
KubeCon
CloudNativeCon
Europe 2023
TiKV

Multi-Cluster Observability with Service Mesh

That Is a Lot of Moving Parts!?

Ryota Sawada
UPSIDER, Inc.

KubeCon | CloudNativeCon
Europe 2023

Agenda

1. **Introduction**
   - About Me / About UPSIDER / About This Talk
2. **Observability with Service Mesh**
   - Standard Setup / Multi-Cluster
3. **Multi-Cluster Challenges**
   - Inter-Cluster Metrics / Data Source / Retention / Cardinality
4. **Quick Review on Key Projects**
   - Prometheus / Istio / Thanos
5. **Solutions and Demo**
   - Overview / Demo
6. **Summary**
   - Recap / Alternatives / Other Considerations

Photo by Max van den Oetelaar on Unsplash

1. Introduction

1. Introduction

About **Me**

- 🏔️ Platform Engineer at UPSIDER, Inc.
- ❤️‍🔥 Passionate about learning, building, and sharing all things technology
- 😬 First time speaking at in-person event
- 👶 Has a son and two cats

Find me **@rytswd** at usual places

@rytswd **@rytswd** **@rytswd@hachyderm.io**

About UPSIDER

- 💳 Fintech startup, providing B2B payment solutions
- 🚀 Adopted by 15,000+ companies
- 🗾 Headquarter based in Tokyo, Japan
- 🌏 Remote team around the world
- 🌤️ Cloud Native End User from early 2019
- 🤖 Multi-Cloud and Multi-Cluster system

🥰 We are hiring!

About This Talk

📚 Prerequisites

Basic ideas about:

👉 Observability

- Metrics 🌟
- Traces
- Logs

👉 Service Mesh

- Traffic management
- Security
- Observability 🌟

✅ This talk is about

- Service Mesh behaviours
- Multi-cluster scenarios
- Observability challenges
- Our learnings at UPSIDER
- “It is a lot of moving parts”

❌ This talk is NOT about

- Istio deep dive
- Prometheus deep dive
- “Silver bullet setup”

💡 More at:
github.com/rytswd/kubecon-eu-2023

2. Observability with Service Mesh

2. Observability with Service Mesh

Standard Setup

🤔 What does "**Observability**" look like with Service Mesh?

