
mittwald.
1001 WAYS to
SHOOT YOURSELF IN THE FOOT
with ON-PREM KUBERNETES
MARTIN HELMICH @mittwald
SLAC2024 BERLIN
May 6th, 2024
CC-BY Anja Vrečko

mittwald.

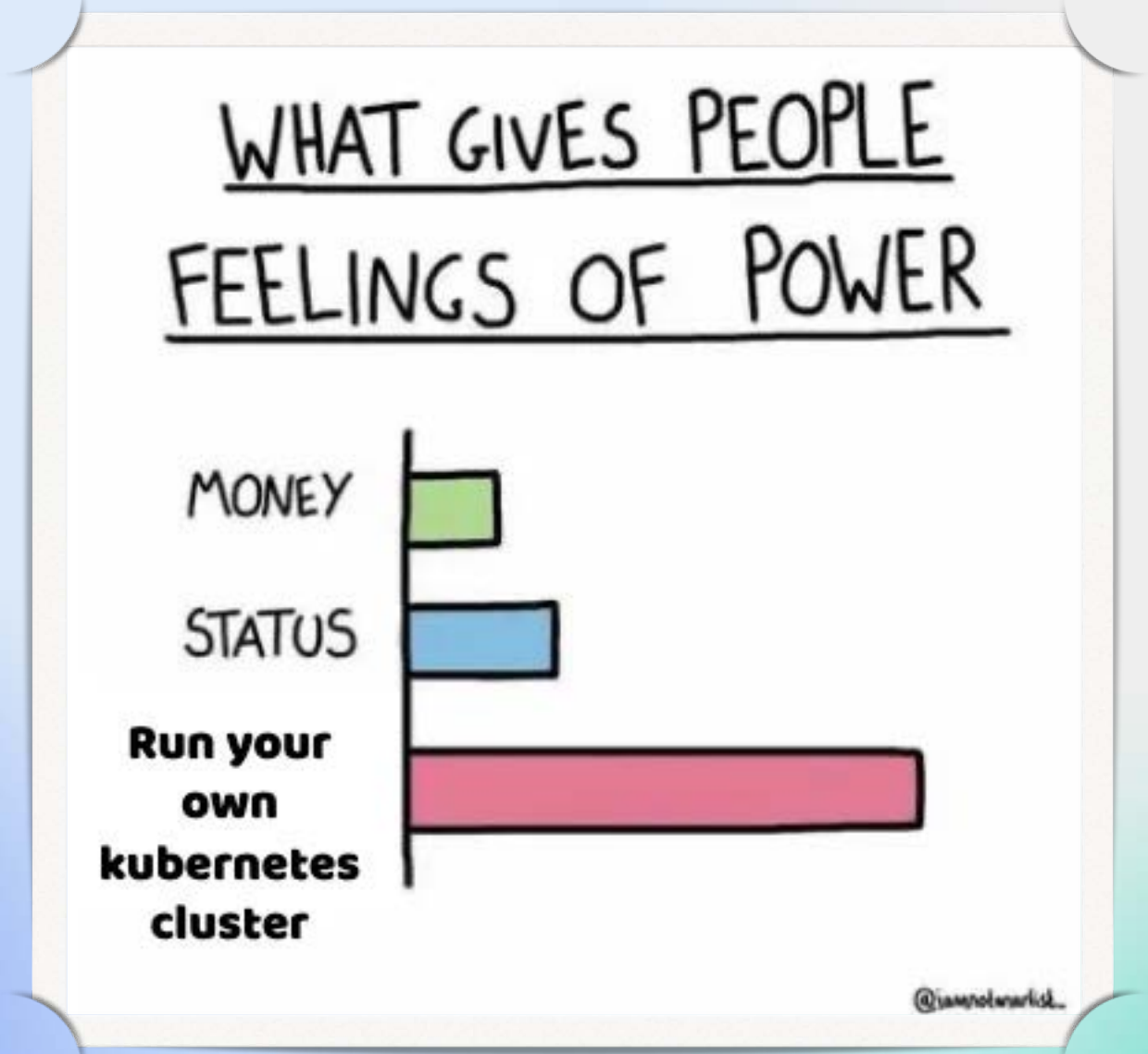
MARTIN HELMICH

Head of Architecture &
Developer Relations

Lecturer, Software Engineering
& Cloud Computing

Sci-Fi-Nerd, Metalhead,
Amateur Woodworker

mittwald.

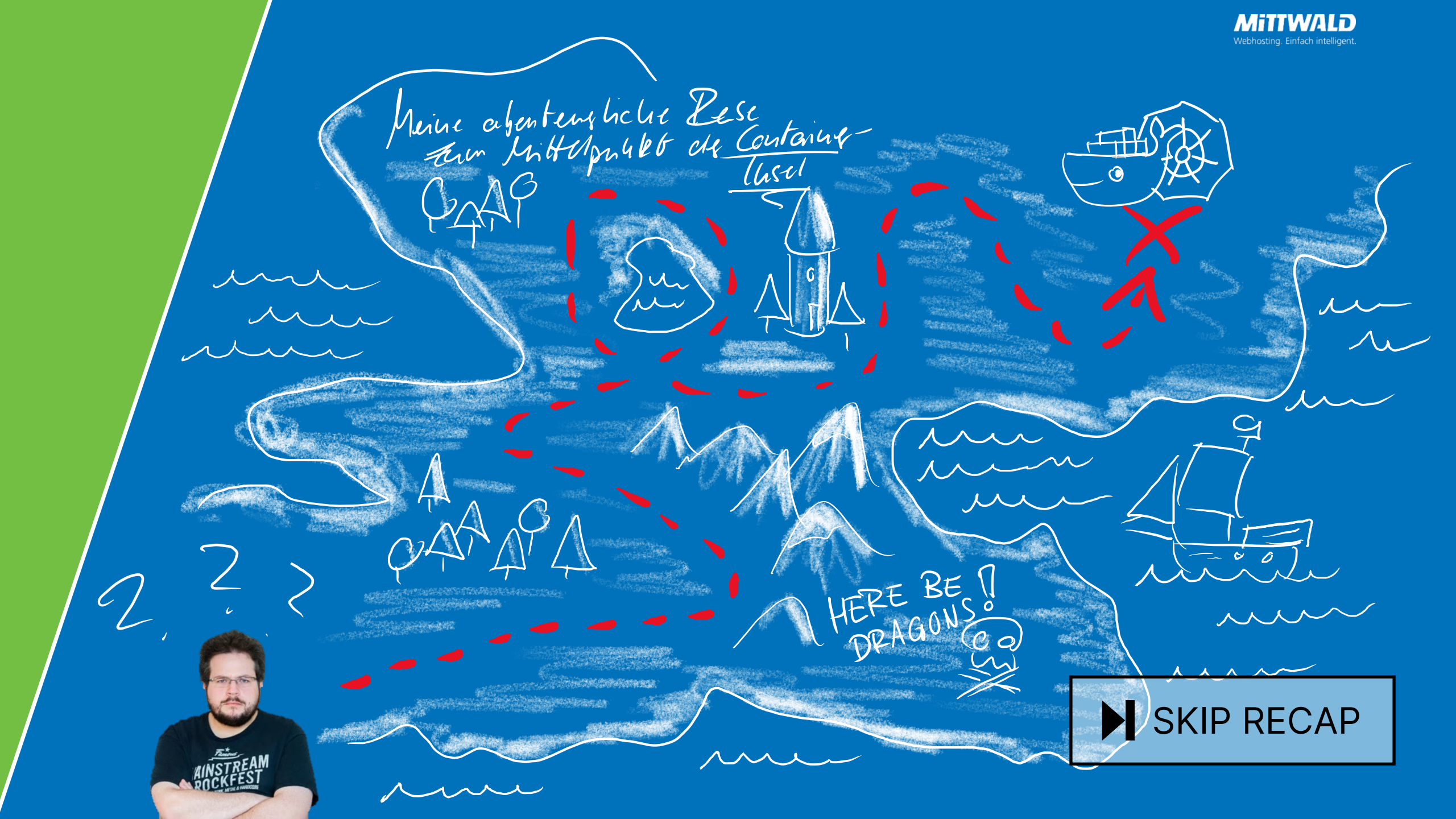

WHAT GIVES PEOPLE FEELINGS OF POWER
MONEY
STATUS
Run your own kubernetes cluster

mittwald.

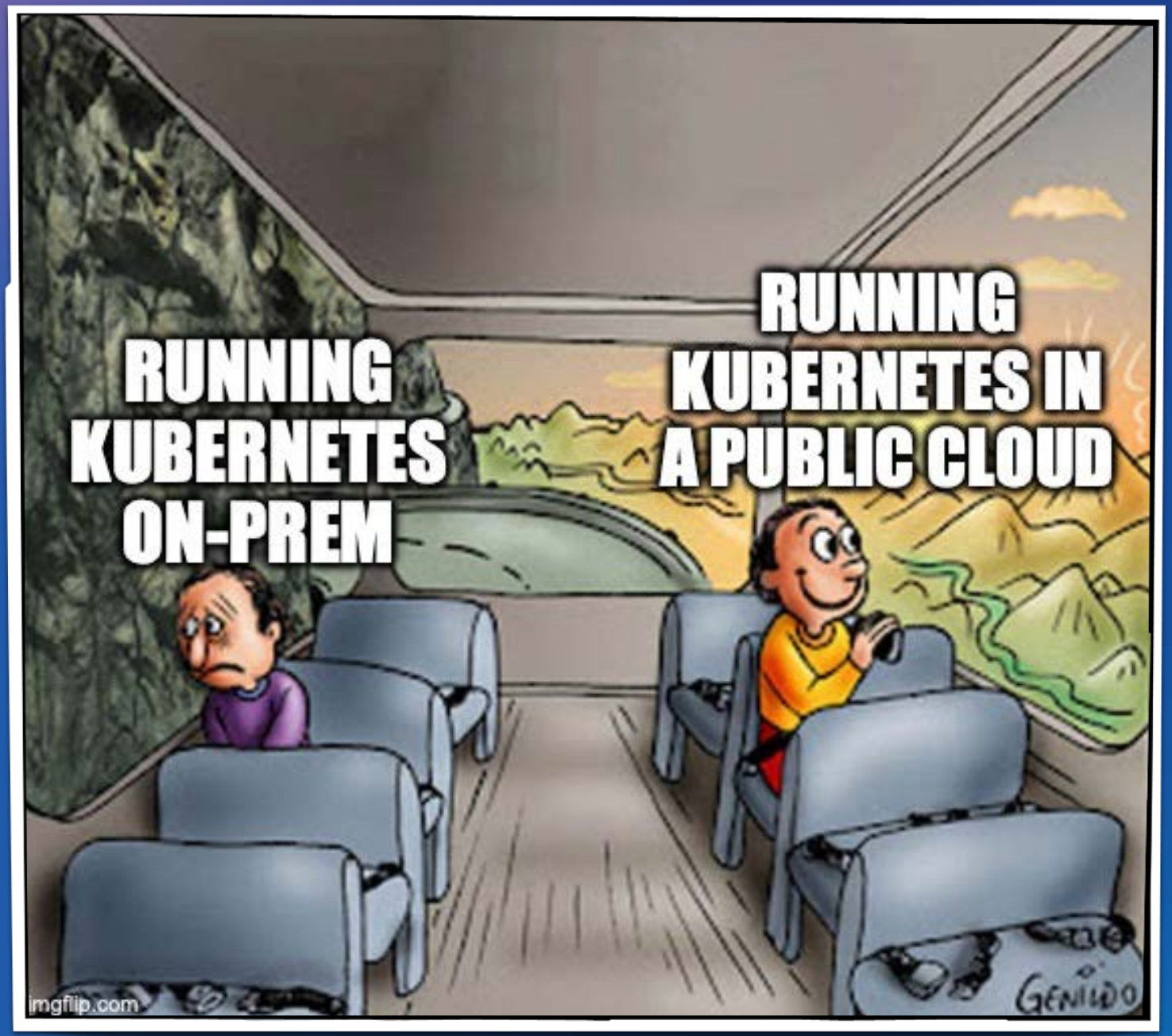
PREVIOUSLY ON

MARTINS ADVENTURES WITH KUBERNETES

SKIP RECAP



mittwald.


