PERCONA

Best Practices For Running Databases on Kubernetes

March 15,2024

Peter Zaitsev,
Founder, Percona

Big Picture

PERCONA

Cloud

Proprietary Clouds bring Great Usability at Great Cost

PERCONA

Have We Been Here Before?

2000s

2020s

PERCONA

Operating Systems

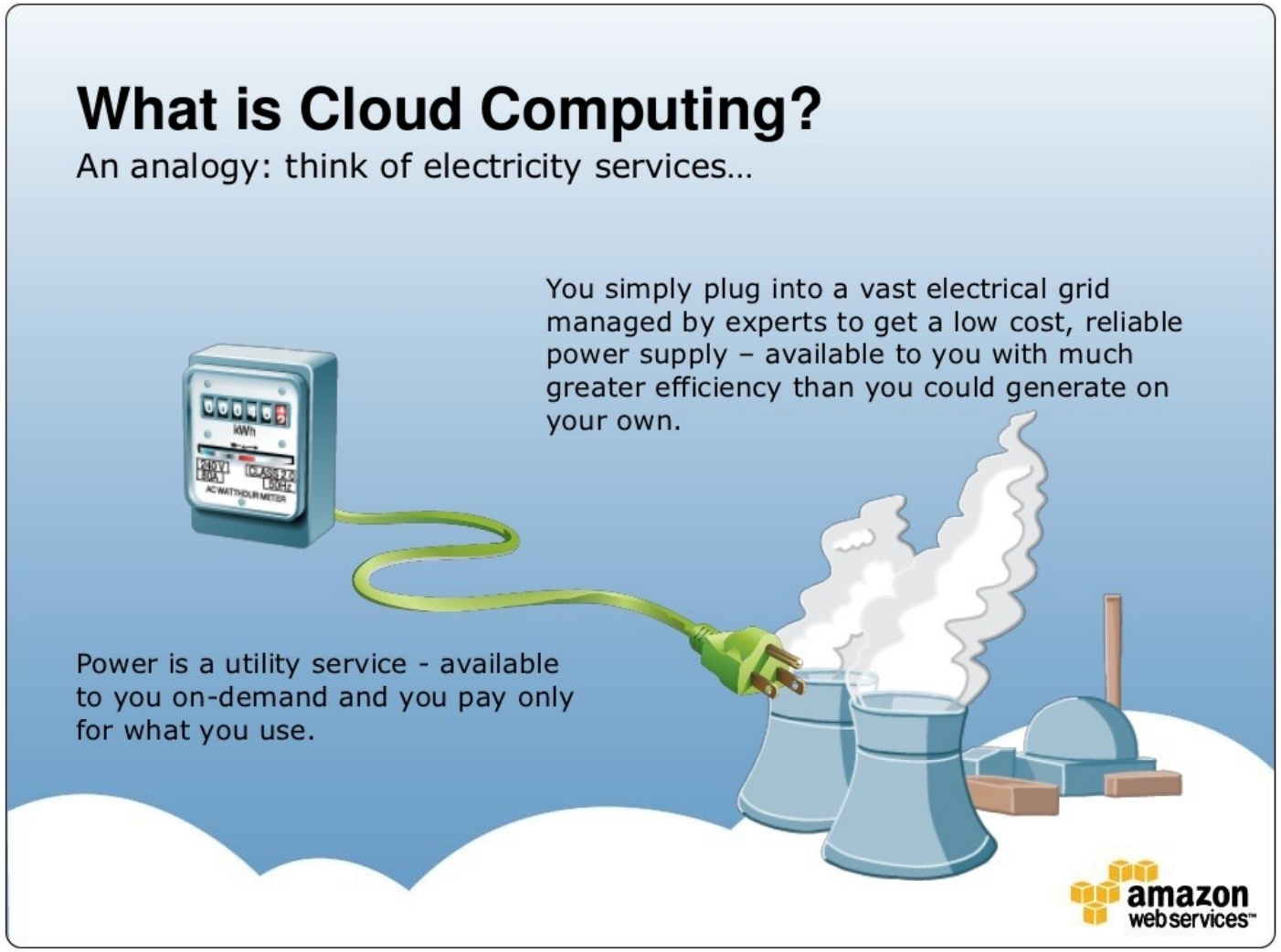
Open Source Catches Up Again

- **Lock-in with Cloud Vendor**
- **Use Proprietary Solutions**
- **Highly Differentiated Cloud**

- **Freedom to Run Anywhere**
- **Use Open Source**
- **Cloud Is Commodity**
- **Customer**
- **Choice of Vendors**

PERCONA

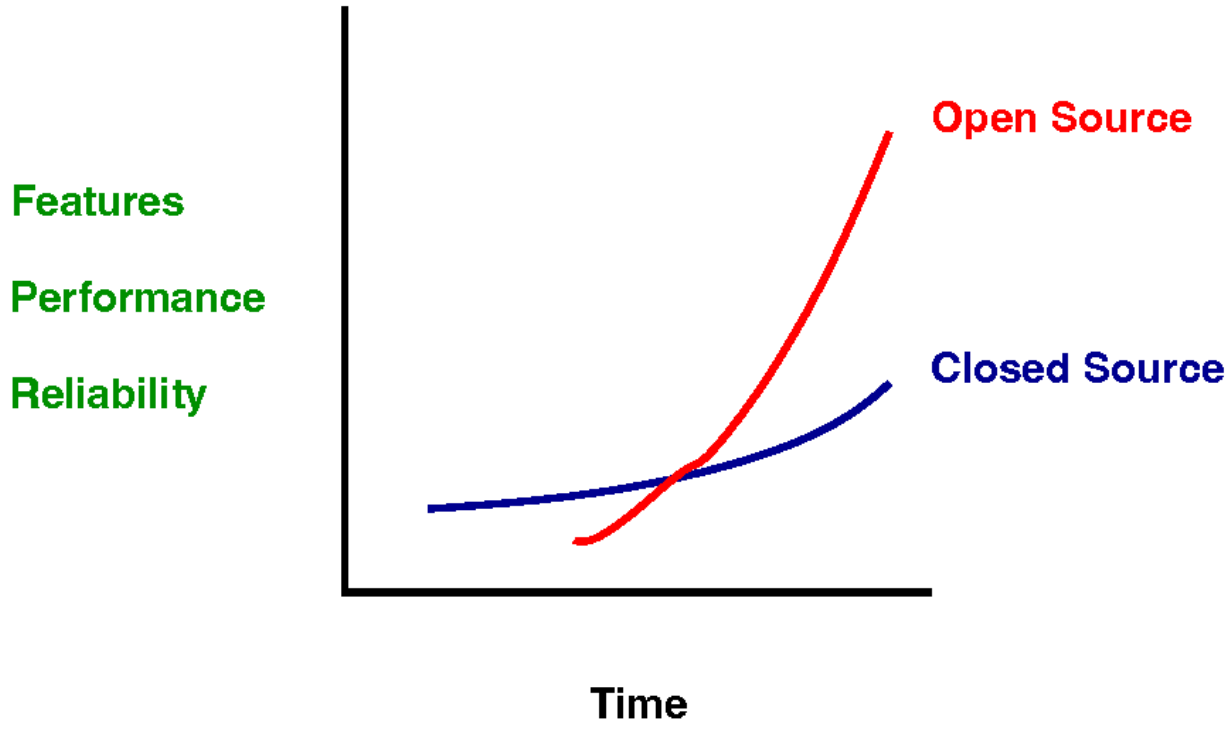
Giving Cloud Its Originally Intended Role of Commodity Infrastructure