🌱 Focus on a single cluster

⚓ Deploy Istio and Prometheus

🔬 Visualise connections

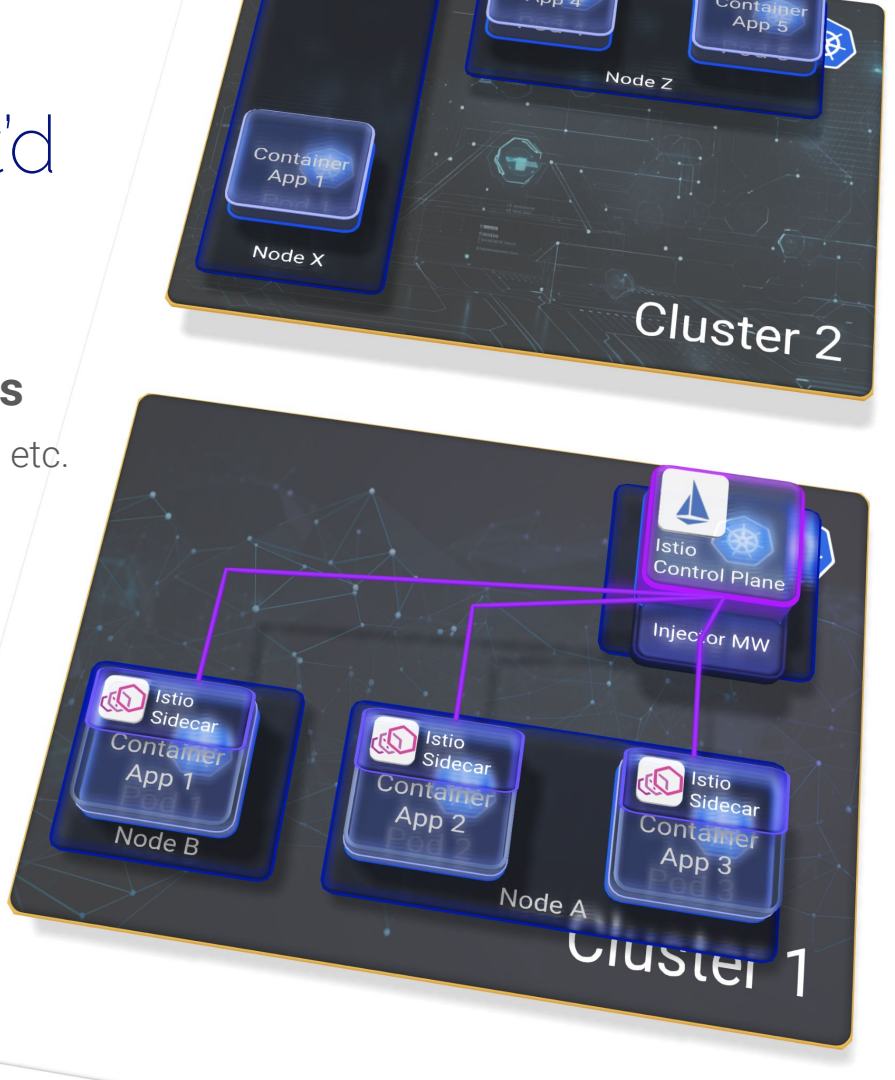
Standard Setup Cont'd

1 Deploy Istio Control Plane

We get Istio Sidecars injected to each Pod

Standard Setup Cont'd

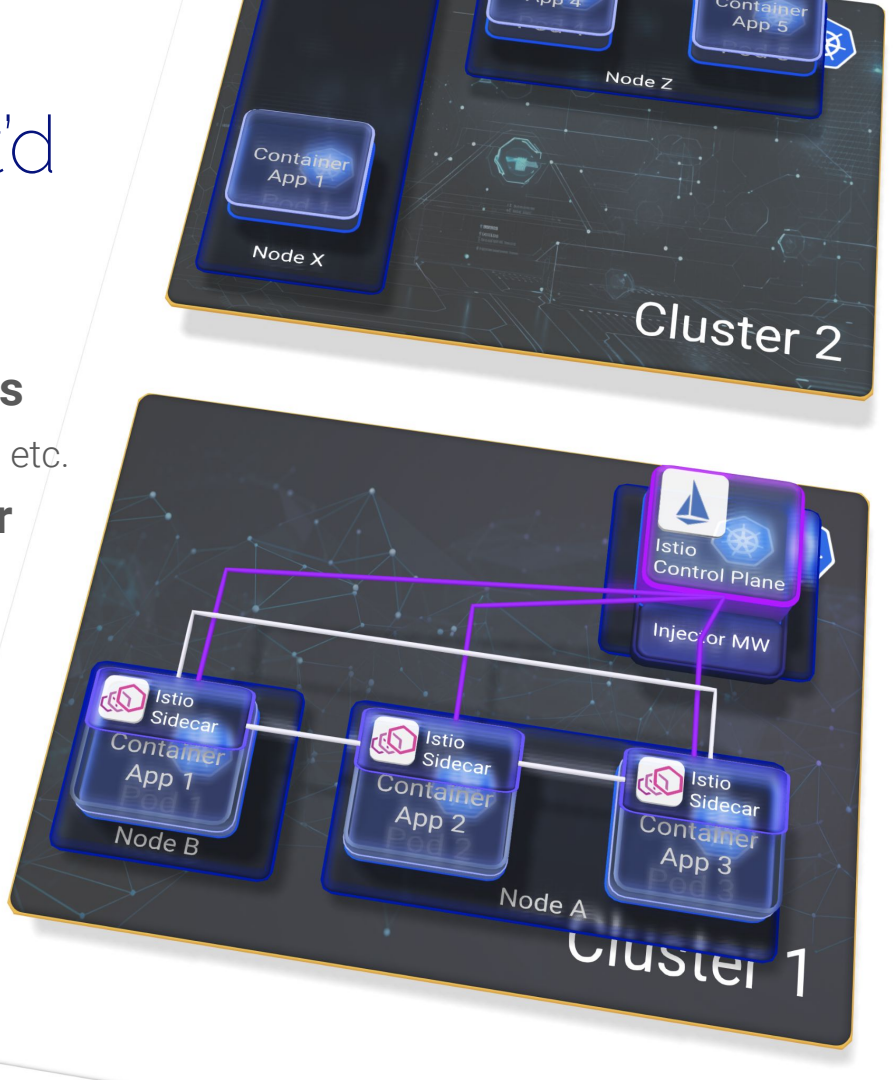
1. **Deploy Istio Control Plane**
   We get Istio Sidecars injected to each Pod
2. **Istio Control Plane Connects to Sidecars**
   Sidecars will get routing details, credentials, metrics, etc.



2. Observability with Service Mesh

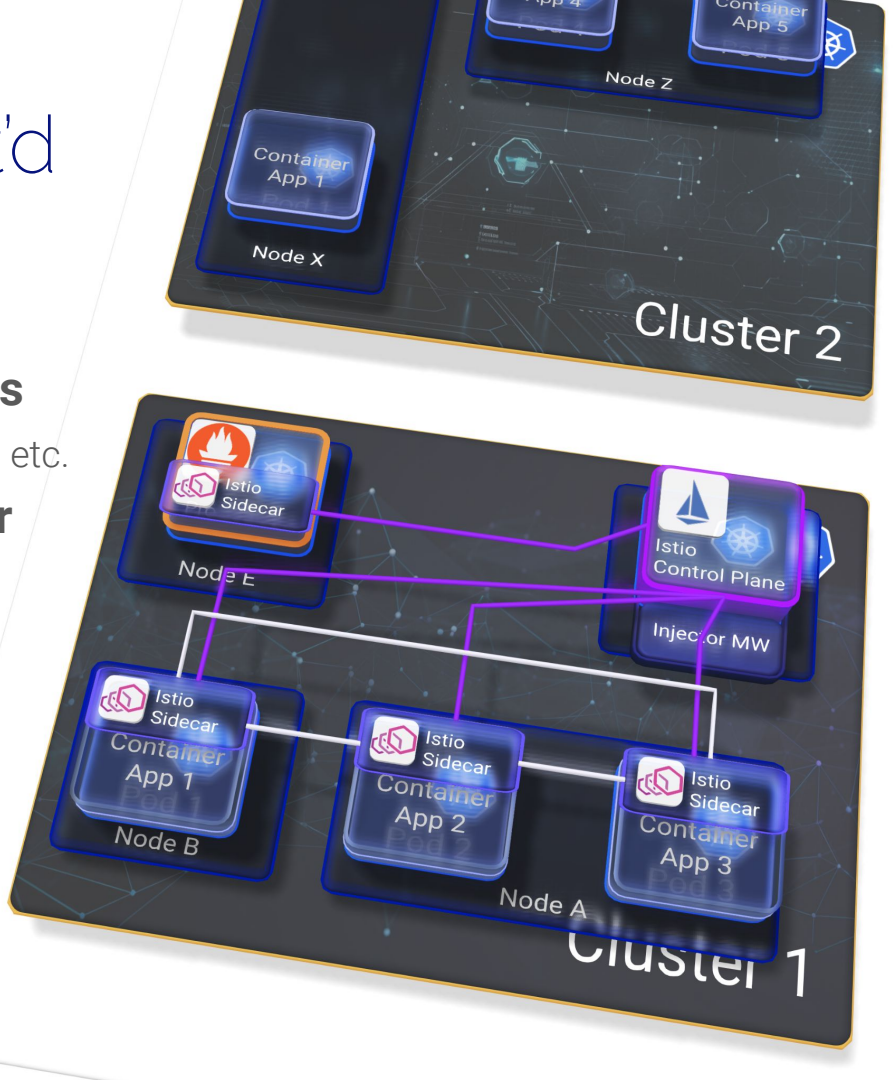
Standard Setup Cont'd

1. **Deploy Istio Control Plane**
   We get Istio Sidecars injected to each Pod
2. **Istio Control Plane Connects to Sidecars**
   Sidecars will get routing details, credentials, metrics, etc.
3. **Istio Data Plane Connects to Each Other**
   Sidecars handles all the connections