RUNNING KUBERNETES ON-PREM
RUNNING KUBERNETES IN A PUBLIC CLOUD
imgflip.com
GENIUS

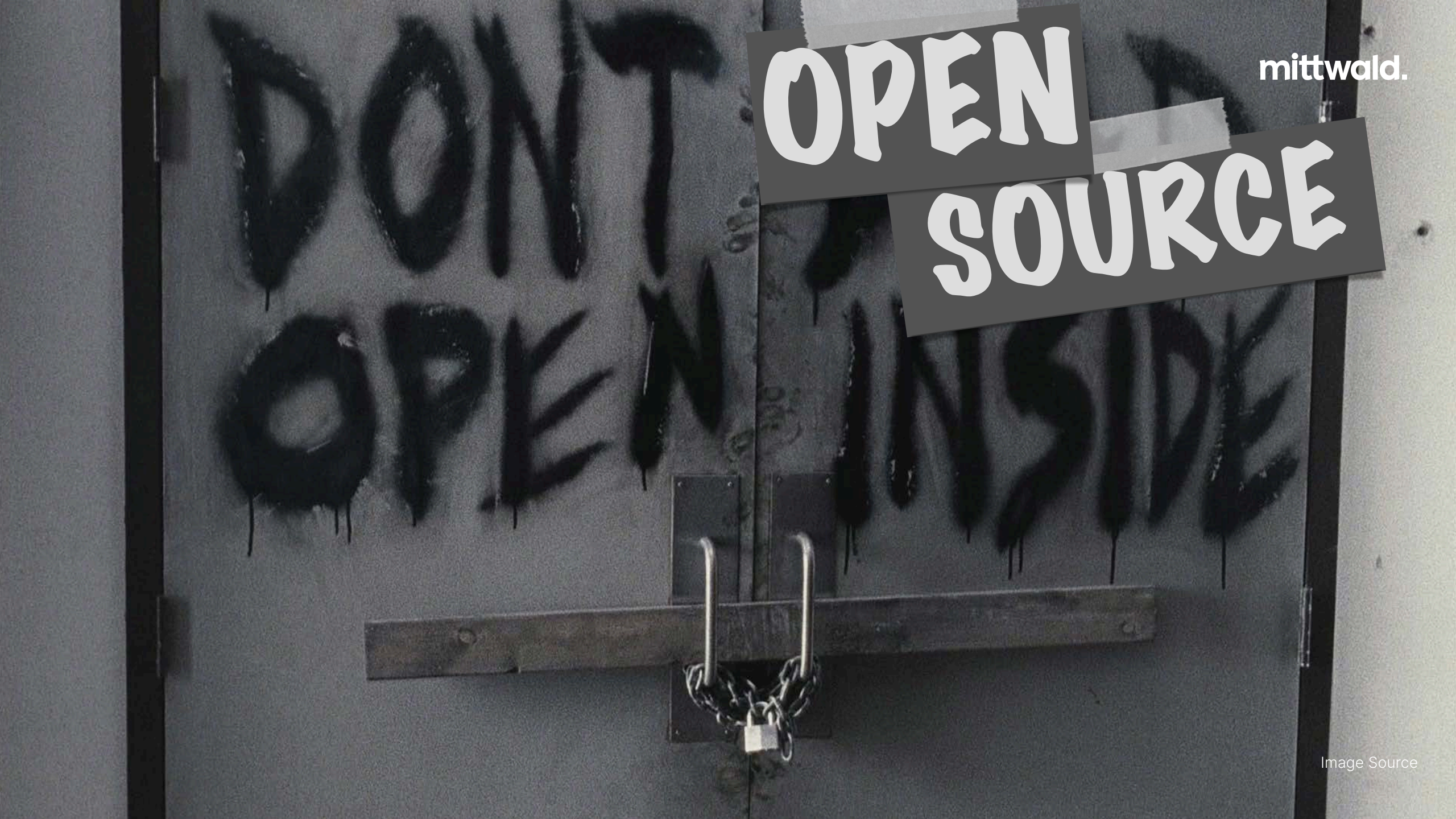
mittwald.

Content Warning:

The following PowerPoint presentation contains material that may be harmful or traumatizing to some audiences.

mittwald.

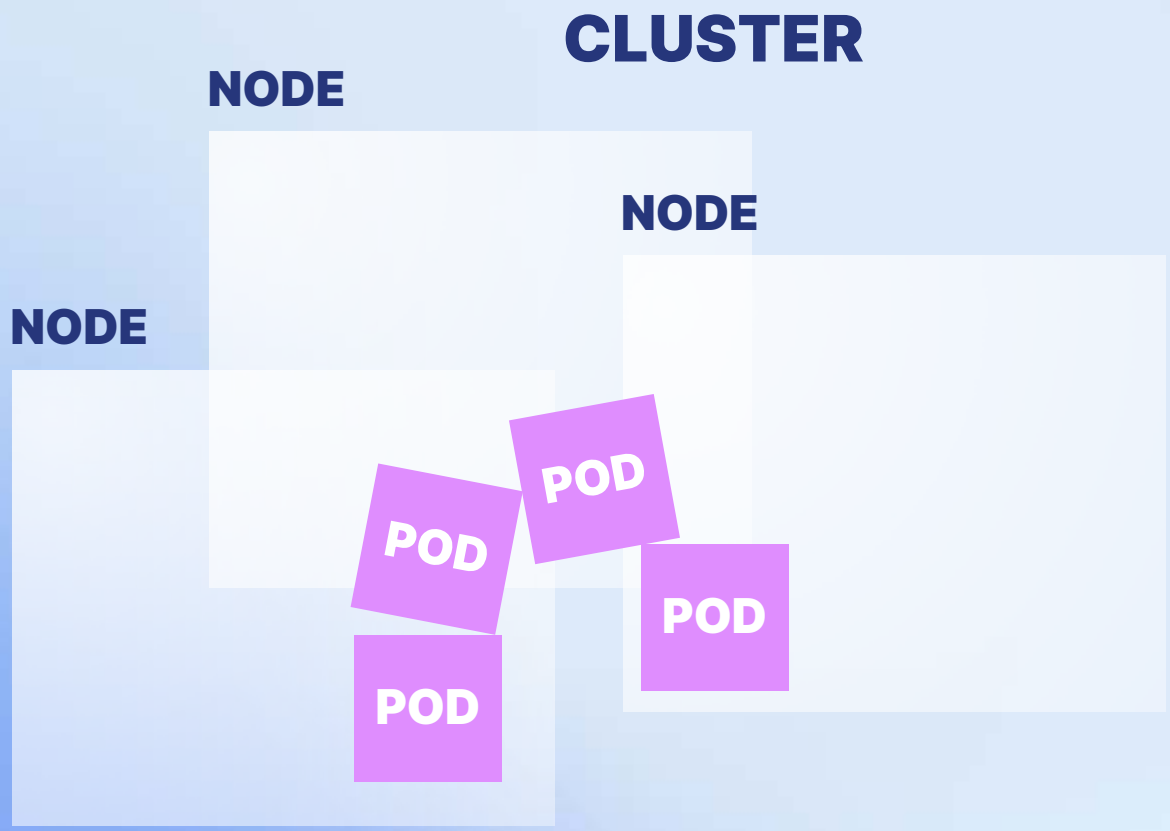
YOUR MILEAGE MAY VARY

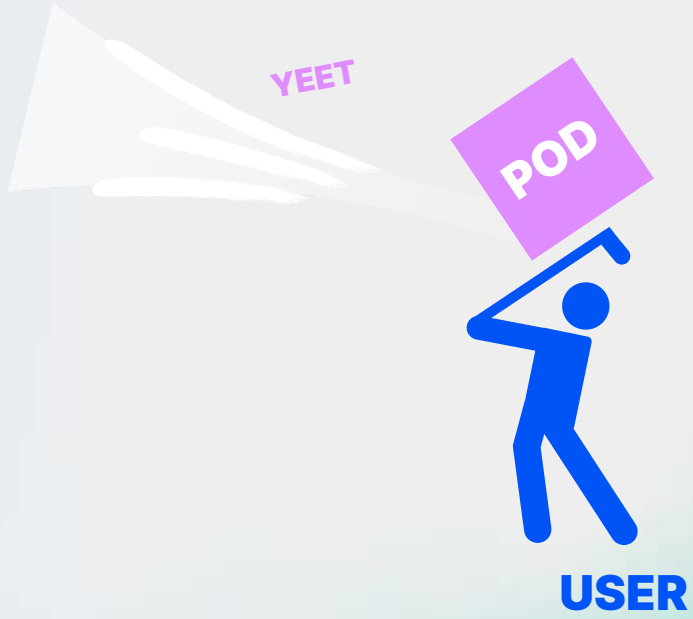


mittwald.

Image Source

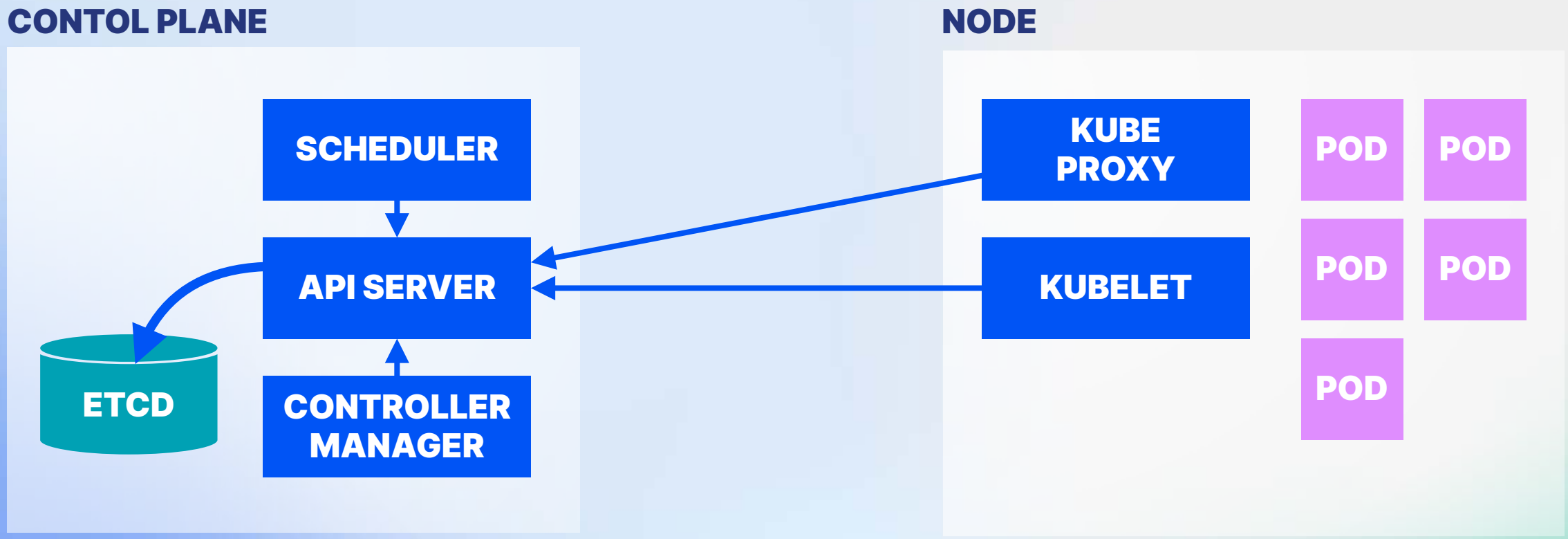
mittwald.