PERCONA

Open Source and Proprietary

Rise of Open Source



https://momjian.us/main/writings/pgsql/forever.pdf

PERCONA

Kubernetes

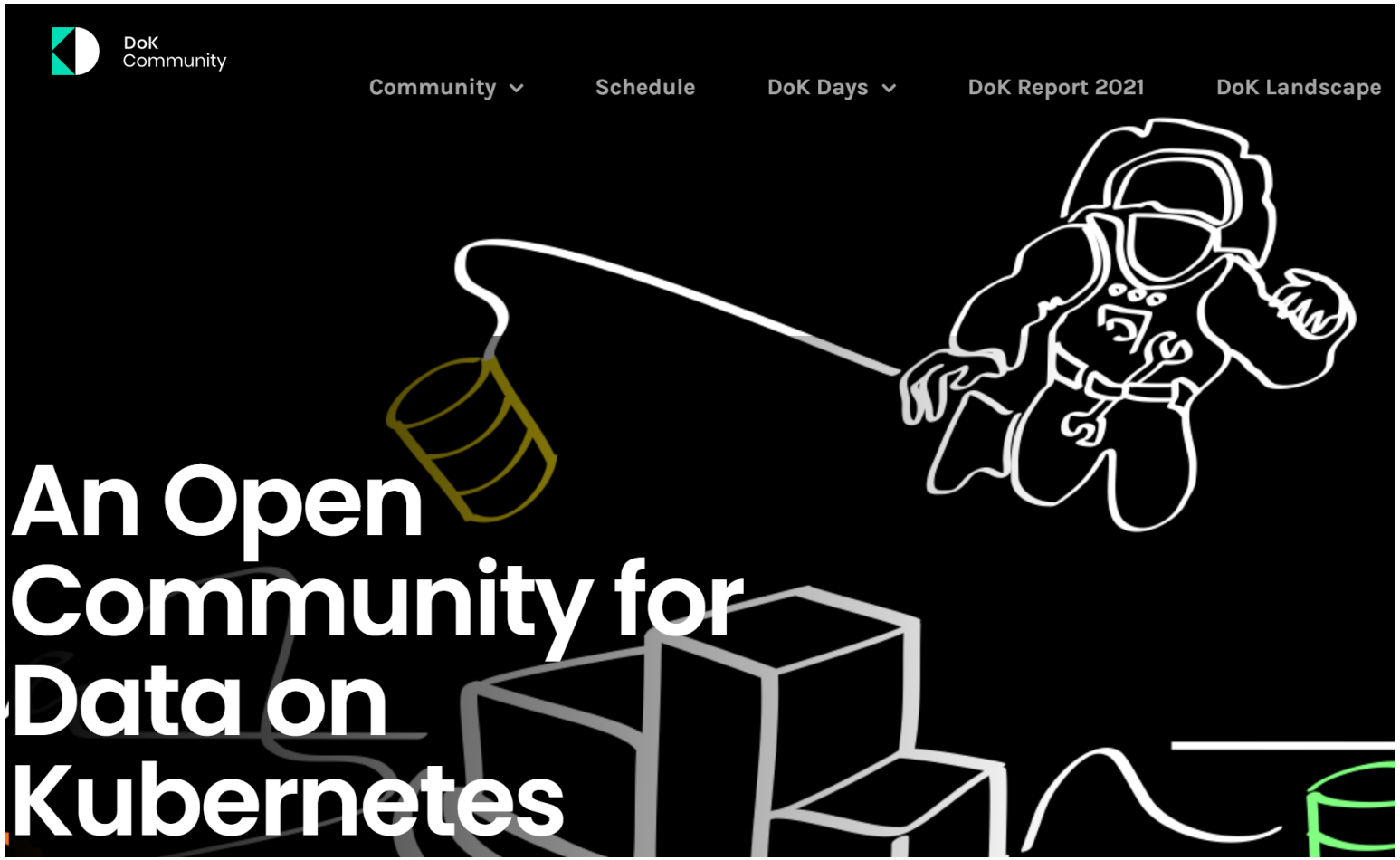
kubernetes

Kubernetes is universally available

Kubernetes is getting better for stateful applications

Kubernetes Operators are available for most popular Open Source Databases

Data on Kubernetes



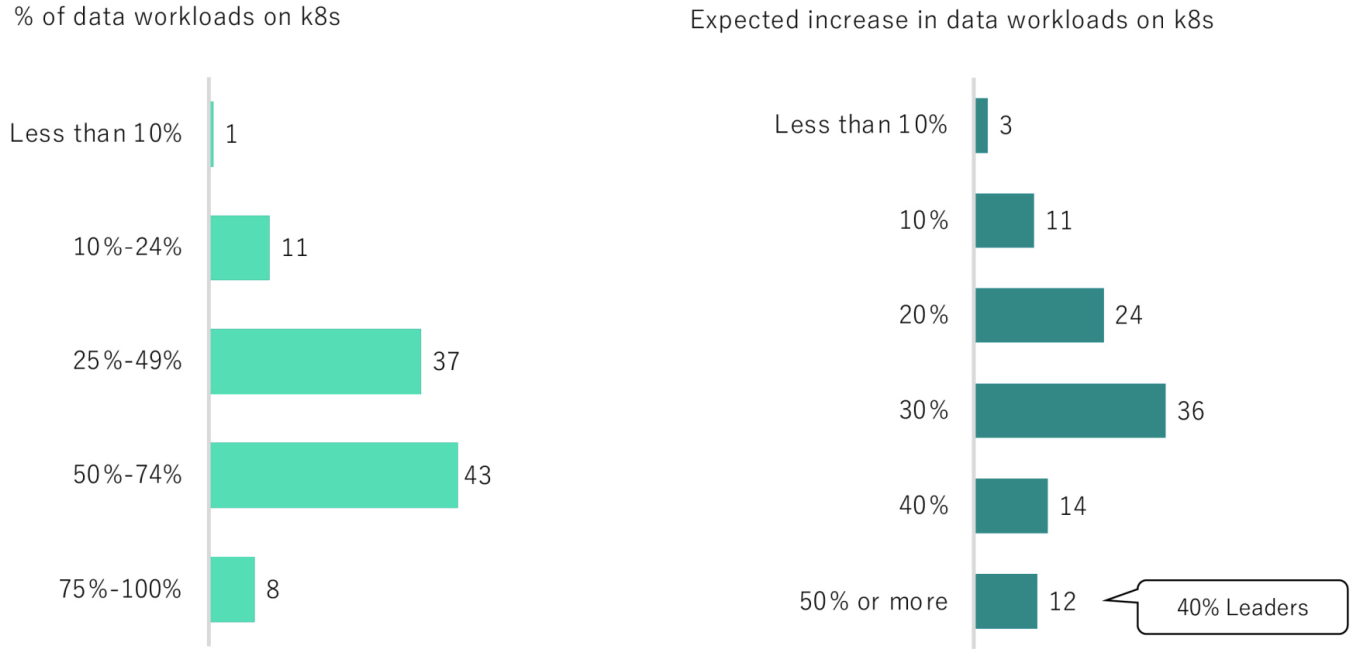
https://dok.community/

PERCONA

Data on Kubernetes

DoK workload %s are already high, and expected to increase

Leaders are chomping at the DoK bit



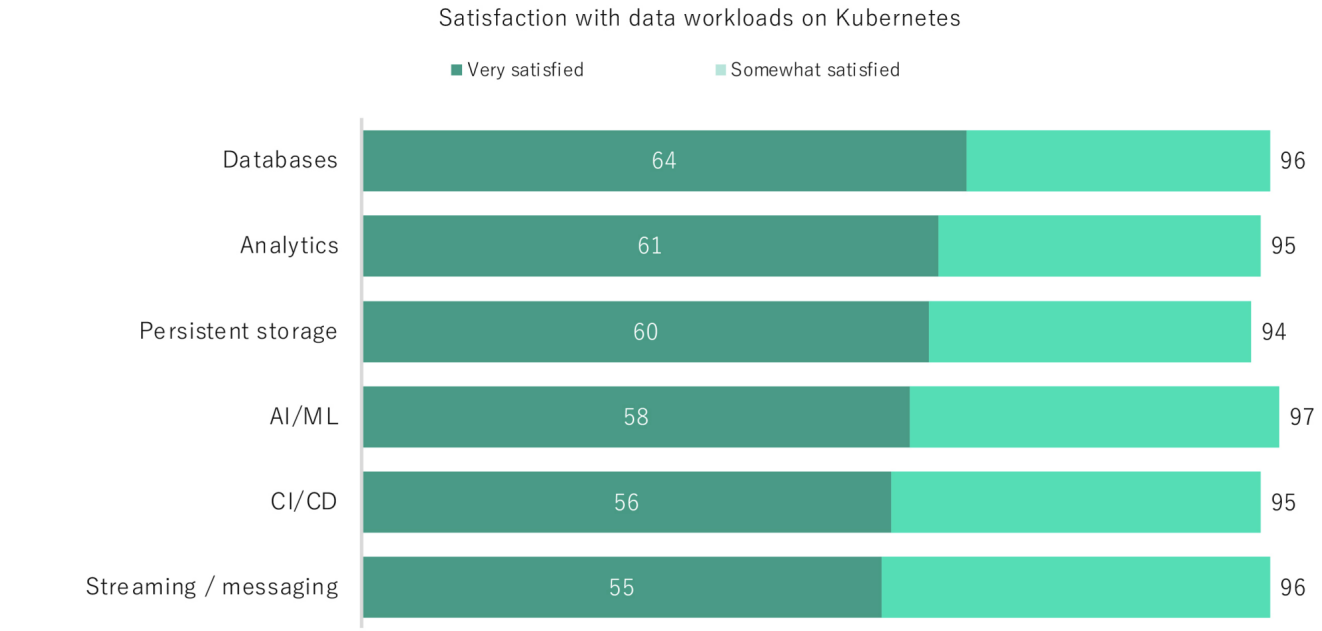
https://dok.community/wp-content/uploads/2022/10/DoK_Report_2022.pdf

PERCONA

And Happy about That...