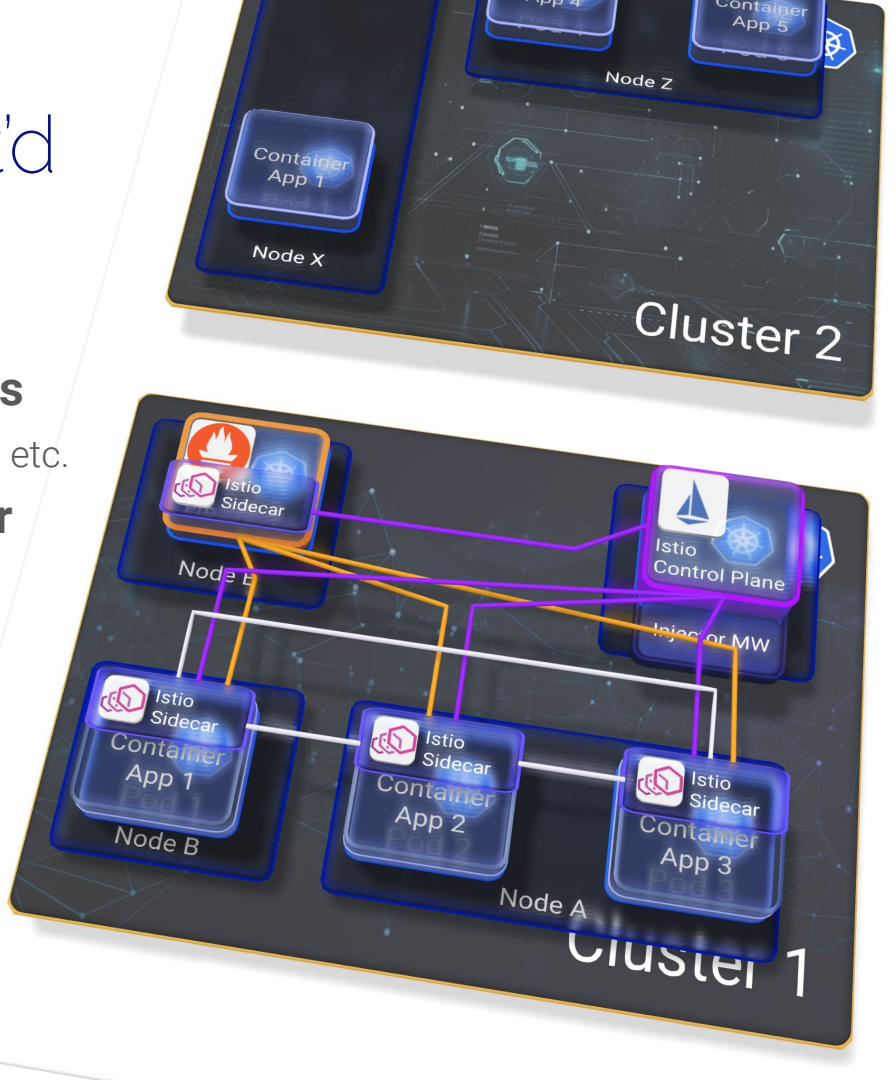
Standard Setup Cont'd

1. **Deploy Istio Control Plane**
   We get Istio Sidecars injected to each Pod
2. **Istio Control Plane Connects to Sidecars**
   Sidecars will get routing details, credentials, metrics, etc.
3. **Istio Data Plane Connects to Each Other**
   Sidecars handles all the connections
4. **Deploy Prometheus**
   Istio Sidecar will be also deployed with Prometheus



Standard Setup Cont'd

1. **Deploy Istio Control Plane**
   We get Istio Sidecars injected to each Pod
2. **Istio Control Plane Connects to Sidecars**
   Sidecars will get routing details, credentials, metrics, etc.
3. **Istio Data Plane Connects to Each Other**
   Sidecars handles all the connections
4. **Deploy Prometheus**
   Istio Sidecar will be also deployed with Prometheus
5. **Prometheus Retrieves Metrics**
   Istio certificates need to be used in strict mTLS env
   Ref: [Istio documentation on Prometheus]



Standard Setup Cont'd

1 Deploy Istio Control Plane

We get Istio Sidecars injected to each Pod

2 Istio Control Plane Connects to Sidecars

Sidecars will get routing details, credentials, metrics, etc.