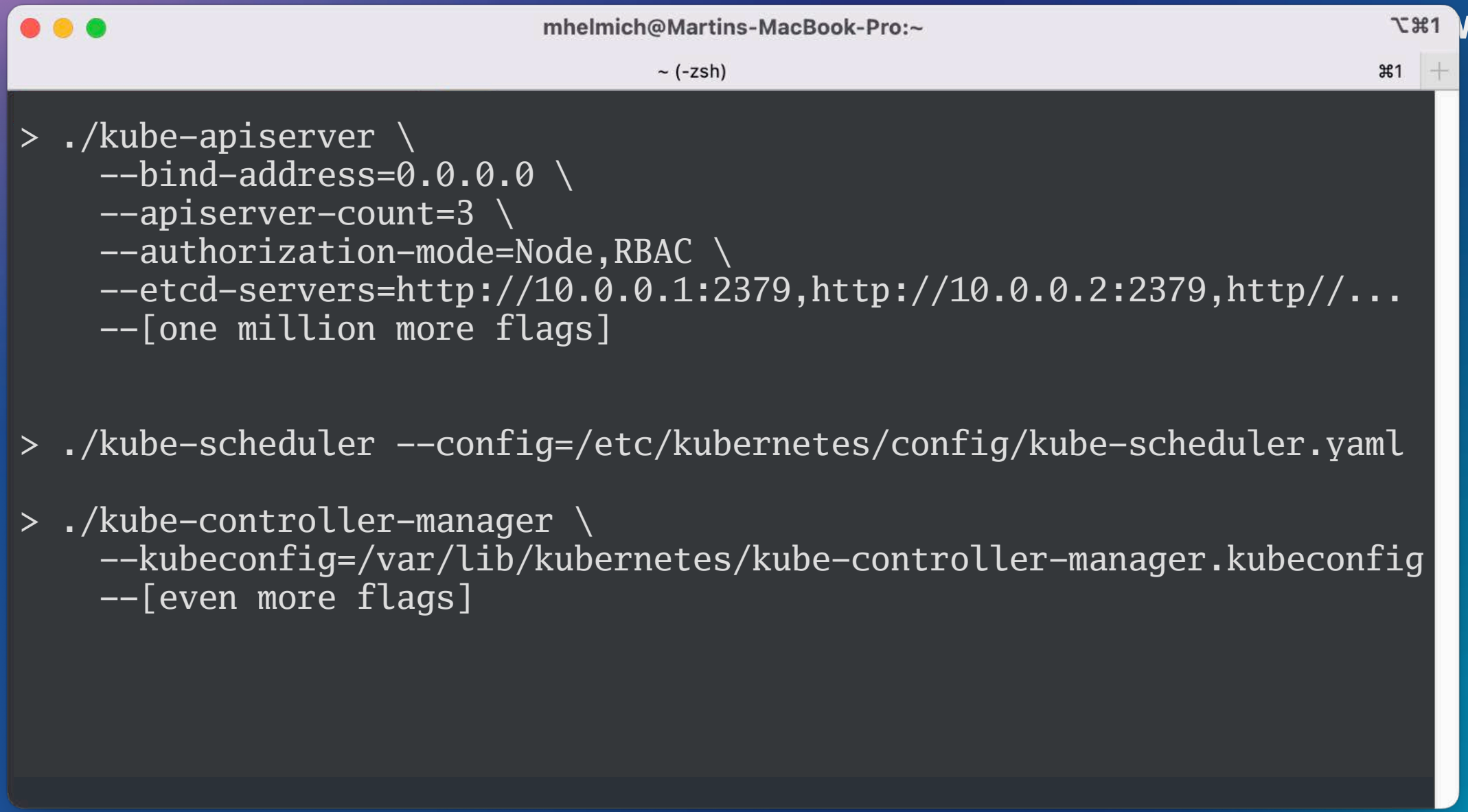


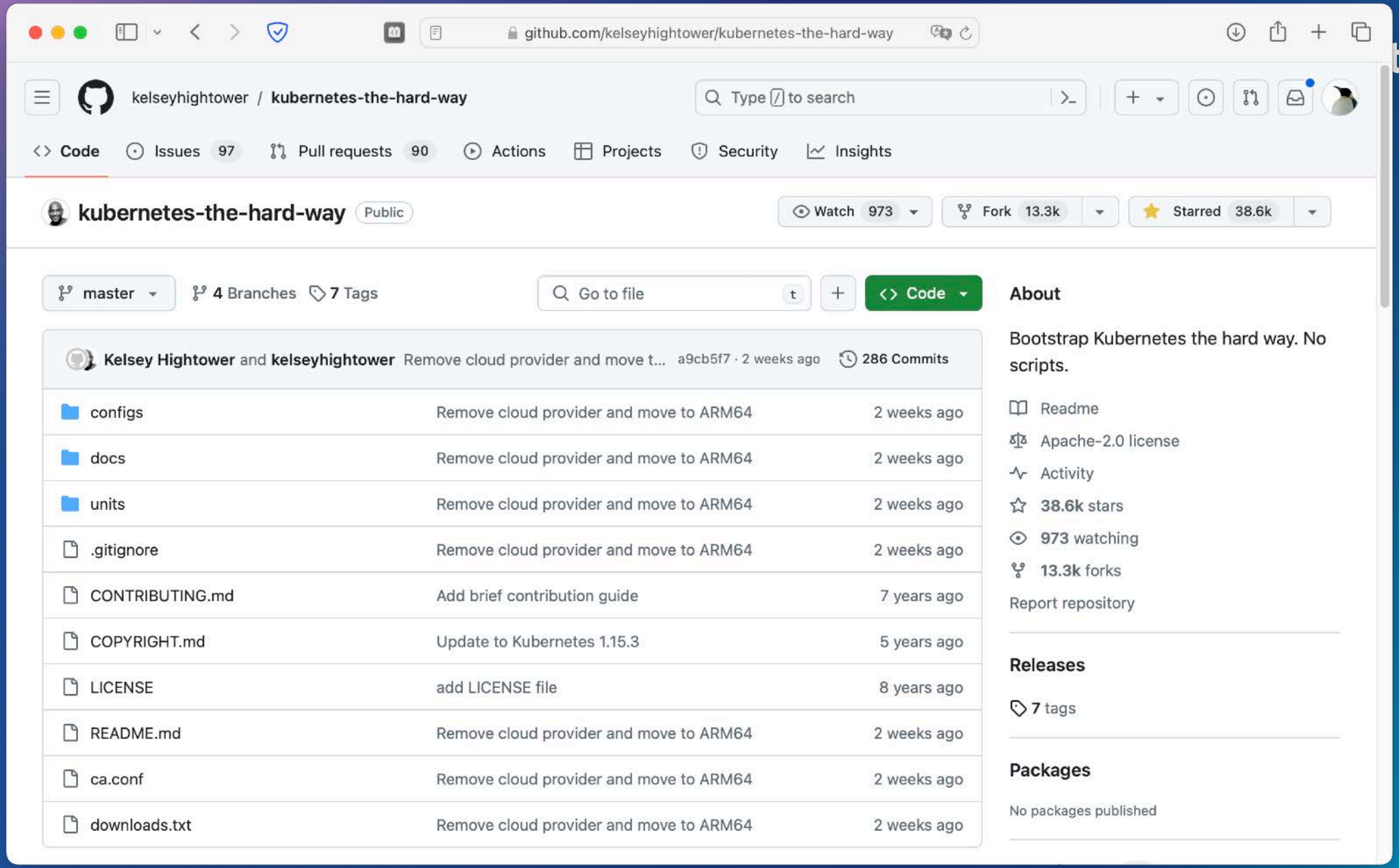
BASIC KUBERNETES PRINCIPLES (SLIGHTLY SIMPLIFIED)

mittwald.

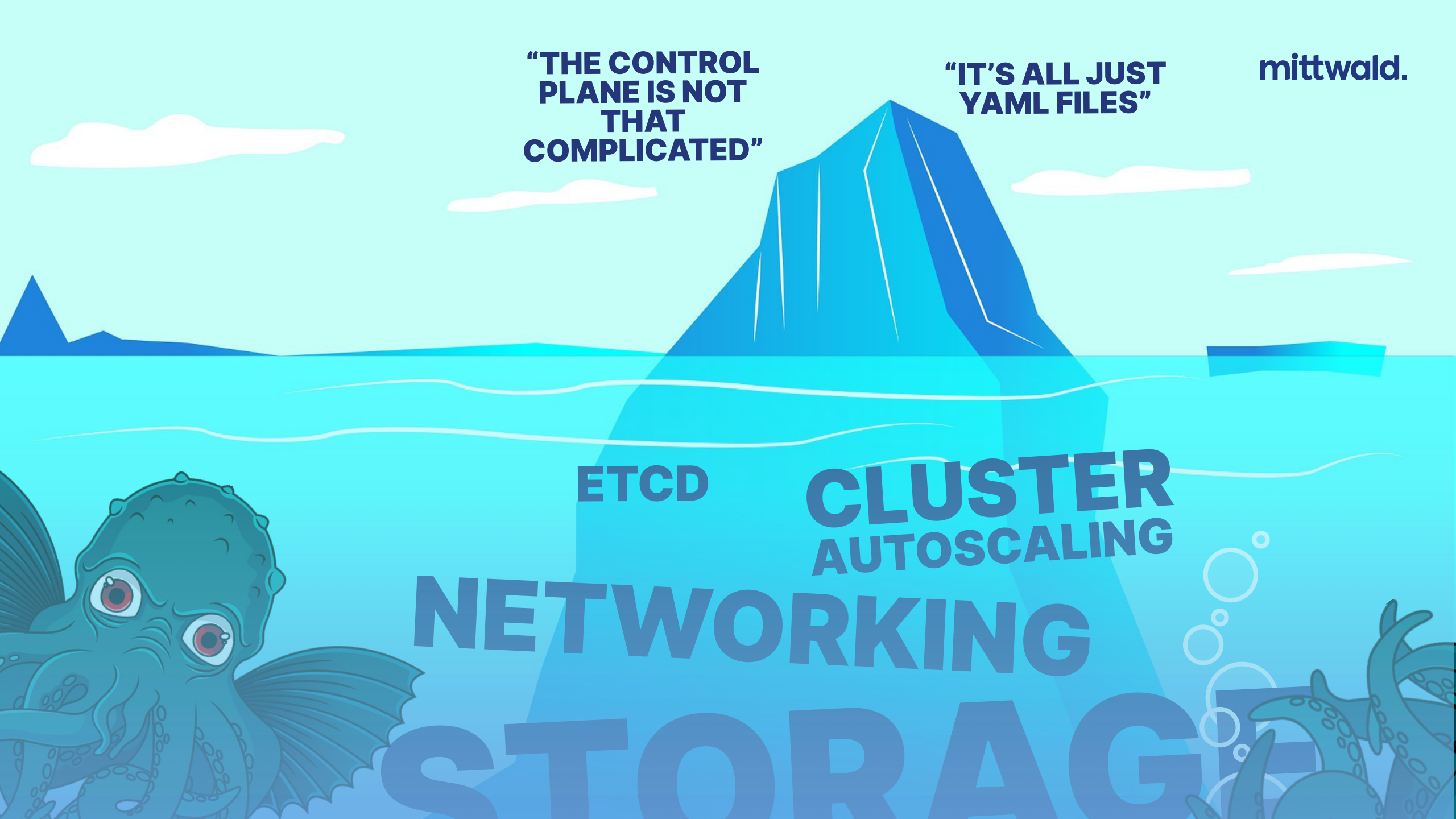


BASIC KUBERNETES ARCHITECTURE


mhelmich@Martins-MacBook-Pro:~
⌥⌘1
~ (-zsh)
⌘1
> ./kube-apiserver \
--bind-address=0.0.0.0 \
--apiserver-count=3 \
--authorization-mode=Node,RBAC \
--etcd-servers=http://10.0.0.1:2379,http://10.0.0.2:2379,http//...
--[one million more flags]
> ./kube-scheduler --config=/etc/kubernetes/config/kube-scheduler.yaml
> ./kube-controller-manager \
--kubeconfig=/var/lib/kubernetes/kube-controller-manager.kubeconfig
--[even more flags]



https://github.com/kelseyhightower/kubernetes-the-hard-way


“THE CONTROL PLANE IS NOT THAT COMPLICATED”
“IT’S ALL JUST YAML FILES”
mittwald.
ETCD
CLUSTER AUTOSCALING
NETWORKING
STORAGE