Satisfaction = Reality - Expectations

DoK is winning the expectations battle



Q. In general, how satisfied are you with using Kubernetes to run each of the following data workloads in your organization? Use a scale from 1 to 5 where 5 means "very satisfied" and 1 means "not at all

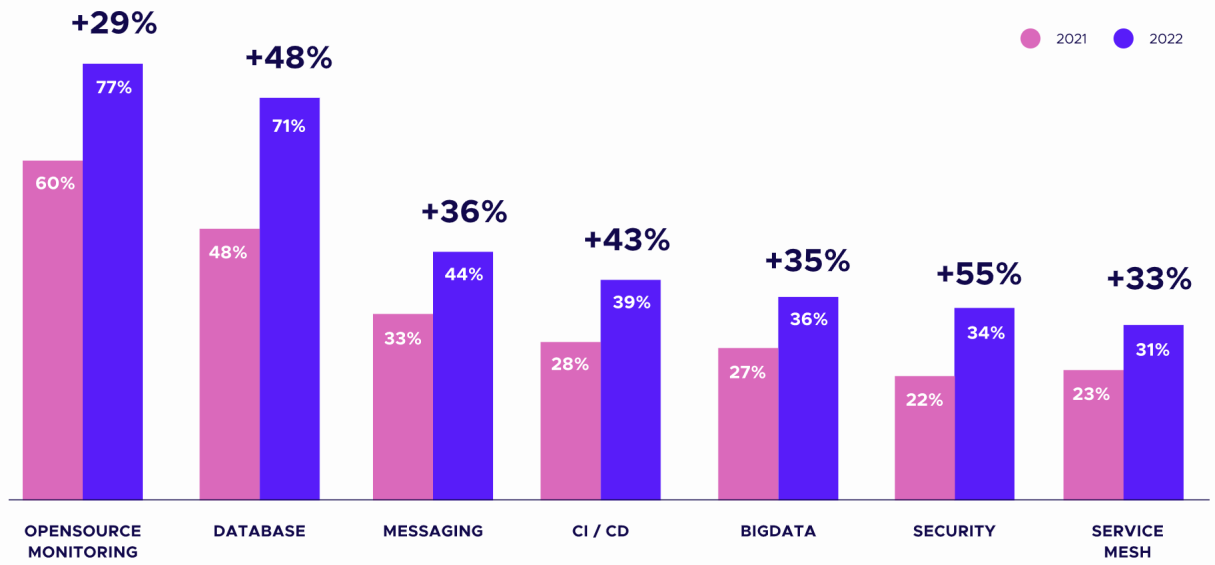
PERCONA

Quite a growth in Adoption!

KUBERNETES GROWTH AREAS

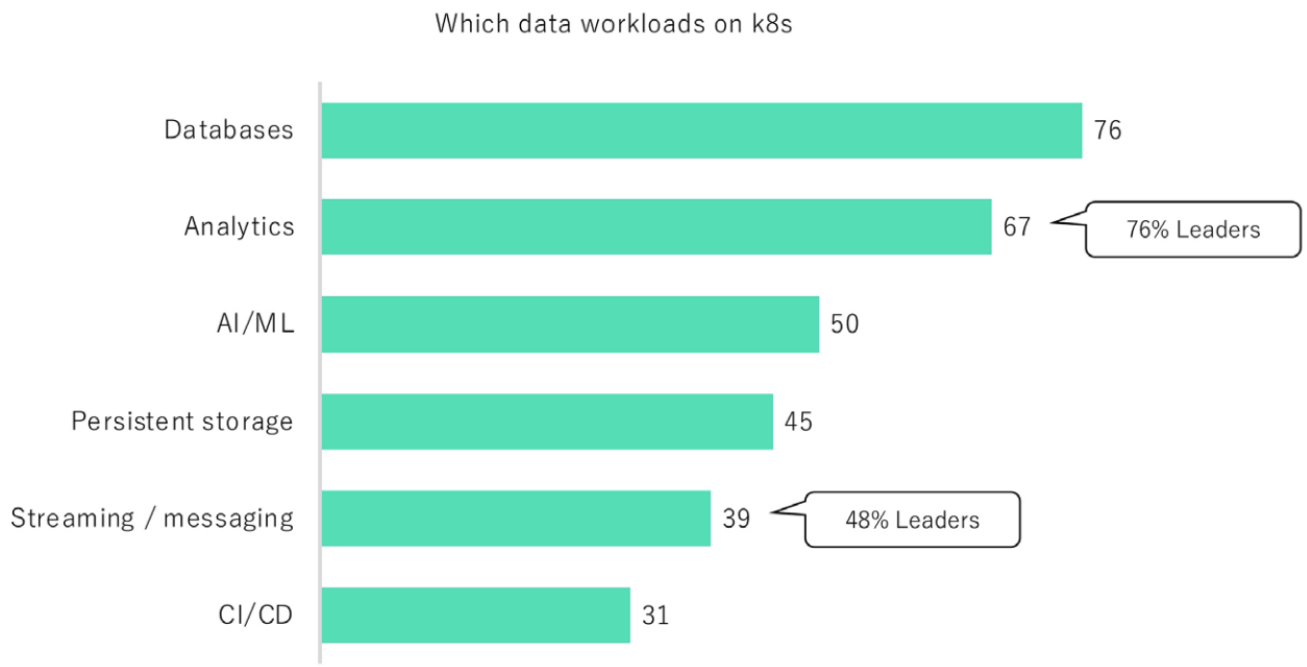
Focusing on non-application workloads, enterprises used an increasing variety of technologies. This reflects the need to enhance Kubernetes with better observability, security, and service-to-service communications. Other technologies enable specific use cases like CI/CD tools or databases.