3 Istio Data Plane Connects to Each Other

Sidecars handles all the connections

4 Deploy Prometheus

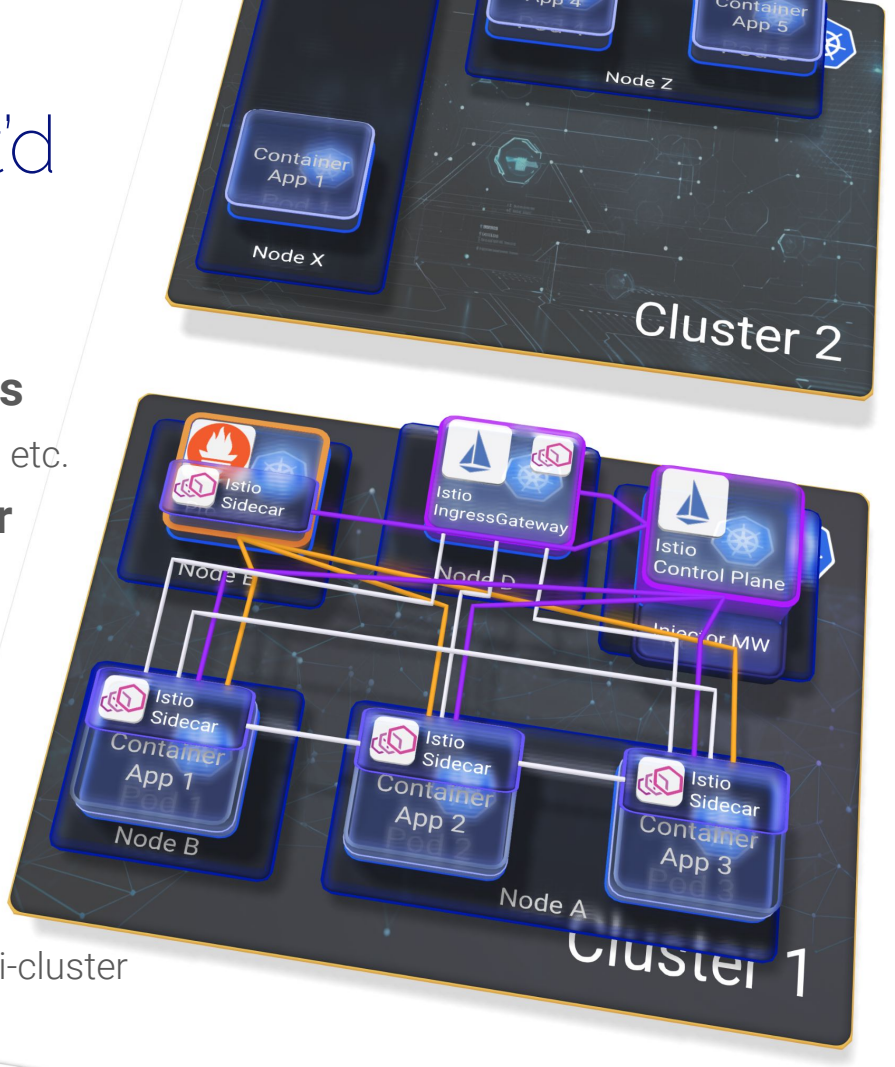
Istio Sidecar will be also deployed with Prometheus

5 Prometheus Retrieves Metrics

Istio certificates need to be used in strict mTLS env

Ex Deploy Istio IngressGateway

Not strictly necessary, but will be necessary for multi-cluster



2. Observability with Service Mesh

Multi-Cluster

🌿 Add another cluster

🚢 Deploy separate Istio and Prometheus

NOTE: This is only one pattern out of many

🏄 Better HA and higher uptime

✨ Multi-Cloud / multi-”managed Kubernetes”

🤹 Allow workload migration to other cluster for better cluster management

💡 At UPSIDER, we also separate clusters for sensitive information management

Multi-Cluster Cont'd

1. **Deploy Separate Istio Control Plane**
   This only demonstrates one installation pattern
2. **Deploy Prometheus**
   Separate Prometheus for another cluster

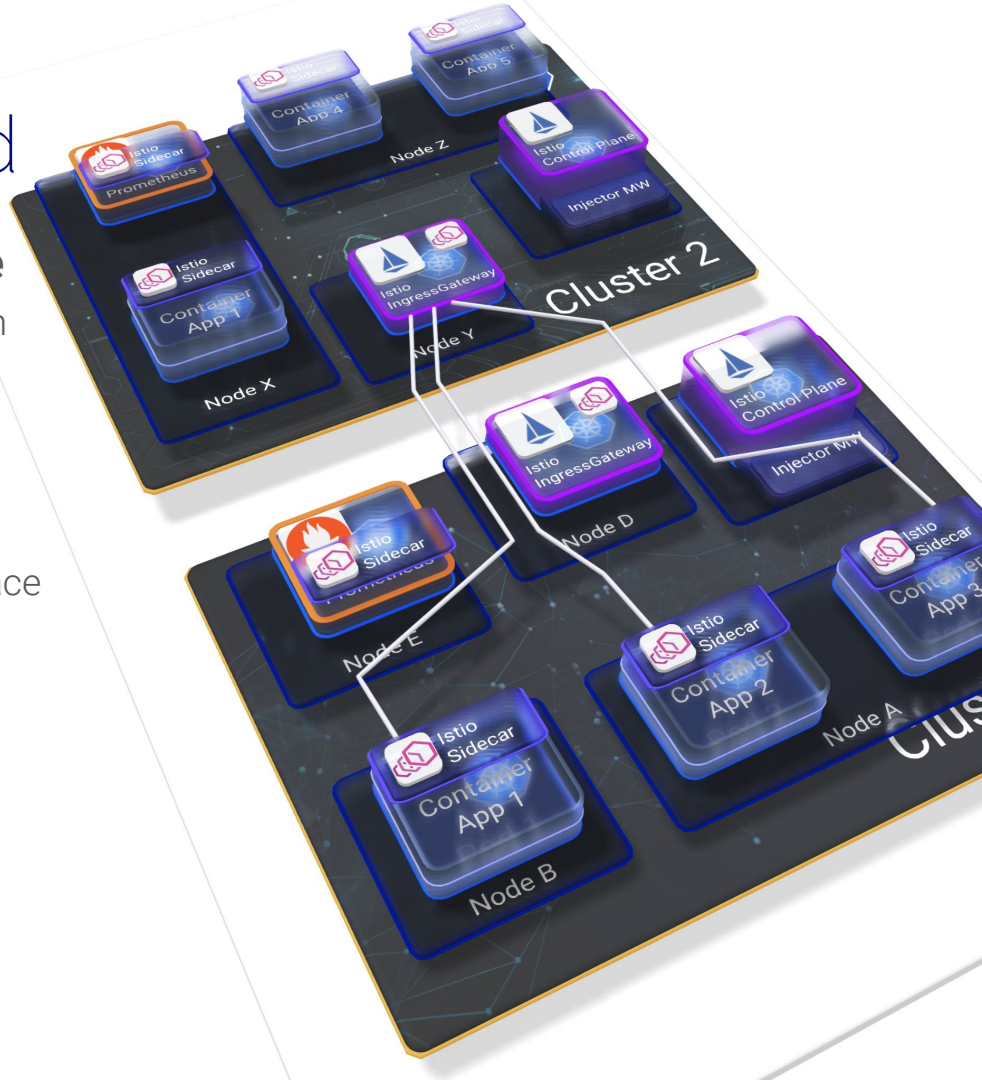
2. Observability with Service Mesh

Multi-Cluster Cont'd

1. **Deploy Separate Istio Control Plane**
   This only demonstrates one installation pattern
2. **Deploy Prometheus**
   Separate Prometheus for another cluster
3. **Deploy Istio IngressGateway**
   Multi-Cluster connection needs this to be in place