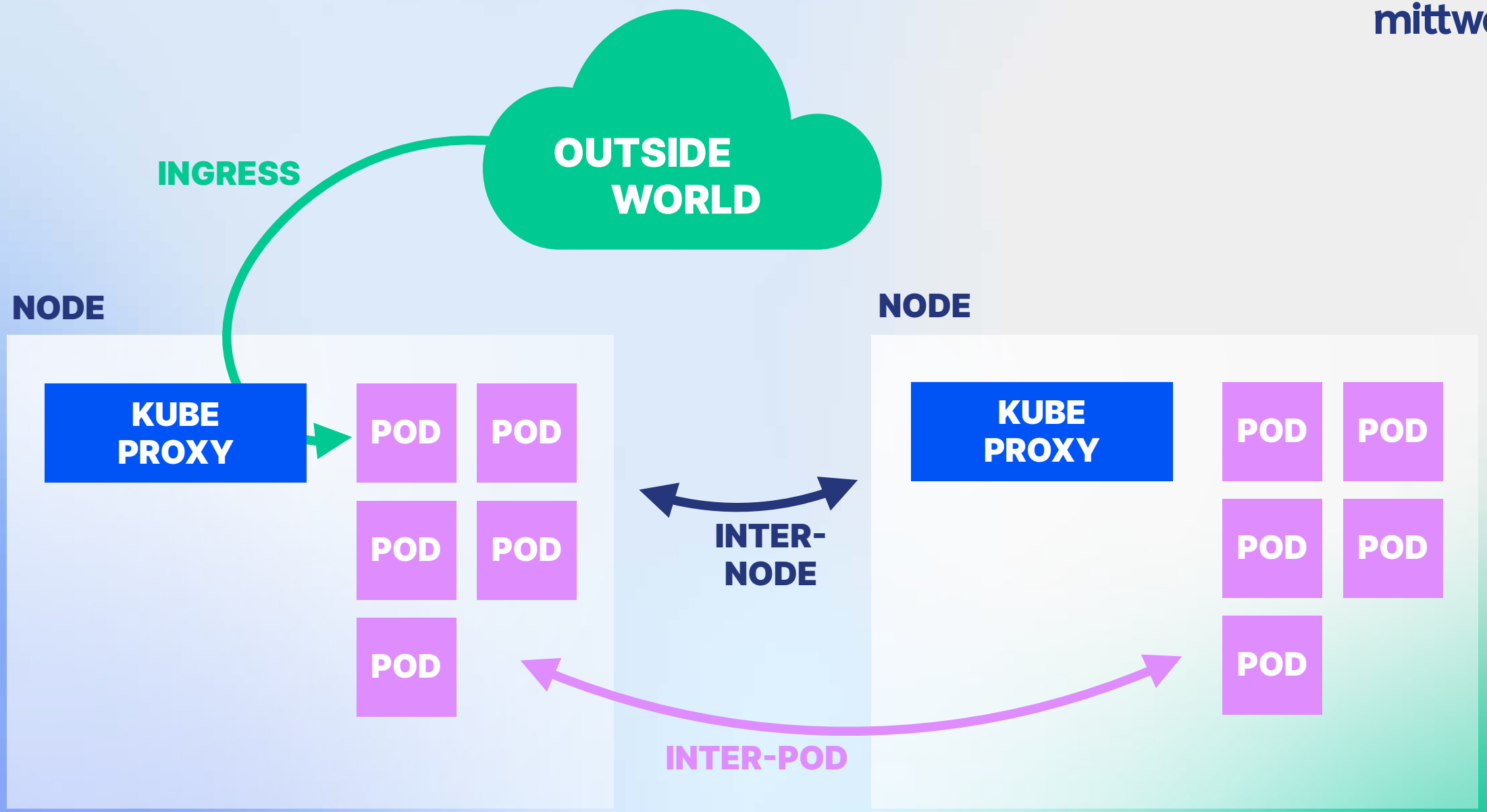
mittwald.

NETWORKING

mittwald.


Hi! My name is Jason
Ah! Nice to meet you, Jason
Jason
FRIENDS
FAMILY
Basic Kubernetes networking
So, Jarvis, wanna hear the difference between ClusterIP and NodePort?

mittwald.

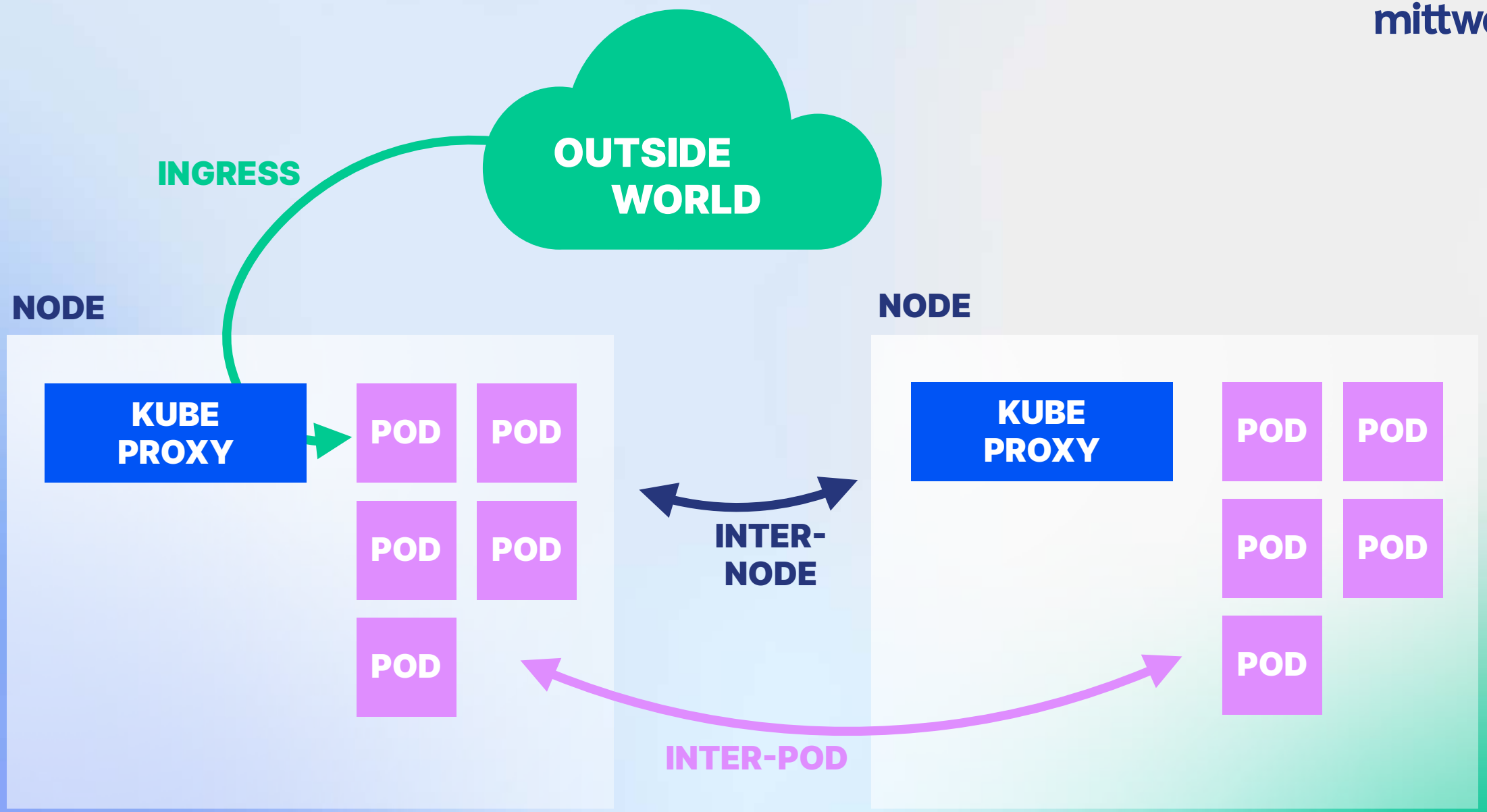

OUTSIDE WORLD
INGRESS
NODE
KUBE PROXY
POD
POD
POD
POD
POD
INTER-NODE
NODE
KUBE PROXY
POD
POD
POD
POD
POD
INTER-POD

MANUAL ROUTING

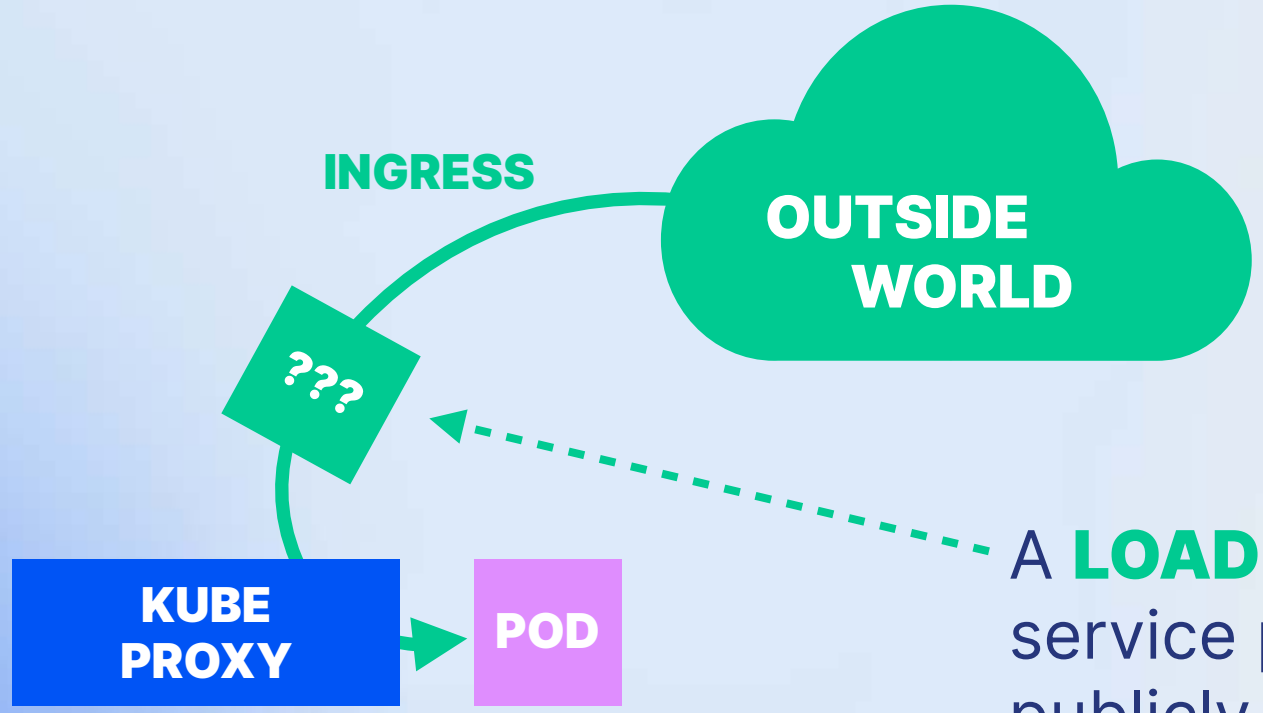
WEAVE CALICO AWS VPC

CILIUM FLANNEL AZURE CNI

mittwald.


OUTSIDE WORLD
INGRESS
NODE
KUBE PROXY
POD
POD
POD
POD
POD
INTER-NODE
NODE
KUBE PROXY
POD
POD
POD
POD
POD
INTER-POD

mittwald.