Across all categories, **open source projects rank among the most frequently used solutions**.



https://www.cncf.io/reports/cncf-annual-survey-2022/

PERCONA



Databases on Kubernetes

Power of Kubernetes Operators

Power of Kubernetes Operators

- **Day One Automation**
  - Installation and Initial Configuration
- **Day Two Automation**
  - Backups, Scaling, Self Healing, Upgrades

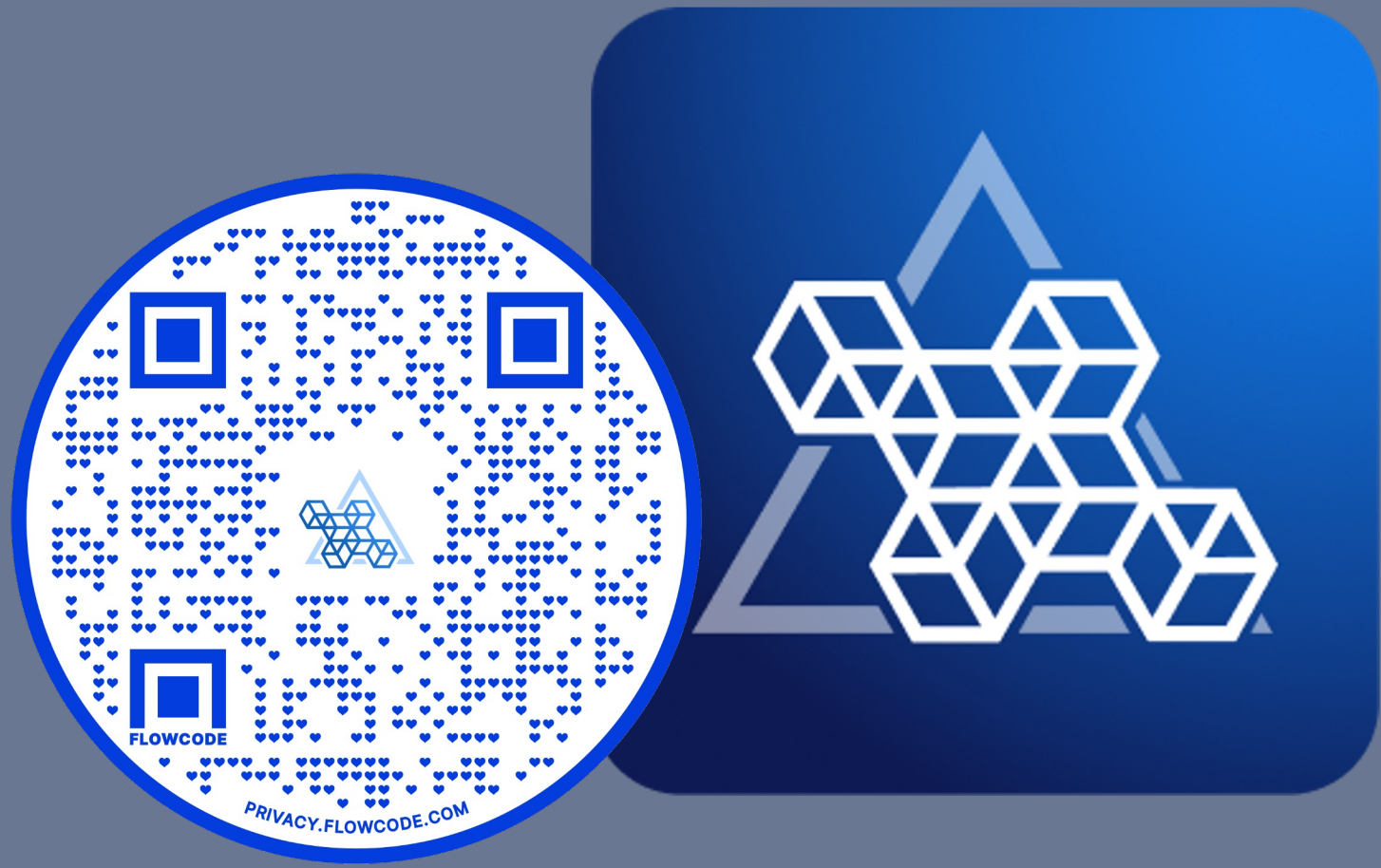
PERCONA

Percona Kubernetes Operators

Get look and feel for Basics ?