Multi-Cluster Cont'd

1. **Deploy Separate Istio Control Plane**
   This only demonstrates one installation pattern
2. **Deploy Prometheus**
   Separate Prometheus for another cluster
3. **Deploy Istio IngressGateway**
   Multi-Cluster connection needs this to be in place
4. **Connect Between Clusters**
   Connect from Cluster 1 to Cluster 2
   NOTE: EgressGateway is not used



Multi-Cluster Cont'd

1 Deploy Separate Istio Control Plane

This only demonstrates one installation pattern

2 Deploy Prometheus

Separate Prometheus for another cluster

3 Deploy Istio IngressGateway

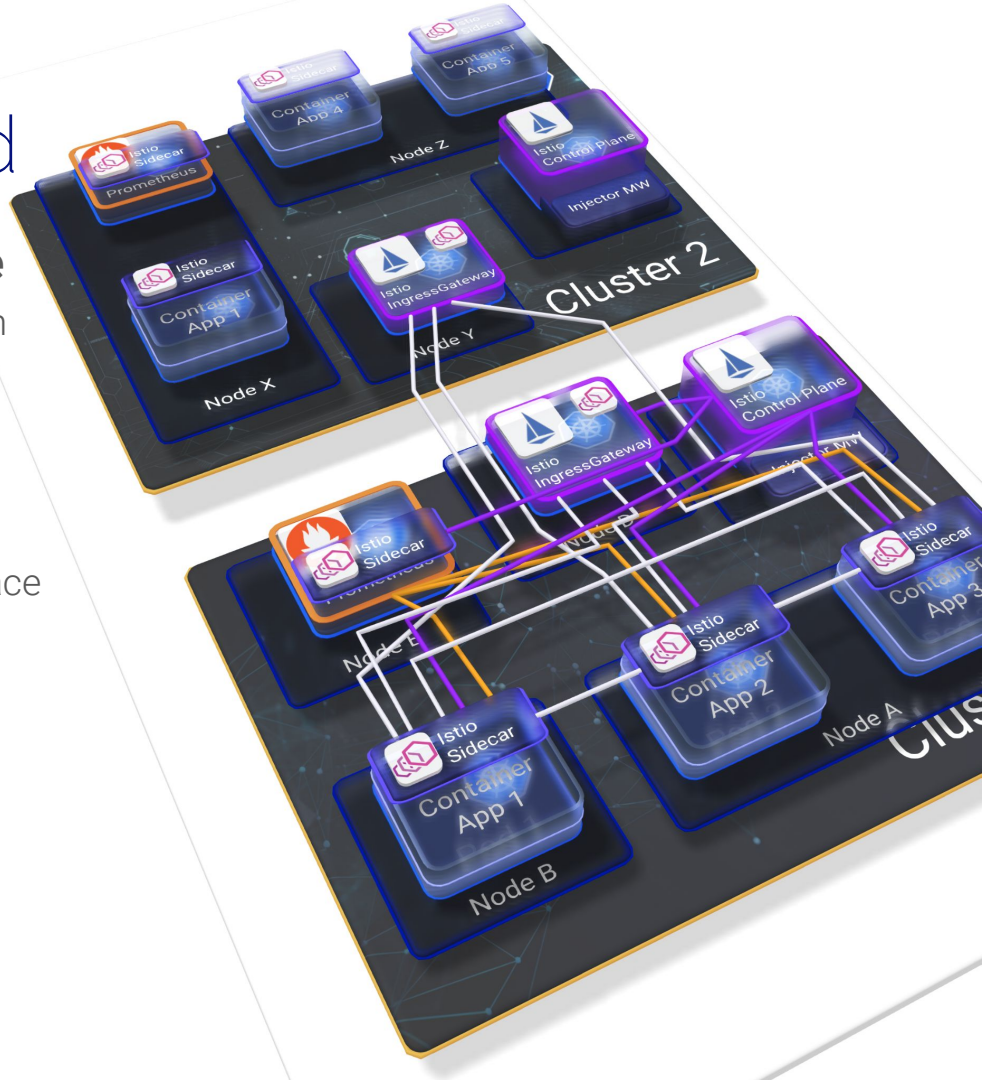
Multi-Cluster connection needs this to be in place

4 Connect Between Clusters

Connect from Cluster 1 to Cluster 2

Ex Visualise Altogether

This is still missing many connections



3. Multi-Cluster Challenges

3. Multi-Cluster Challenges

Inter-Cluster Metrics

- 🐾 Traffic leaves Cluster A to Cluster B
- 🍃 Each cluster can have its own observability setup
- 🤔 Who would be responsible for inter-cluster metrics?
- 🤯 How can we merge multiple metrics from a number of clusters?

3. Multi-Cluster Challenges

Cardinality

- 🦚 Service Mesh allows getting detailed metrics
- ☐ It is more of a challenge for Service Mesh rather than multi-cluster setup
- 🎛️ There will be so much data associated for every cluster
- 🥵 Observability dashboard could have performance issues

3. Multi-Cluster Challenges

Retention

⏰ Latest metrics can be fetched from running services

💾 Metrics need to be retained for historical investigation and analysis

😵 Where and how should we store metrics?