A **LOAD BALANCER** service provides a publicly routed IP address to a **SERVICE** object

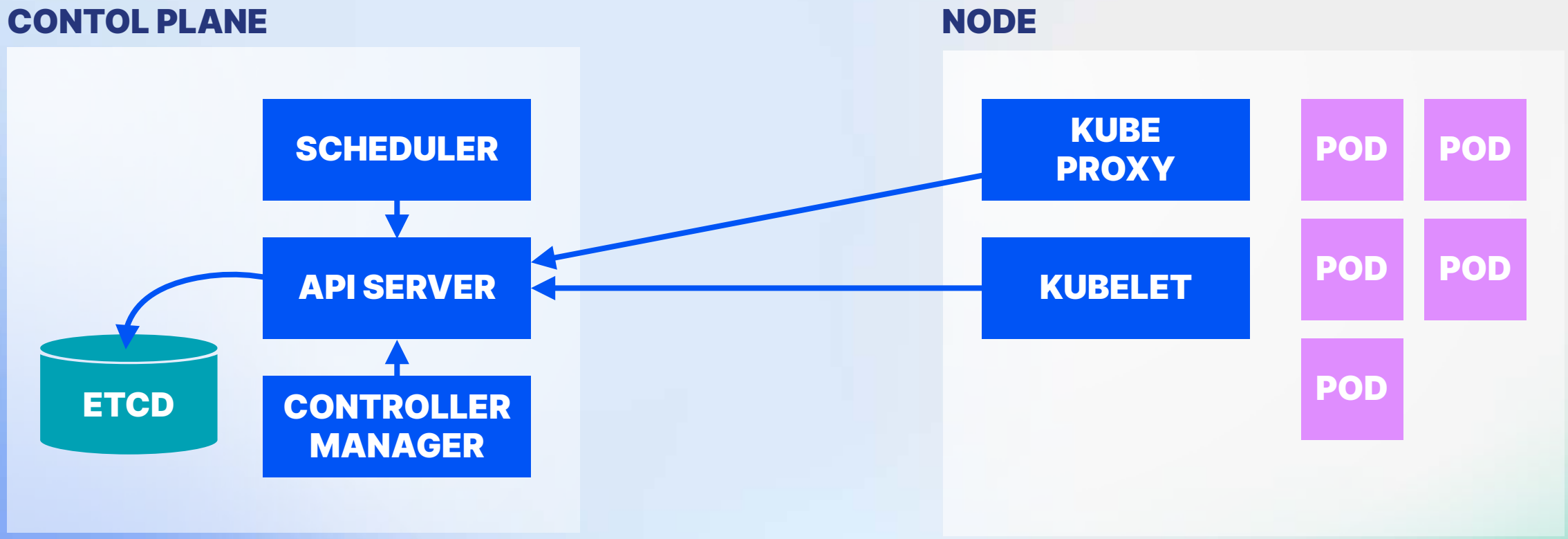
mittwald.

```yaml
apiVersion: v1
kind: Service
metadata:
  name: example-service
spec:
  type: LoadBalancer
  selector:
    app: example
  ports:
    - port: 80
      targetPort: 8080
```

What **ACTUALLY HAPPENS** when you create a LoadBalancer service?

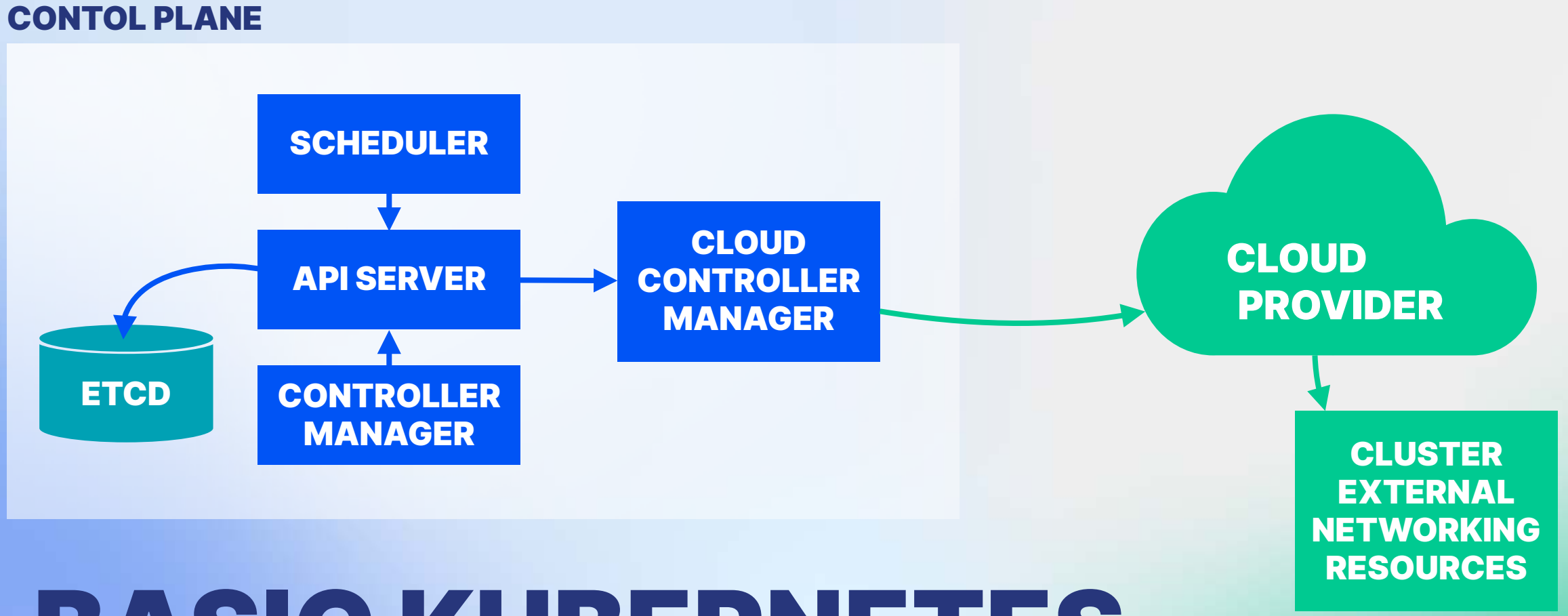

mittwald.
ALWAYS HAS BEEN
WAIT, IT'S ALL JUST IPTABLES?

mittwald.

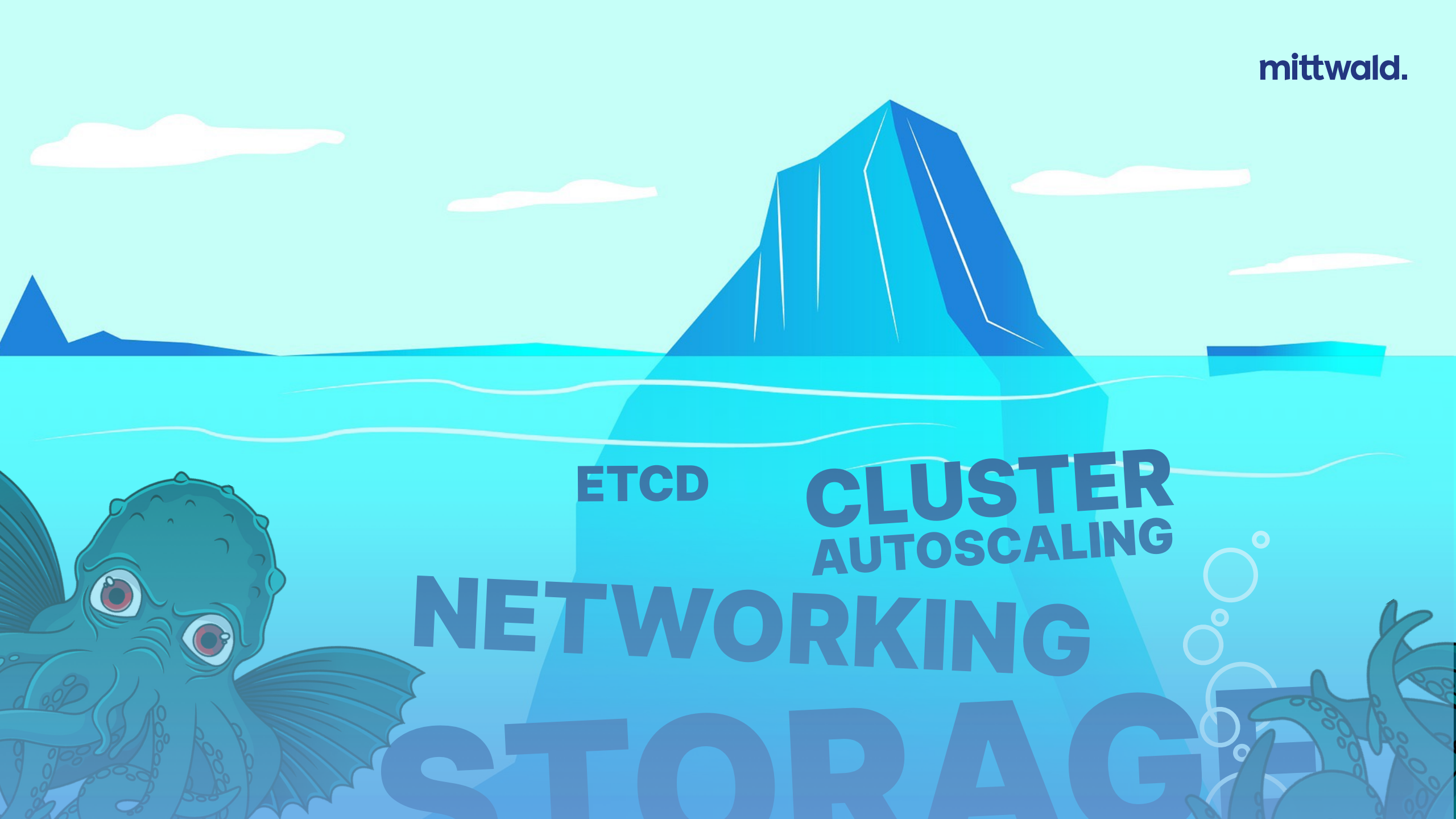


BASIC KUBERNETES ARCHITECTURE

mittwald.