Kubernetes with Operators is Easy and Powerful

https://per.co.na/PXCMinkube



Many Modern DBaaS Are Built on Kubernetes Operators

Altinity

influxdata

MariaDB

crunchy data

planetscale

PERCONA

Best Practices

#1 Use Operators

For Production Deployment you need to ensure great “Day 2” Automation

PERCONA

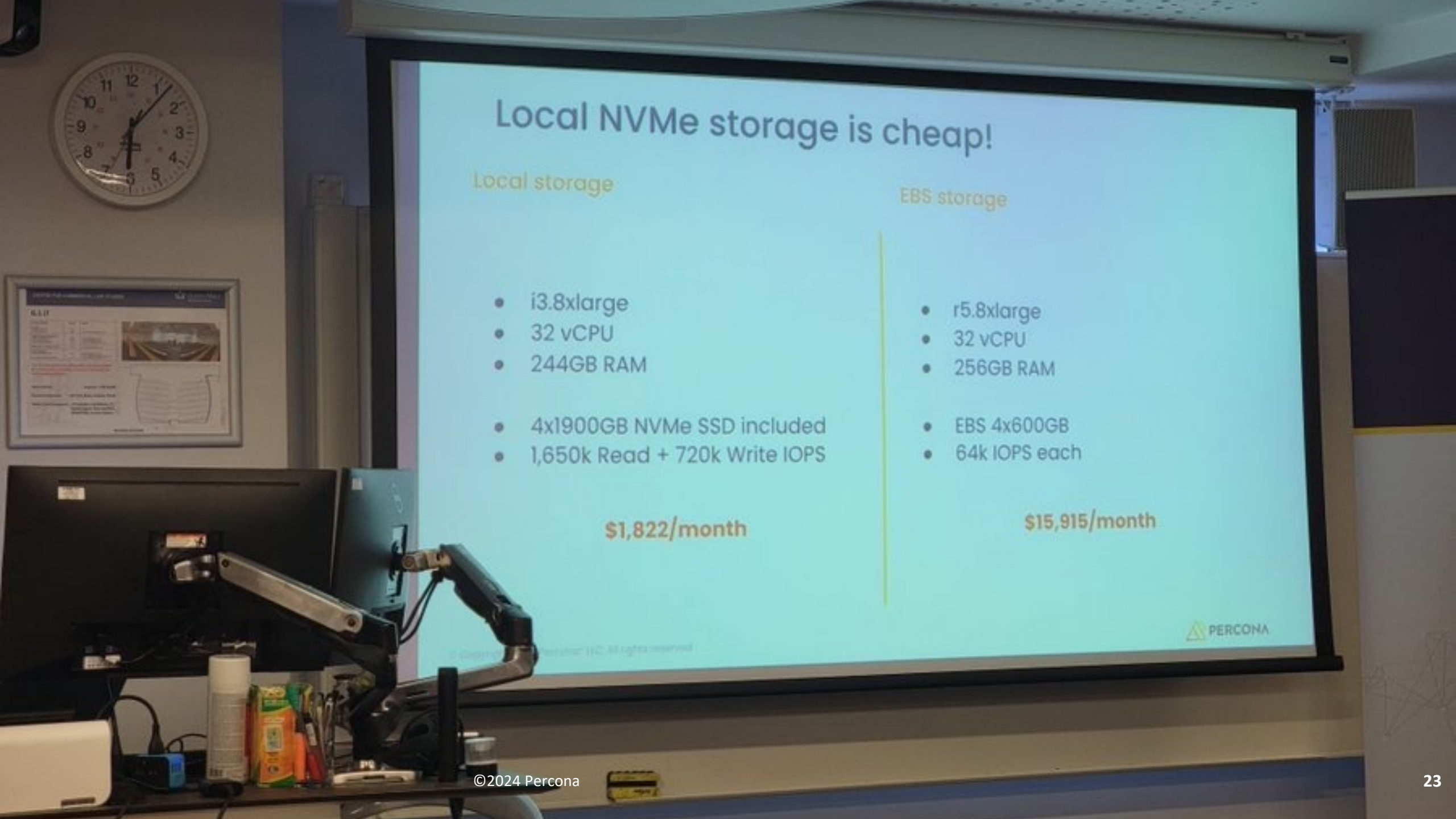
#2 Setup High Availability

Relying on Single Instance in Kubernetes environment is even more dangerous

#3 Keep Persistent Data Persistent

Persistent Volumes; Local Disk or Fast Remote Storage

PERCONA



#4 Keep Data per Pod Small

50TB of data connected to single POD is not a good idea

PERCONA

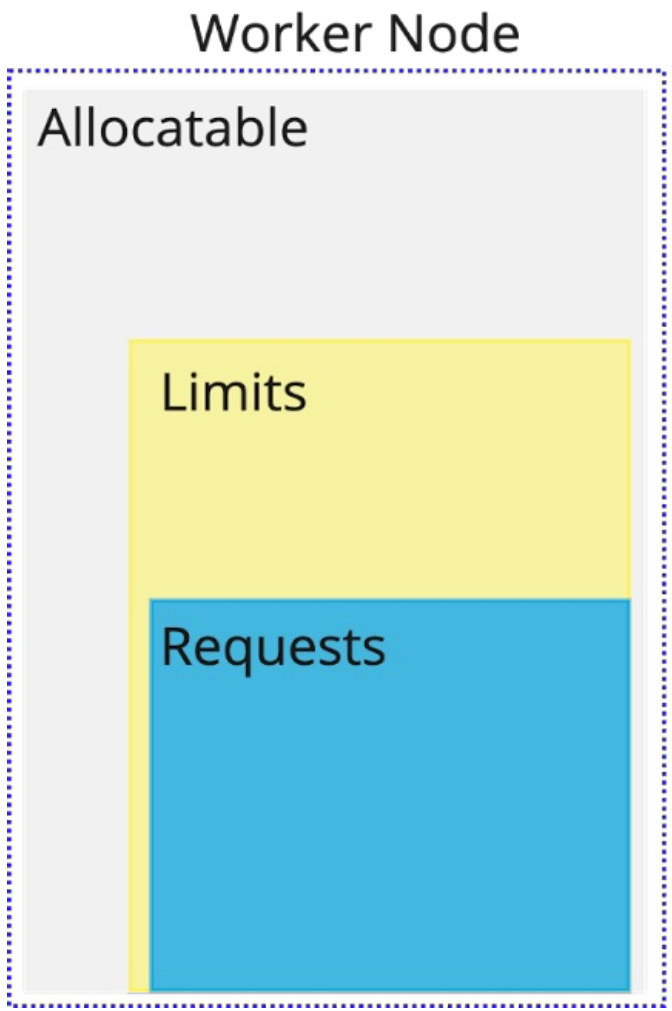
#5 Use Appropriate Node Sizes

Kubernetes or Not Databases often need “Big Iron” more than Apps

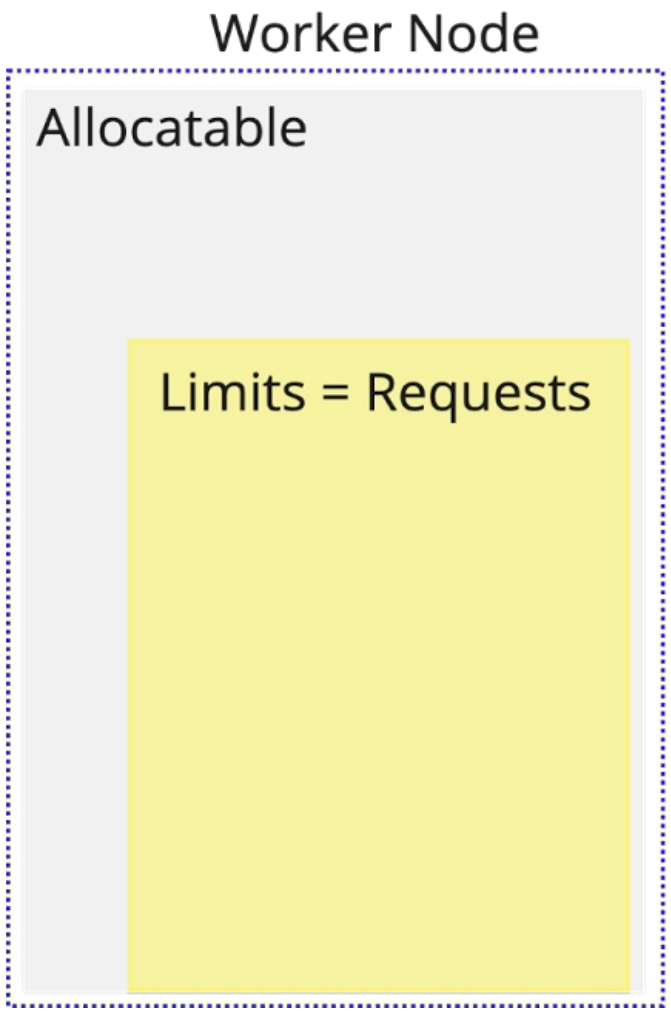
#6 Configure Resource Requests and Limits