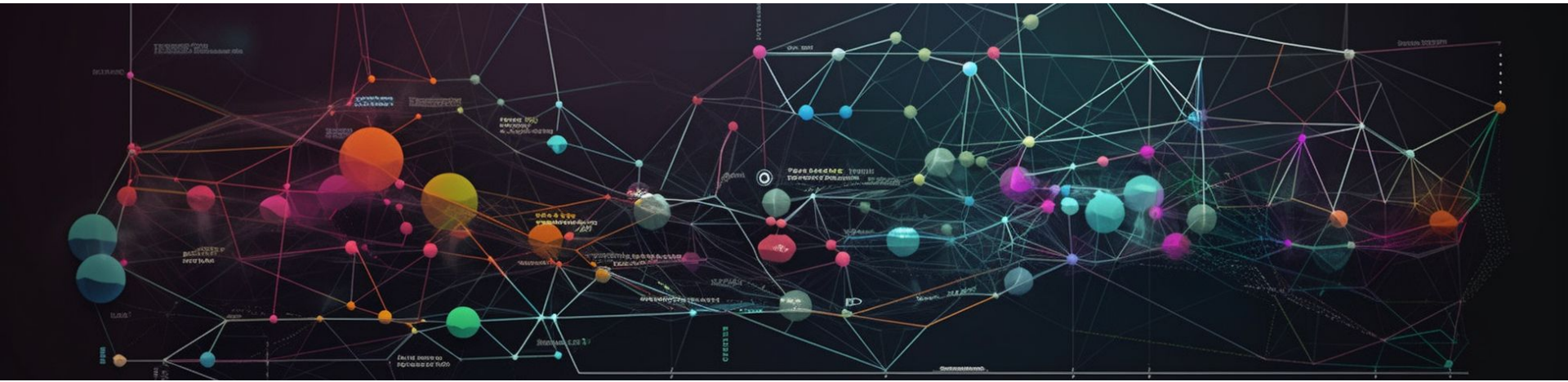
🤒 If cardinality is too high, data retention cost can easily add up

3. Multi-Cluster Challenges

Data Source

- There could be one service deployed in a number of clusters
- Dashboard needs to be aware of data source origin for all metrics
- Metrics deduplication needs to happen per data source

4. Quick Review on Key Projects



Istio

- **CNCF Incubating Project**
  Accepted in Sep 2022
- **Service Mesh with Envoy**
  Envoy, CNCF Graduated Project, plays a significant role for Data Plane
- **Mutual TLS (mTLS)**
  mTLS is an important piece for multi-cluster to work
- **Various Multi-Cluster Model**
  Multi-Primary, Primary-Remote, etc.
- **Sidecar and Ambient Deployment**
  Support for “ambient mesh” was introduced in Sep 2022

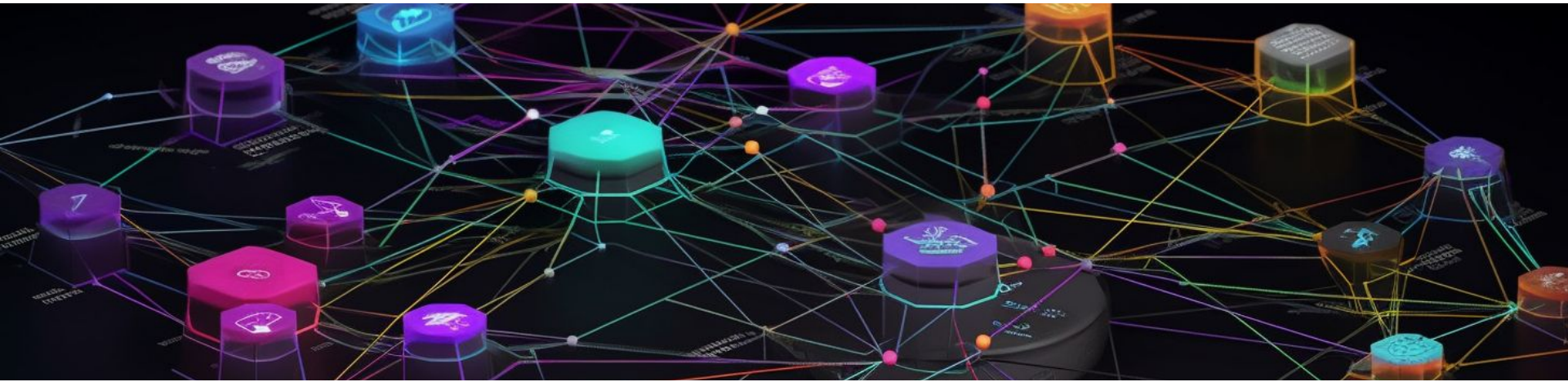
Prometheus

- **CNCF Graduated Project**
  Graduated in Aug 2018, second project to graduate, following Kubernetes
- **Remote Read / Write APIs**
  Allows flexible integration
- **Prometheus Operator**
  Deploy as many Prometheus as you need with simple CRD
- **Many Prometheus Related Services**
  Thanos, Cortex, Grafana Mimir, etc. are all related services

Thanos

- **CNCF Incubating Project**
  Accepted in July 2019
- **Long Term Storage for Prometheus**
  Handles heavy lifting of data retention
- **Sidecar or Receiver**
  Deployment model can be adjusted according to business needs
- **Performant Queries**
  Designed with query specific components such as query-frontend

5. Solutions and Demo



Overview

One solution for demo, using 3 Kubernetes cluster

Istio

- Deploy using manifests
- Multi-Primary on different networks
- Strict mTLS
- Sidecar deployment
- Non-demo profile
  (Prometheus and other components will not be deployed together)

