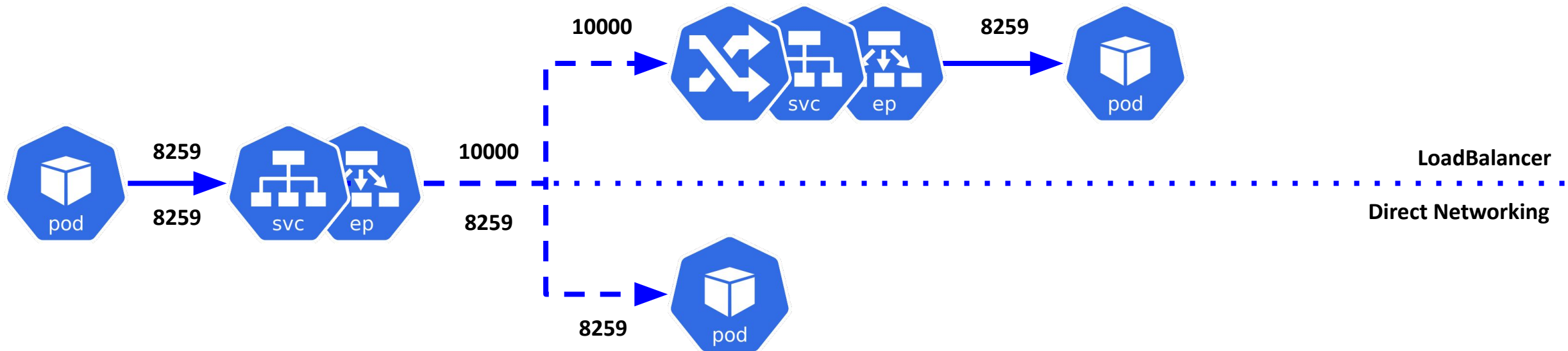


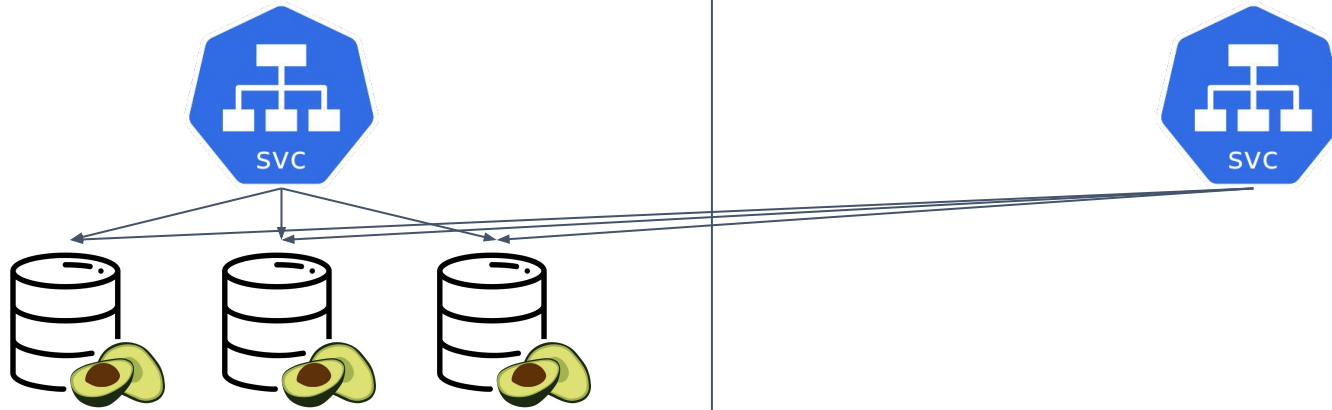
Migration between Regions

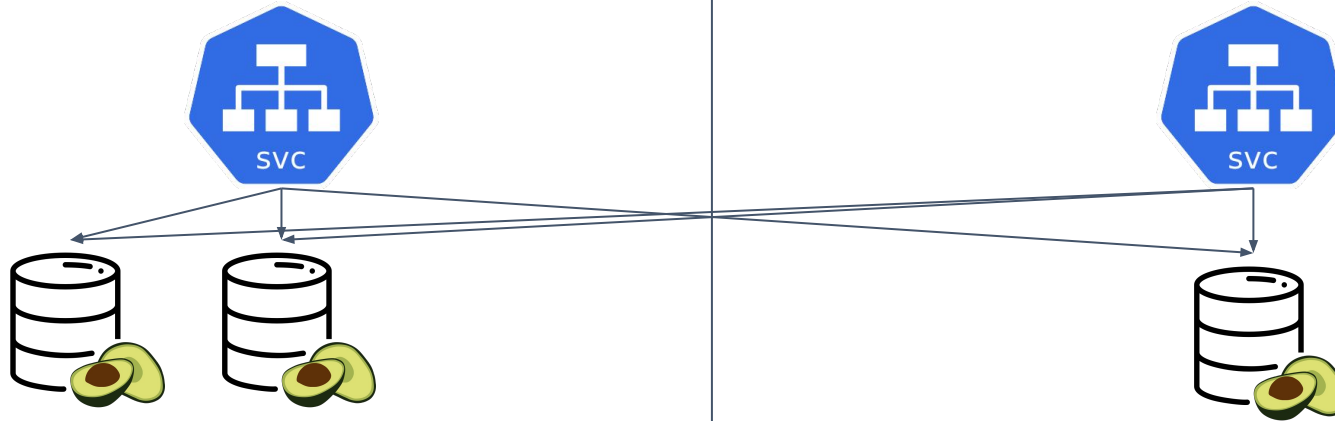


Network Perspective