Or you may have non uniform Performance and Severe Impact on other Pods

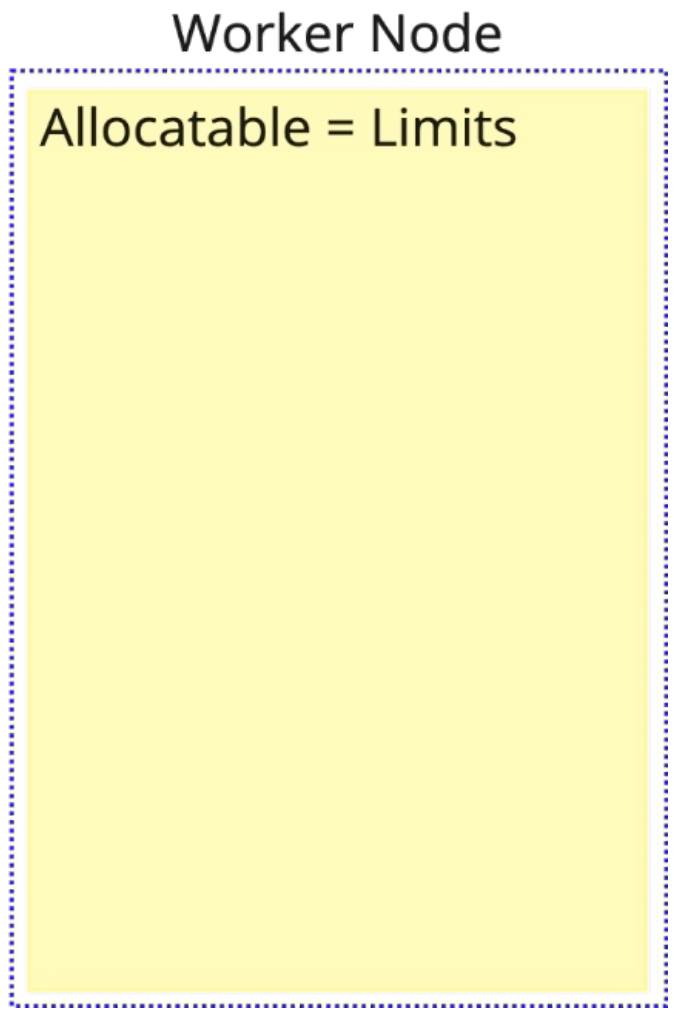
Shared resources



Guaranteed QoS



Dedicated node



PERCONA

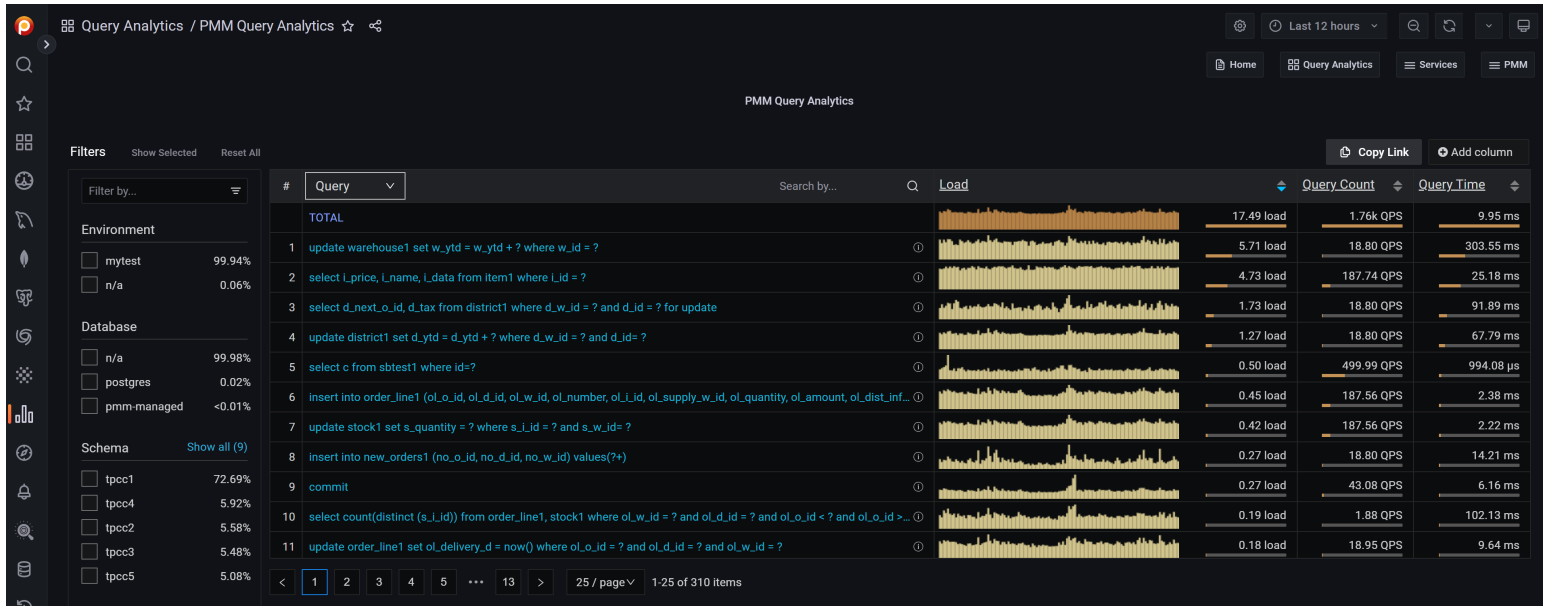
#7 Use Proper Anti-Affinity

3 Node Cluster Running on Single Physical Node is not Great High Availability Solution

#8 Tune your Database

Database Configuration, Indexes, Queries needs to be taken care of as usually

Query Analytics in Percona Monitoring and Management



PERCONA

#9
Understand
how to Scale

Some Databases can be “Scaled Out” others only “Scaled-Up” and Scaled for Reads

#10 Control Eviction with Pod Priority

Rescheduling Database Pod Can be Expensive, so better ensure it does not happen too often

PERCONA

#11 Do not Expose your Database unless you have to

Unintended Publicly Accessible Data is leading cause of Security Leaks

PERCONA

#12 Enable Encryption

Data at Rest and Data in Transit. Does not cause huge Overhead those days

#13 Use Kubernetes Secrets

Great way to pass database access credentials to your application

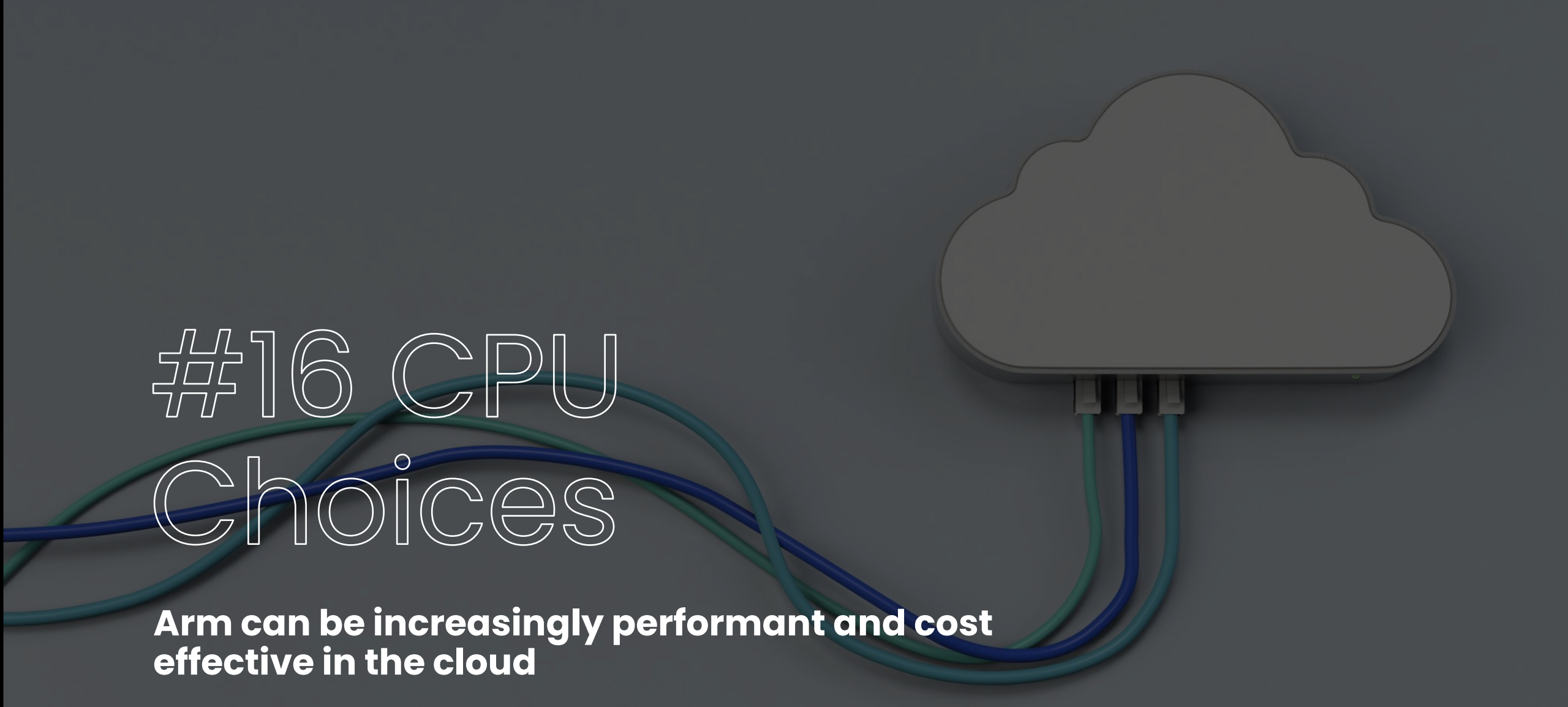
#14 Do not forget Backups

Clustering Does not Eliminate need for backups. Do them. Good Operators make it Easy

#15 Consider New Generation Databases

Databases designed to be run on Cloud Native Infrastructure are Coming - Vitess, Neon, Yugabyte, TiDB