BASIC KUBERNETES ARCHITECTURE

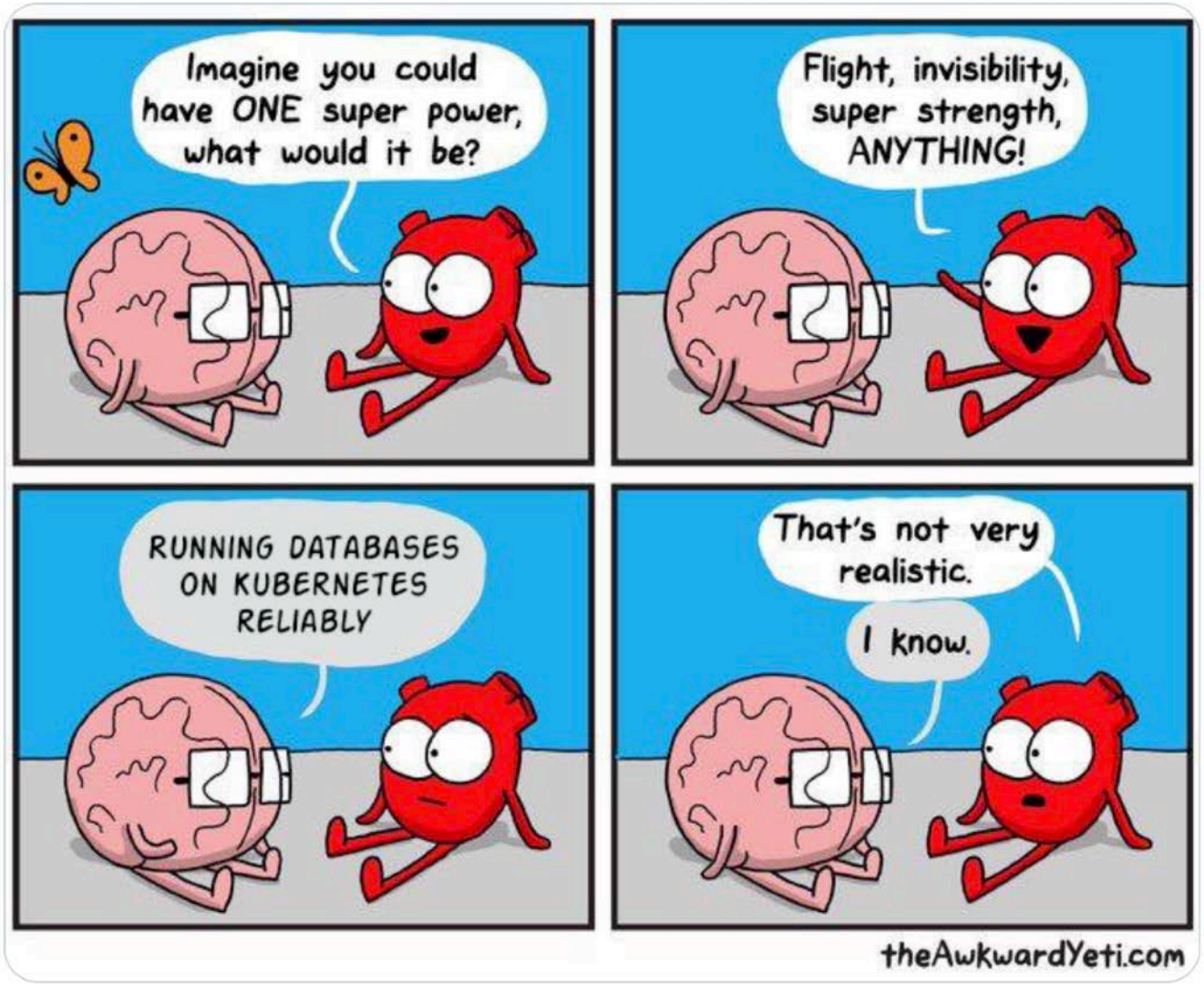

mittwald.
ETCD
CLUSTER
AUTOSCALING
NETWORKING
STORAGE

mit

STORAGE

Daniele Polencic — @danielepolencic@hachyderm.io
@danielepolencic

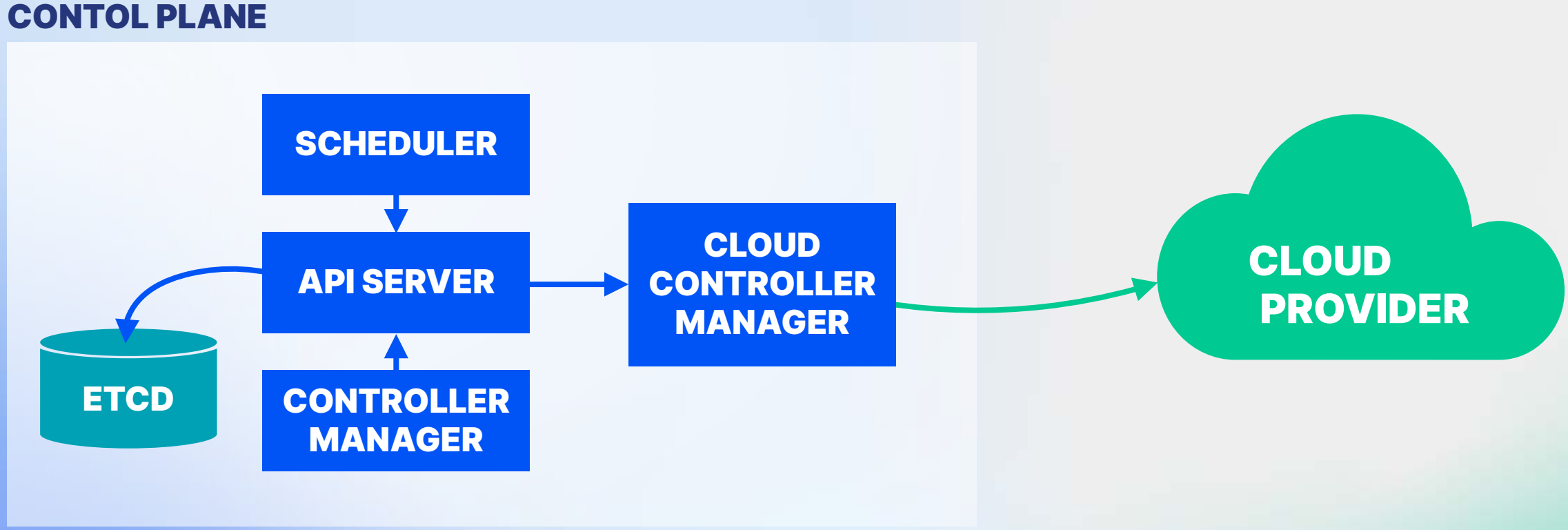
Databases on Kubernetes



5:25 nachm. · 31. Aug. 2019

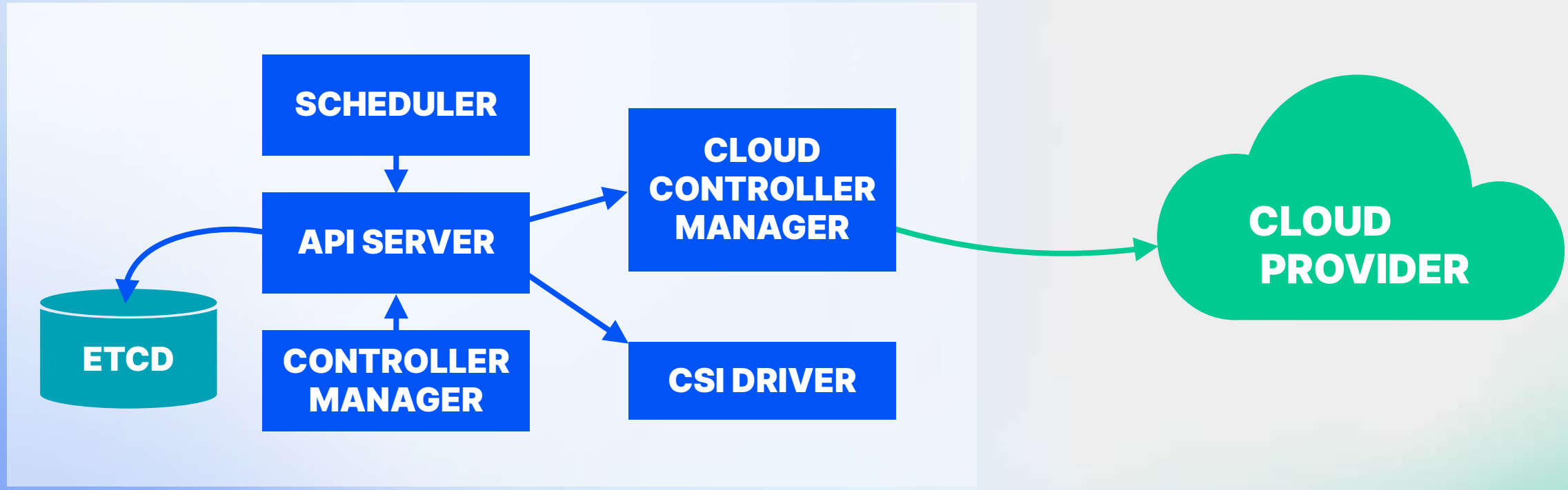
mittwald.

mittwald.



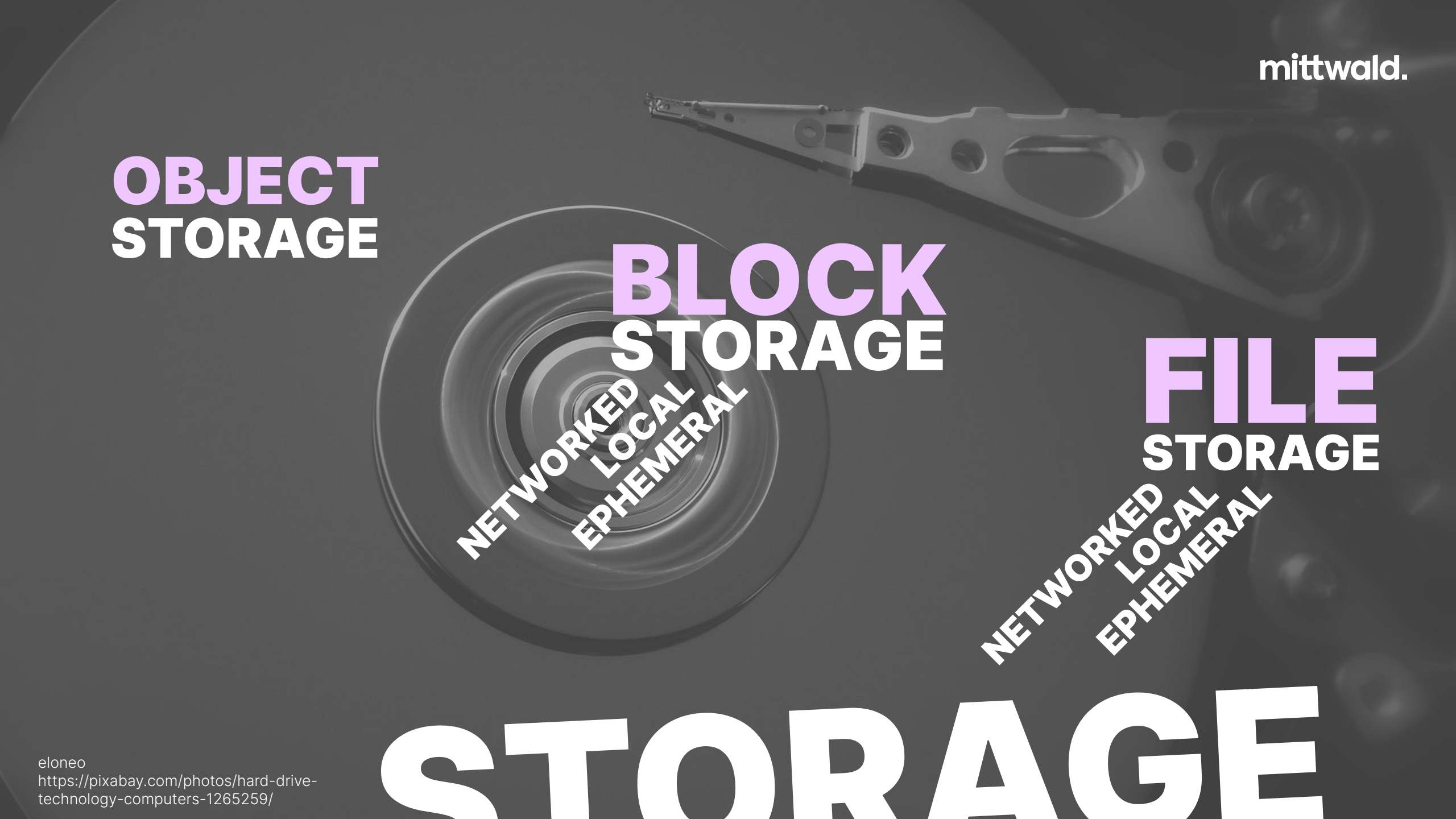
BASIC KUBERNETES ARCHITECTURE

mittwald.



BASIC KUBERNETES ARCHITECTURE

mittwald.



OBJECT STORAGE

BLOCK STORAGE

NETWORKED
LOCAL
EPHEMERAL

FILE STORAGE

NETWORKED
LOCAL
EPHEMERAL

STORAGE

eloneo
https://pixabay.com/photos/hard-drive-technology-computers-1265259/

Drivers

The following are a set of CSI driver which can be used with Kubernetes:

NOTE: If you would like your driver to be added to this table, please open a pull request in this repo updating this file. Other Features is allowed to be filled in Raw Block, Snapshot, Expansion, Cloning and Topology. If driver did not implement any Other Features, please leave it blank.