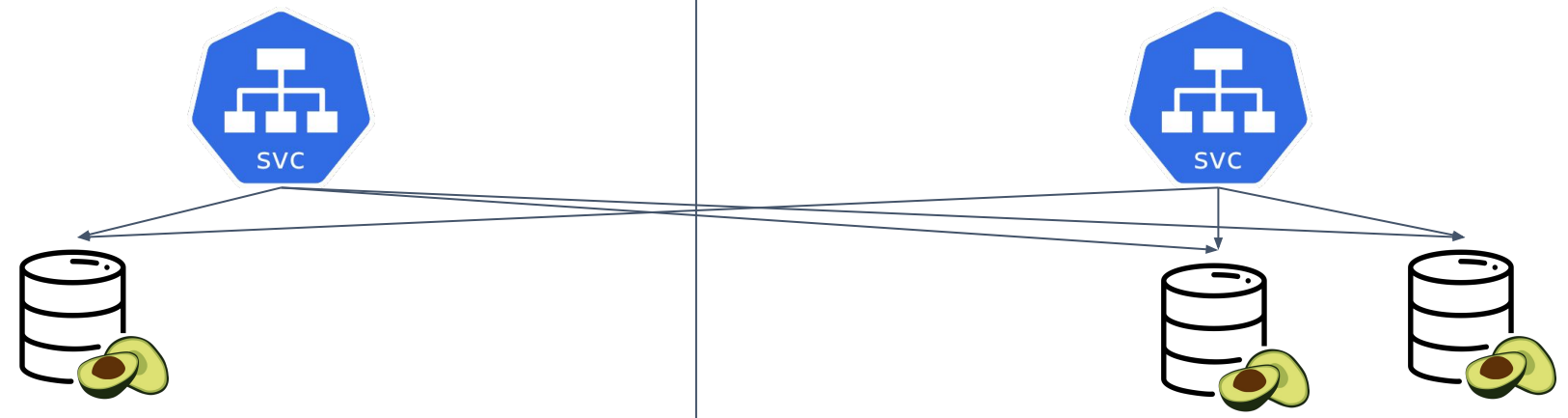
Using Services without selector we can easily manage IPs behind Service.
It can point directly to Pod IP or to virtual IP & Port on LoadBalancer.