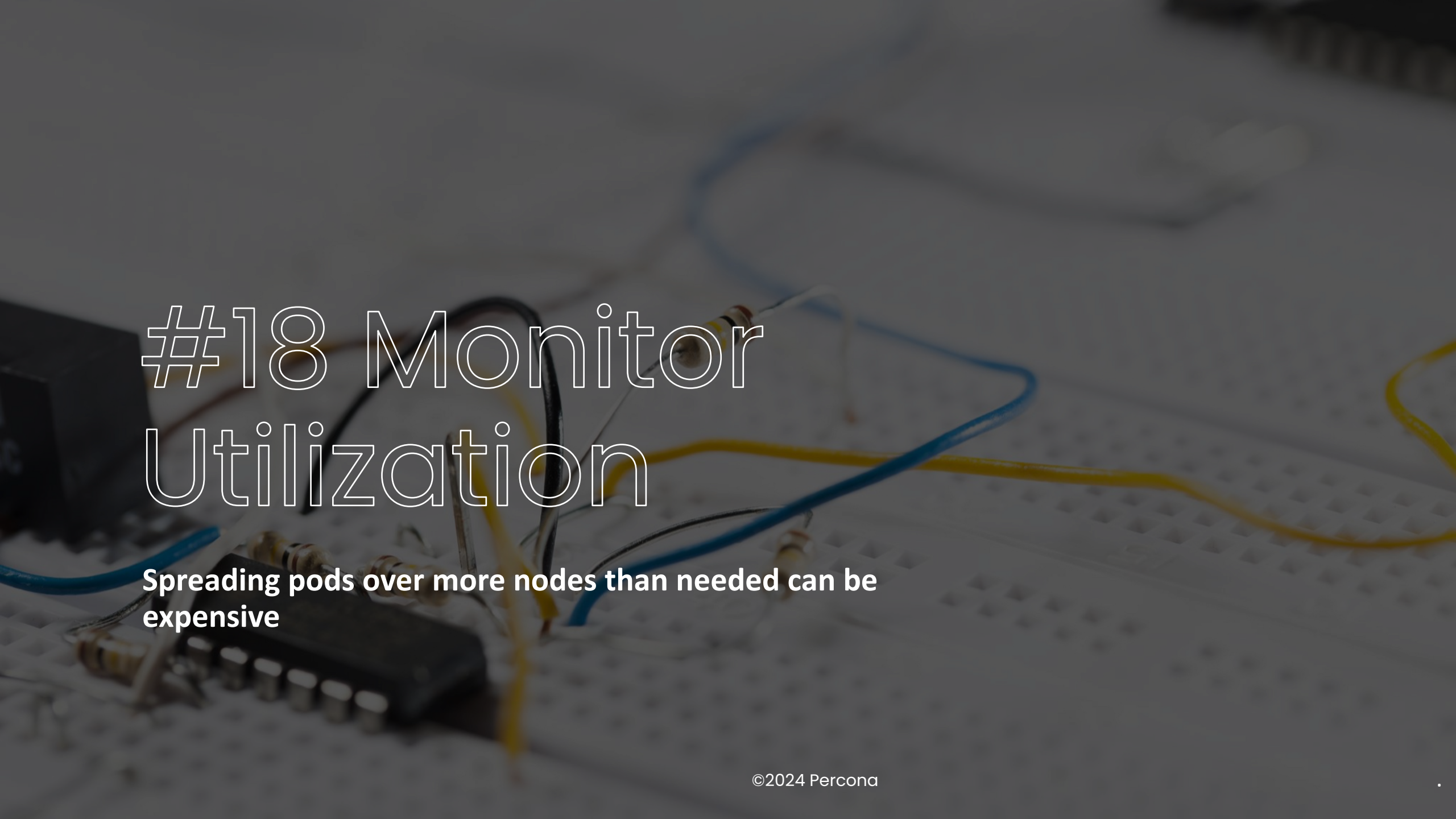
#16 CPU Choices



Arm can be increasingly performant and cost effective in the cloud

#17 Pick Right Kubernetes deployment mode

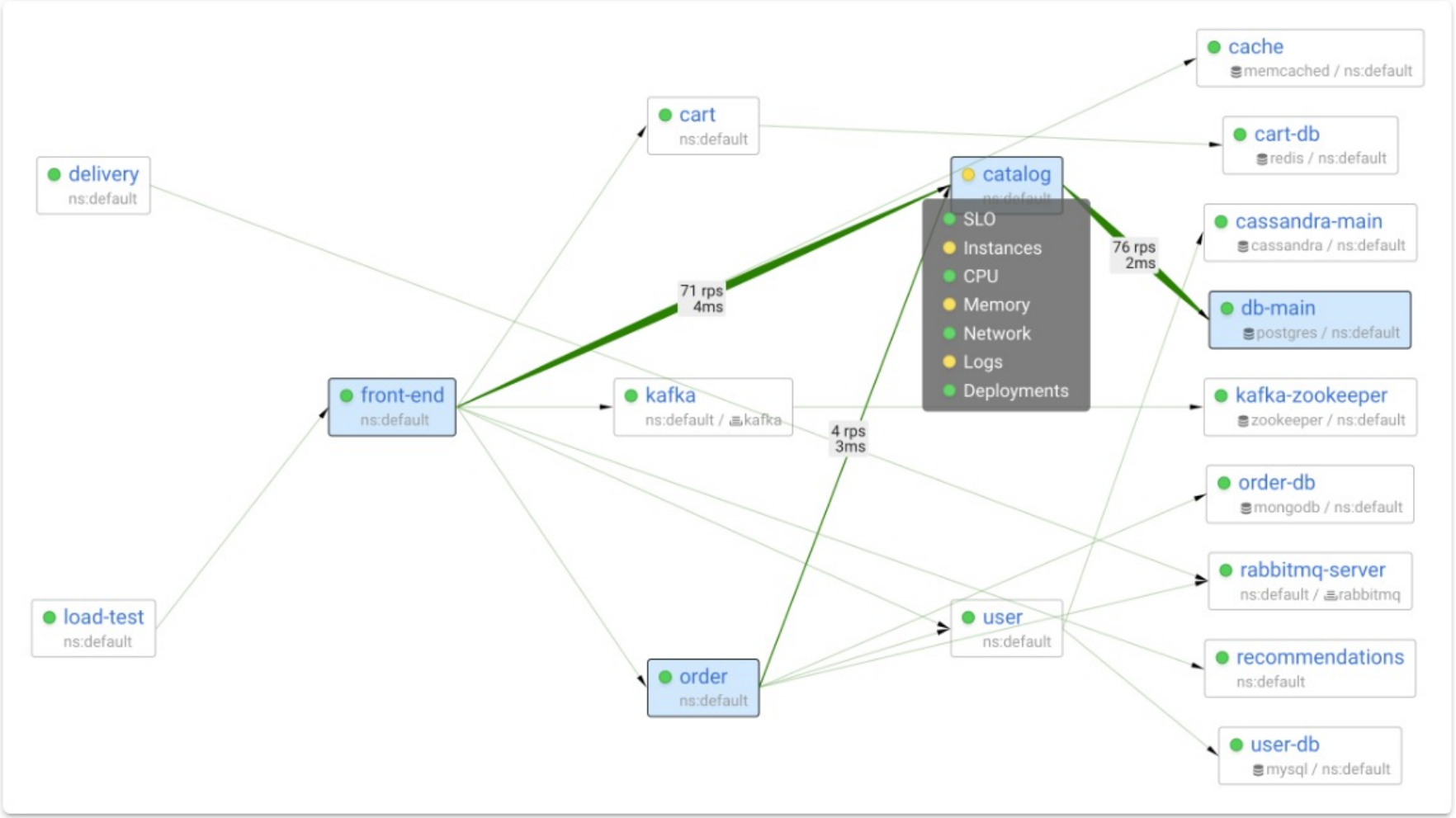
Managed and Self-Managed both have their merits



#18 Monitor Utilization

Spreading pods over more nodes than needed can be expensive

#19 Observe...