Prometheus

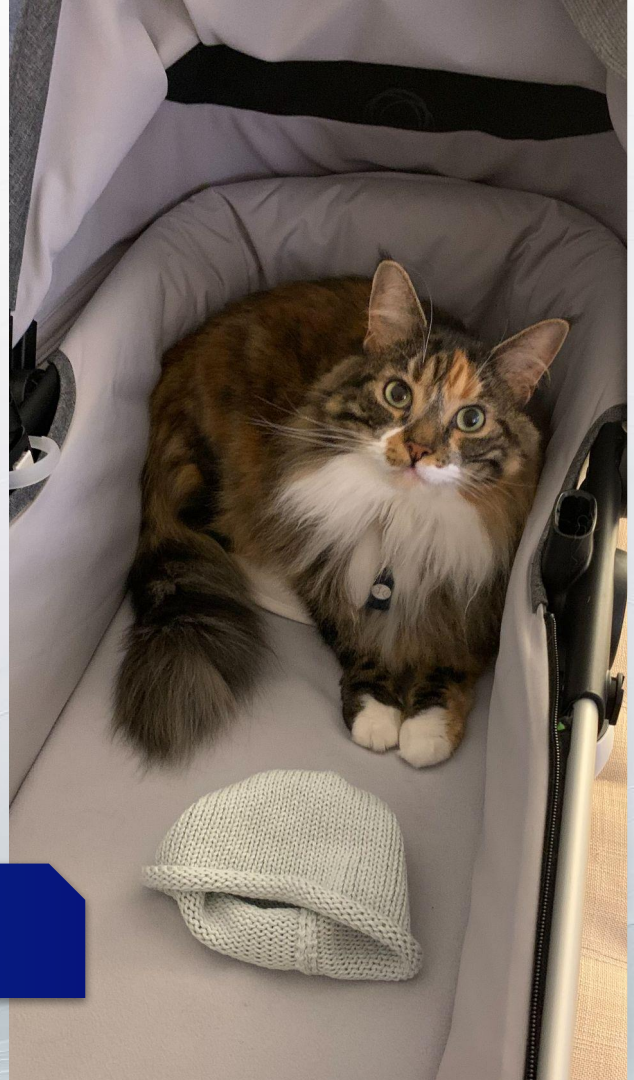
- Deploy using Prometheus Operator
- Prometheus Federation for high cardinality
- Remote Write

Thanos

- Deploy using Helm
- Use Receiver rather than Sidecar

5. Solutions and Demo

Demo Time!



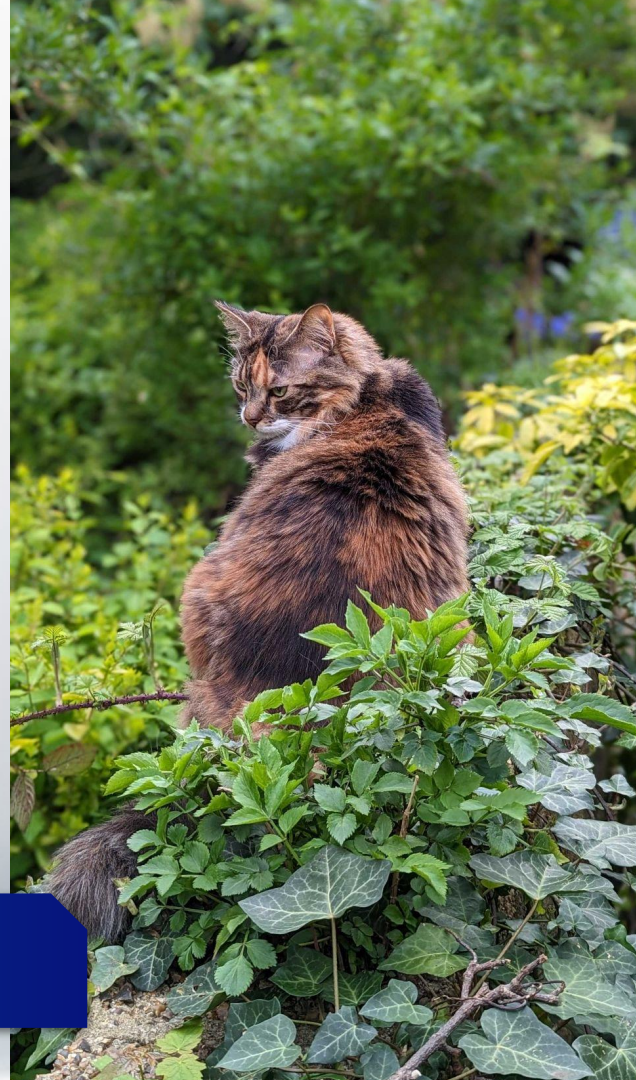
💡 The Demo setup is available at:
github.com/rytswd/kubecon-eu-2023

6. Summary

6. Summary

Recap

- **There Are A Lot of Moving Parts!!**
  A few clusters already seem too much to comprehend
- **Multi-Cluster Made Easy by Service Mesh**
  Istio certainly made multi-cluster service setup much easier
- **Observability Is Complicated**
  Service Mesh is a part of story, it is not a solution by itself
- **Solution with Central Observability Cluster**
  It is one of many options for multi-cluster environment
- **Multi-Cluster Observability Challenges**
  Check out the demo repository for more details