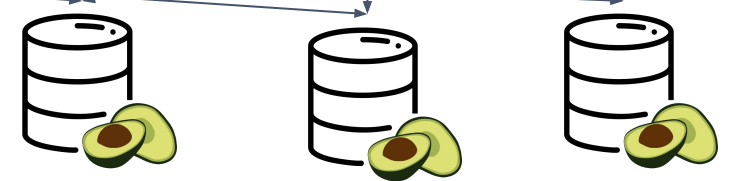




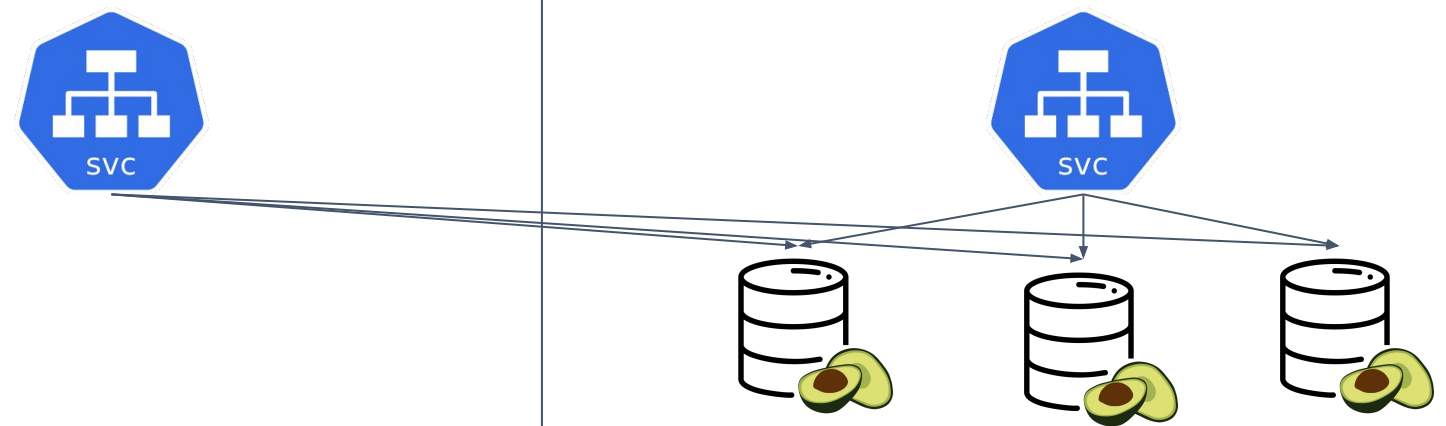
Cilium ClusterMesh



Cilium ClusterMesh



Cilium ClusterMesh



Cilium ClusterMesh

Demo



KubeCon



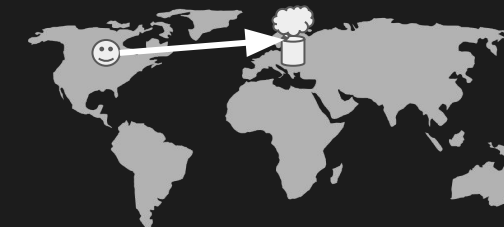
CloudNativeCon

Europe 2022

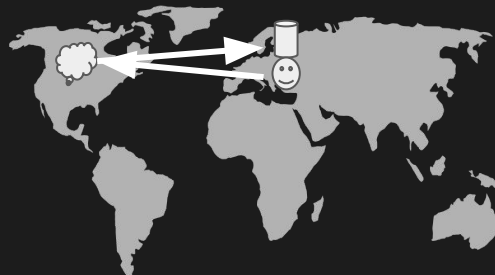
```
EU-EU: Version check: 126.337076ms
EU-EU: Version check: 118.552365ms
EU-EU: Version check: 128.318416ms
EU-EU: Version check: 136.648903ms
EU-EU: Version check: 126.207939ms
EU-EU: Version check: 135.726068ms
EU-EU: Version check: 126.274717ms
EU-EU: Version check: 127.686135ms
EU-EU: Version check: 136.576541ms
EU-EU: Version check: 126.720595ms
EU-EU: Version check: 127.496973ms
EU-EU: Version check: 120.747117ms
EU-EU: Version check: 126.698783ms
EU-EU: Version check: 130.77178ms
EU-EU: Version check: 128.568011ms
EU-EU: Version check: 134.475181ms
EU-EU: Version check: 133.998142ms
EU-EU: Version check: 118.770997ms
EU-EU: Version check: 127.686235ms
EU-EU: Version check: 134.310127ms
```