coroot :~#

https://coroot.com

PERCONA

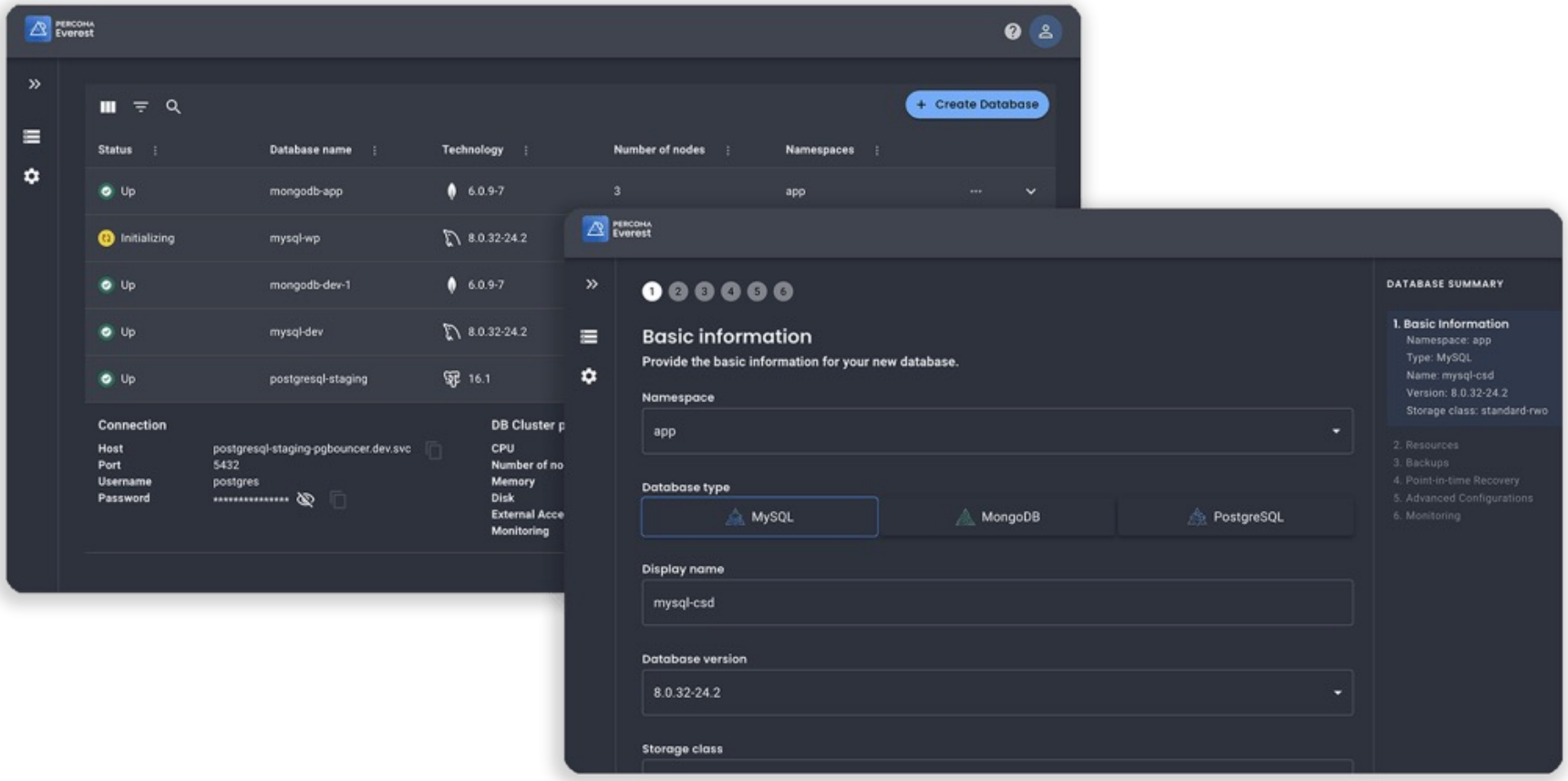
What are yours ?

Day 1 and Day 2 Automation, Toil Reduction Similar to DBaaS

UX is Different, Requires Kubernetes Expertise

Beyond Kubernetes Operators?

Percona Everest – Open Source DBaaS bases on Kubernetes



https://docs.percona.com/everest/index.html

PERCONA

Beyond Single Kubernetes Cluster

karmada

Star 3,981

Open, Multi-Cloud, Multi-Cluster Kubernetes Orchestration

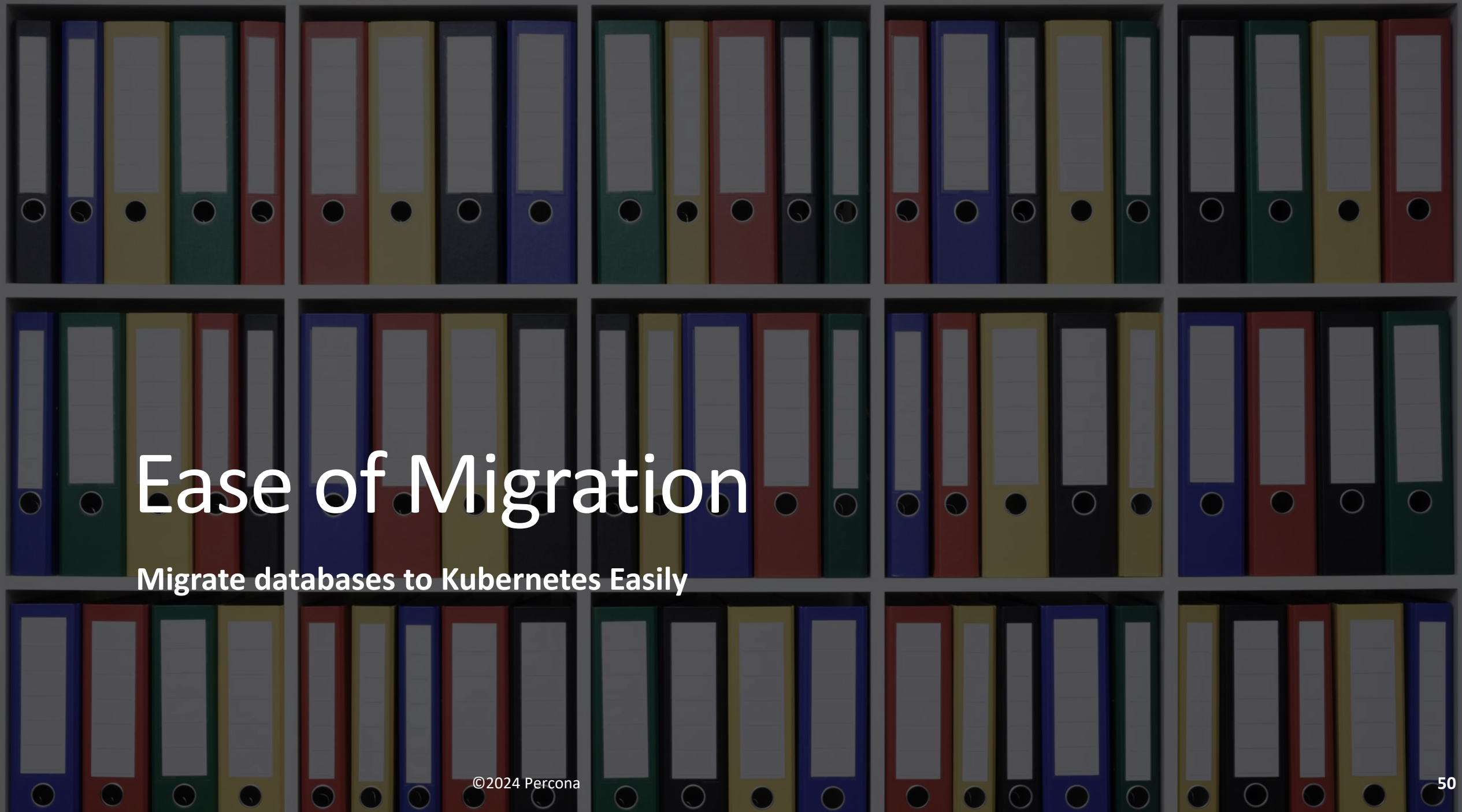
PERCONA

Automatic Volume Size Management

- **Can we automatically resize data volumes ?**

Auto Suspend
(For Development)

Suspend Idle Database Nodes, wake up when request comes in



Ease of Migration

Migrate databases to Kubernetes Easily

Thank you, Let's Connect!

https://www.linkedin.com/in/peterzaitsev/

https://twitter.com/PeterZaitsev

http://www.peterzaitsev.com