💡 The Demo setup is available at:
github.com/rytswd/kubecon-eu-2023

6. Summary

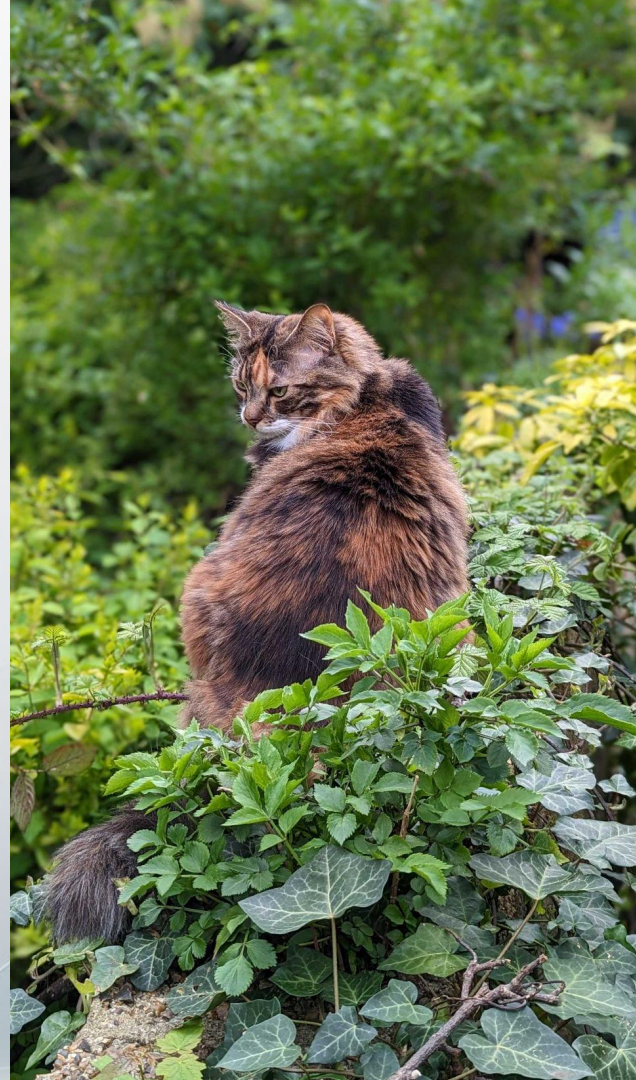
Alternatives

Istio

- Cilium
  https://cilium.io/
  https://isovalent.com/blog/post/cilium-service-mesh/
- Envoy
  https://www.envoyproxy.io/
- Linkerd
  https://linkerd.io/

Thanos

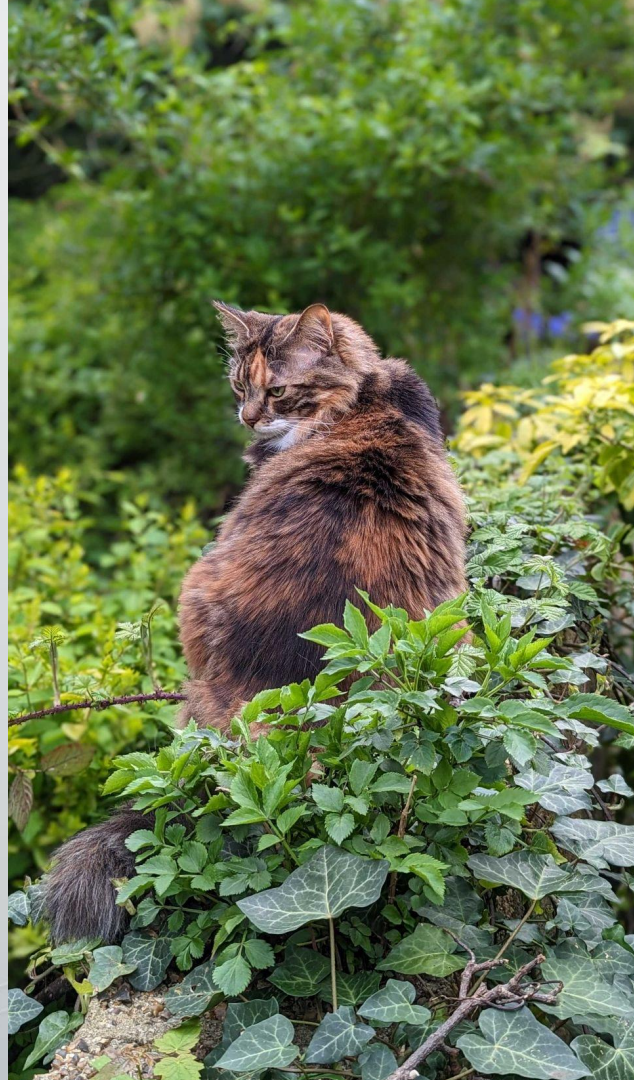
- Cortex
  https://cortexmetrics.io/
- Grafana Mimir
  https://grafana.com/oss/mimir/
- M3
  https://m3db.io/
- VictoriaMetrics
  https://victoriametrics.com/products/open-source/



6. Summary

Other Considerations

- **Alert Management**
  Alerts should be managed per cluster in general
- **Configuration Management**
  There are so many moving parts, GitOps can help make sense
- **Other Observability Aspects**
  We did not cover traces and logs
- **Service Mesh Specific Solutions**
  Istio has Kiali, and Cilium has Hubble
- **Multi-Tenancy**
  Data segregation is another challenge
- **OpenTelemetry**
  Support for OpenTelemetry has been growing
- **Sidecar or Sidecarless**



Appendix

Appendix

Other Talks at KubeCon

Wednesday, April 19 • 16:30 - 17:05

Operate Multi-Tenancy Service Mesh with ArgoCD in Production

Lin Sun, Solo.io & Faseela K, Ericsson Software Technology

Wednesday, April 19 • 17:25 - 18:00

OTel Me About Metrics: A Metrics 101 Crash Course

Reese Lee, New Relic

Thursday, April 20 • 16:30 - 17:05

Future of Istio - Sidecar, Sidecarless or Both?

Neeraj Poddar, Solo.io

Friday, April 21 • 14:55 - 15:30

Effortless Open Source Observability with Cilium, Prometheus and Grafana - LGTM!

Raymond de Jong & Anna Kapuścińska, Isovalent

Friday, April 21 • 16:00 - 16:35

Future of Service Mesh - Sidecar or Sidecarless or Proxyless?

Idit Levine & Yuval Kohavi, Solo.io; Keith Mattix II, Microsoft; Eric Van Norman, IBM; John Howard, Google

Appendix

References

Repository

https://github.com/rytswd/kubecon-eu-2023

Slide Deck

https://dub.sh/kubecon-eu-2023-mco (This slide deck)

Tools used

- MidJourney - for many image materials
  https://www.midjourney.com/
- Spline - for 3D image
  https://spline.design/
  https://app.spline.design/file/6505bdd6-dcd6-41c5-bff5-40538ec6b16e

Appendix

Special Thanks

Presentation Prep

Maciej Gaweł

Other Support

Gabor Szendrei
Hideto Miki
Hiroto Funakoshi
Miki Masumoto
Vlad Frățilă
Yoshitaka Sakakibara

Thank You!

Ryota Sawada @ rytswd
UPSIDER, Inc.

☺️ Please leave feedback with this QR Code 👉


KubeCon
CloudNativeCon
Europe 2023
TiKV