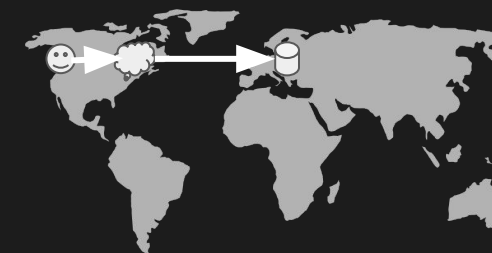
```
US-EU: Version check: 361.125024ms
US-EU: Version check: 360.888828ms
US-EU: Version check: 360.613599ms
US-EU: Version check: 359.056893ms
US-EU: Version check: 358.891485ms
US-EU: Version check: 361.970842ms
US-EU: Version check: 359.026569ms
US-EU: Version check: 359.859833ms
US-EU: Version check: 360.007583ms
US-EU: Version check: 360.698262ms
US-EU: Version check: 361.068818ms
US-EU: Version check: 361.157891ms
US-EU: Version check: 359.540758ms
US-EU: Version check: 359.060478ms
US-EU: Version check: 364.471375ms
US-EU: Version check: 360.736428ms
US-EU: Version check: 360.430665ms
US-EU: Version check: 360.380138ms
US-EU: Version check: 358.881459ms
US-EU: Version check: 359.341218ms
```



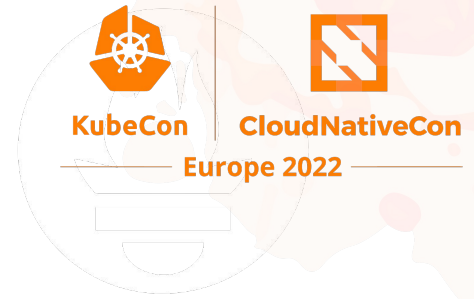
```
EU-US: Version check: 718.097018ms
EU-US: Version check: 705.184403ms
EU-US: Version check: 709.903133ms
EU-US: Version check: 704.439429ms
EU-US: Version check: 709.789228ms
EU-US: Version check: 708.146709ms
EU-US: Version check: 708.687007ms
EU-US: Version check: 704.800378ms
EU-US: Version check: 704.011259ms
EU-US: Version check: 709.986457ms
EU-US: Version check: 703.22943ms
EU-US: Version check: 707.975414ms
EU-US: Version check: 709.34647ms
EU-US: Version check: 710.939729ms
EU-US: Version check: 711.425207ms
EU-US: Version check: 710.820513ms
EU-US: Version check: 709.754996ms
EU-US: Version check: 708.700447ms
EU-US: Version check: 713.927667ms
EU-US: Version check: 709.508458ms
```



```
US-US: Version check: 360.658082ms
US-US: Version check: 359.374911ms
US-US: Version check: 360.482318ms
US-US: Version check: 361.019328ms
US-US: Version check: 358.936874ms
US-US: Version check: 359.226745ms
US-US: Version check: 360.741628ms
US-US: Version check: 360.24982ms
US-US: Version check: 359.414612ms
US-US: Version check: 359.922797ms
US-US: Version check: 359.076539ms
US-US: Version check: 360.916636ms
US-US: Version check: 360.209578ms
US-US: Version check: 360.952281ms
US-US: Version check: 362.511985ms
US-US: Version check: 360.416689ms
US-US: Version check: 360.161465ms
US-US: Version check: 359.397269ms
US-US: Version check: 359.515998ms
US-US: Version check: 360.753769ms
```



Lessons Learned



What does not work:

- Using
 - Kubernetes port-forward
 - exposing pods on host port via public IP
- Manual endpoint juggling
- Using solely one of the two options

What might work (in the future):

- Using one operator per cluster
 - Challenge: Leader Elections
- Distributed Cluster
 - Latency (Agency/leader election)

PromCon
North America 2021

Thanks for Listening!



Test-drive Oasis
14-days for free

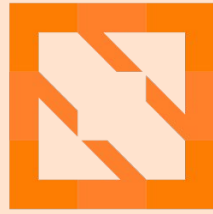


Reach out with Feedback/Questions!

- <https://www.arangodb.com/>



KubeCon



CloudNativeCon

Europe 2022

WELCOME TO VALENCIA