DISCLAIMER: Information in this table has not been validated by Kubernetes SIG-Storage. Users who want to use these CSI drivers need to contact driver maintainers for driver capabilities.

Production Drivers

Name	CSI Driver Name	Compatible with CSI Version(s)	Description	Persistence (Beyond Pod Lifetime)	Supported Access Modes	Dynar Provisio
Alicloud Disk	`diskplugin.csi.alibabacloud.com`	v1.0	A Container Storage Interface (CSI) Driver for Alicloud Disk	Persistent	Read/Write Single Pod	Yes
Alicloud NAS	`nasplugin.csi.alibabacloud.com`	v1.0	A Container Storage Interface (CSI) Driver for Alicloud Network Attached Storage (NAS)	Persistent	Read/Write Multiple Pods	No

mittwald.

OBJECT
STORAGE

BLOCK
STORAGE

FILE
STORAGE

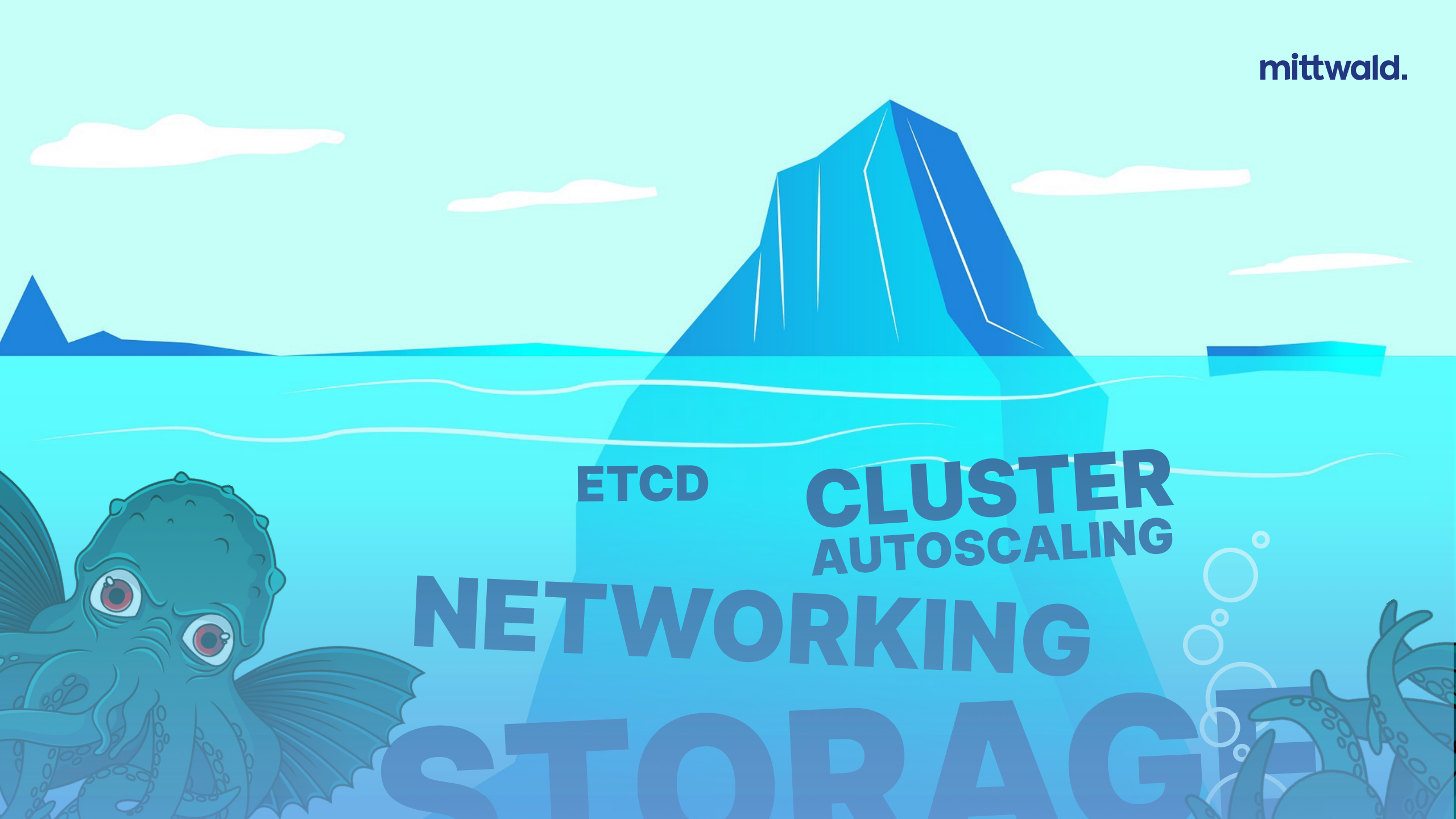
MIND YOUR WORKLOAD

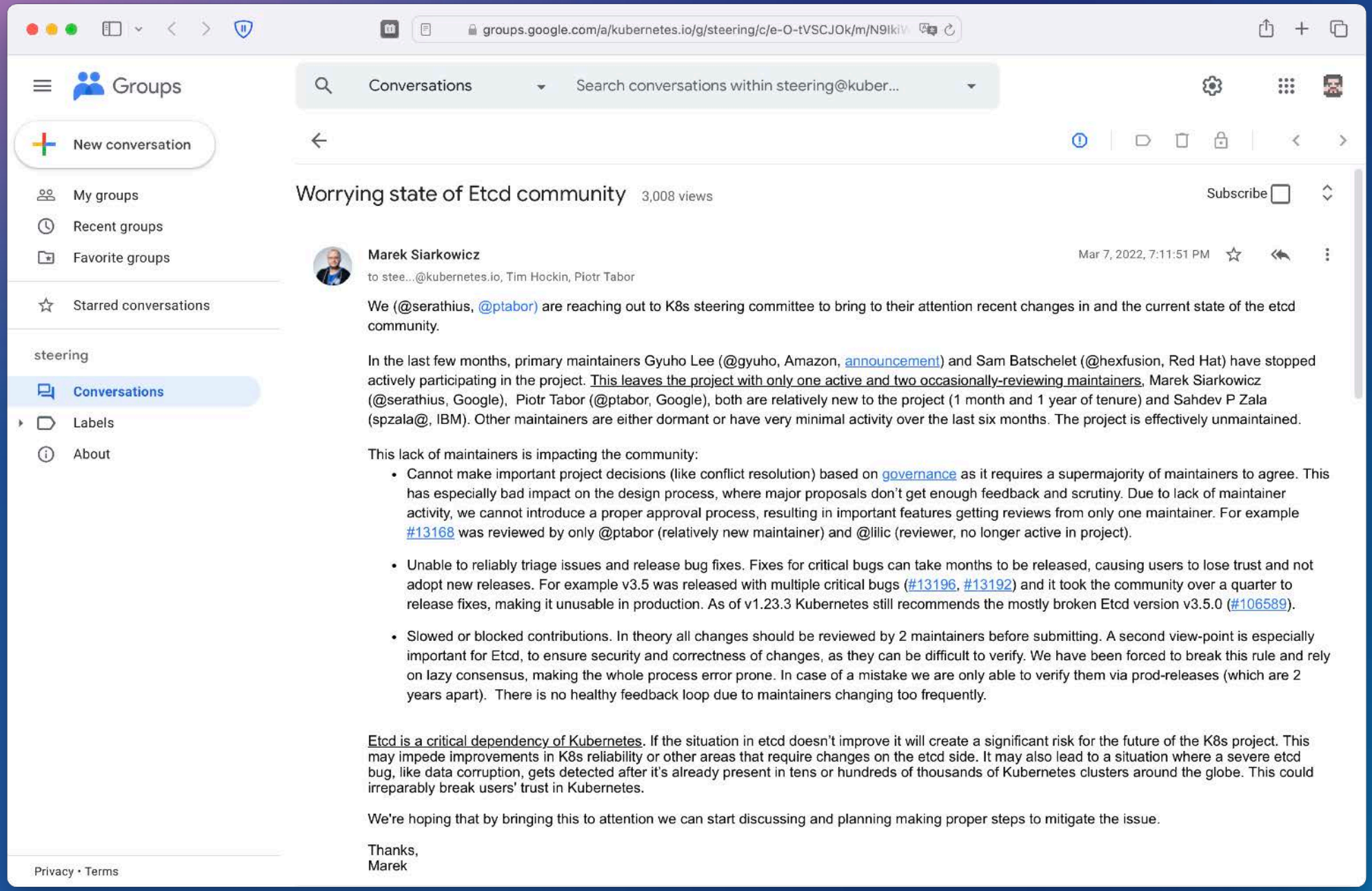
NETWORKED
LOCAL
EPHEMERAL

NETWORKED
LOCAL
EPHEMERAL

STORAGE

eloneo
https://pixabay.com/photos/hard-drive-technology-computers-1265259/


mittwald.
ETCD
CLUSTER
AUTOSCALING
NETWORKING
STORAGE

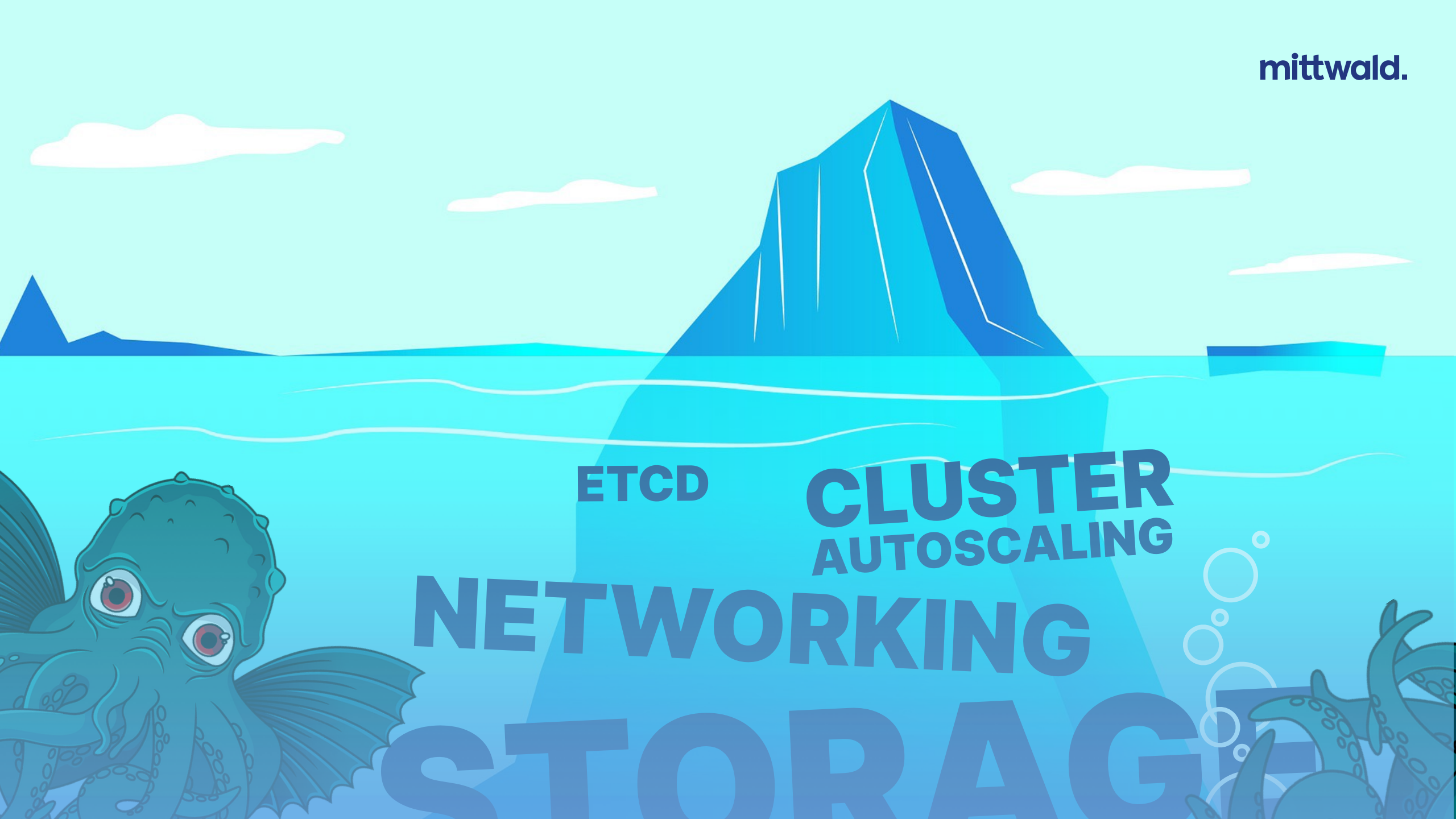

groups.google.com/a/kubernetes.io/g/steering/c/e-O-tVSCJOk/m/N9lkiV
Groups
Conversations
Search conversations within steering@kuber...
New conversation
My groups
Recent groups
Favorite groups
Starred conversations
steering
Conversations
Labels
About
Privacy • Terms
Worrying state of Etcd community
3,008 views
Subscribe
Marek Siarkowicz
to stee...@kubernetes.io, Tim Hockin, Piotr Tabor
Mar 7, 2022, 7:11:51 PM
We (@serathius, @ptabor) are reaching out to K8s steering committee to bring to their attention recent changes in and the current state of the etcd community.
In the last few months, primary maintainers Gyuho Lee (@gyuho, Amazon, announcement) and Sam Batschelet (@hexfusion, Red Hat) have stopped actively participating in the project. This leaves the project with only one active and two occasionally-reviewing maintainers, Marek Siarkowicz (@serathius, Google), Piotr Tabor (@ptabor, Google), both are relatively new to the project (1 month and 1 year of tenure) and Sahdev P Zala (spzala@, IBM). Other maintainers are either dormant or have very minimal activity over the last six months. The project is effectively unmaintained.
This lack of maintainers is impacting the community:
Cannot make important project decisions (like conflict resolution) based on governance as it requires a supermajority of maintainers to agree. This has especially bad impact on the design process, where major proposals don't get enough feedback and scrutiny. Due to lack of maintainer activity, we cannot introduce a proper approval process, resulting in important features getting reviews from only one maintainer. For example #13168 was reviewed by only @ptabor (relatively new maintainer) and @lilic (reviewer, no longer active in project).
Unable to reliably triage issues and release bug fixes. Fixes for critical bugs can take months to be released, causing users to lose trust and not adopt new releases. For example v3.5 was released with multiple critical bugs (#13196, #13192) and it took the community over a quarter to release fixes, making it unusable in production. As of v1.23.3 Kubernetes still recommends the mostly broken Etcd version v3.5.0 (#106589).
Slowed or blocked contributions. In theory all changes should be reviewed by 2 maintainers before submitting. A second view-point is especially important for Etcd, to ensure security and correctness of changes, as they can be difficult to verify. We have been forced to break this rule and rely on lazy consensus, making the whole process error prone. In case of a mistake we are only able to verify them via prod-releases (which are 2 years apart). There is no healthy feedback loop due to maintainers changing too frequently.
Etcd is a critical dependency of Kubernetes. If the situation in etcd doesn't improve it will create a significant risk for the future of the K8s project. This may impede improvements in K8s reliability or other areas that require changes on the etcd side. It may also lead to a situation where a severe etcd bug, like data corruption, gets detected after it's already present in tens or hundreds of thousands of Kubernetes clusters around the globe. This could irreparably break users' trust in Kubernetes.
We're hoping that by bringing this to attention we can start discussing and planning making proper steps to mitigate the issue.
Thanks,
Marek

mittwald.

mittwald.

← etcd on networked block storage →

← etcd on local SSD →

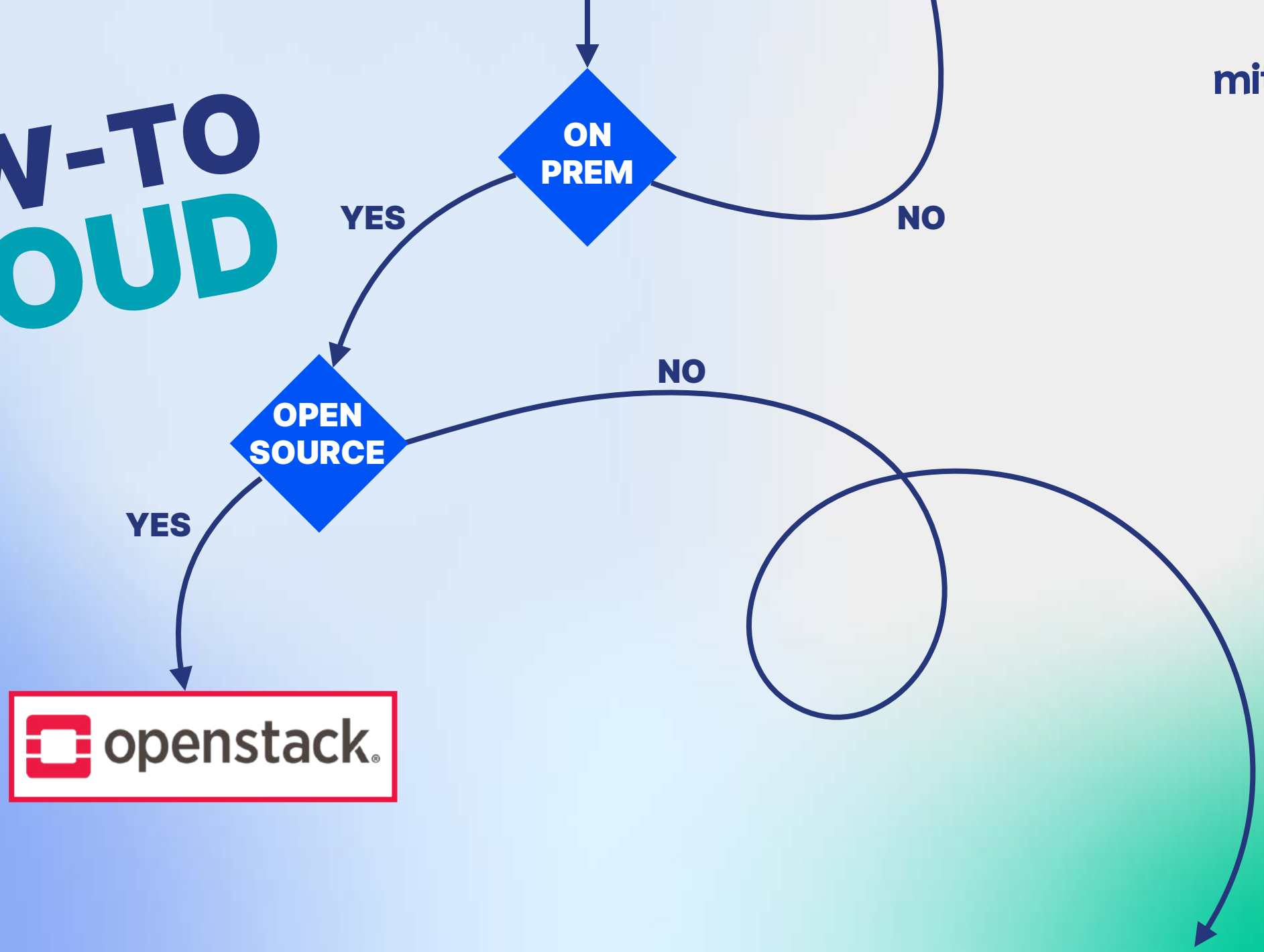

mittwald.
ETCD
CLUSTER
AUTOSCALING
NETWORKING
STORAGE

mittwald.

PROVISIONING CLUSTERS

HOW-TO CLOUD

mittwald.



mittwald.

KUBERMATIC

OPENSTACK MAGNUM

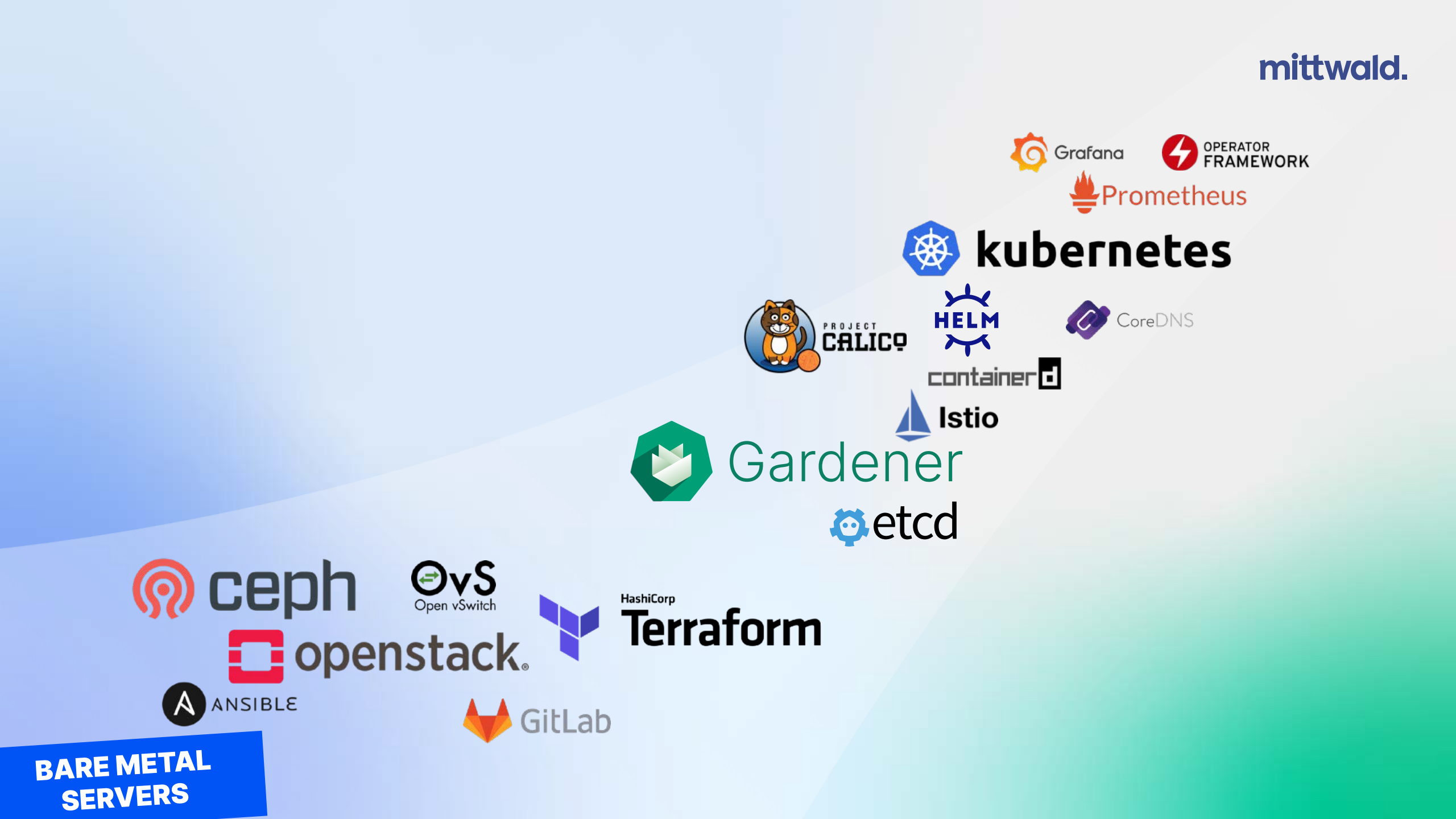
CLUSTER API

GARDENER

RANCHER®
BY SUSE

OPEN-SOURCE CLUSTER PROVISIONING

mittwald.

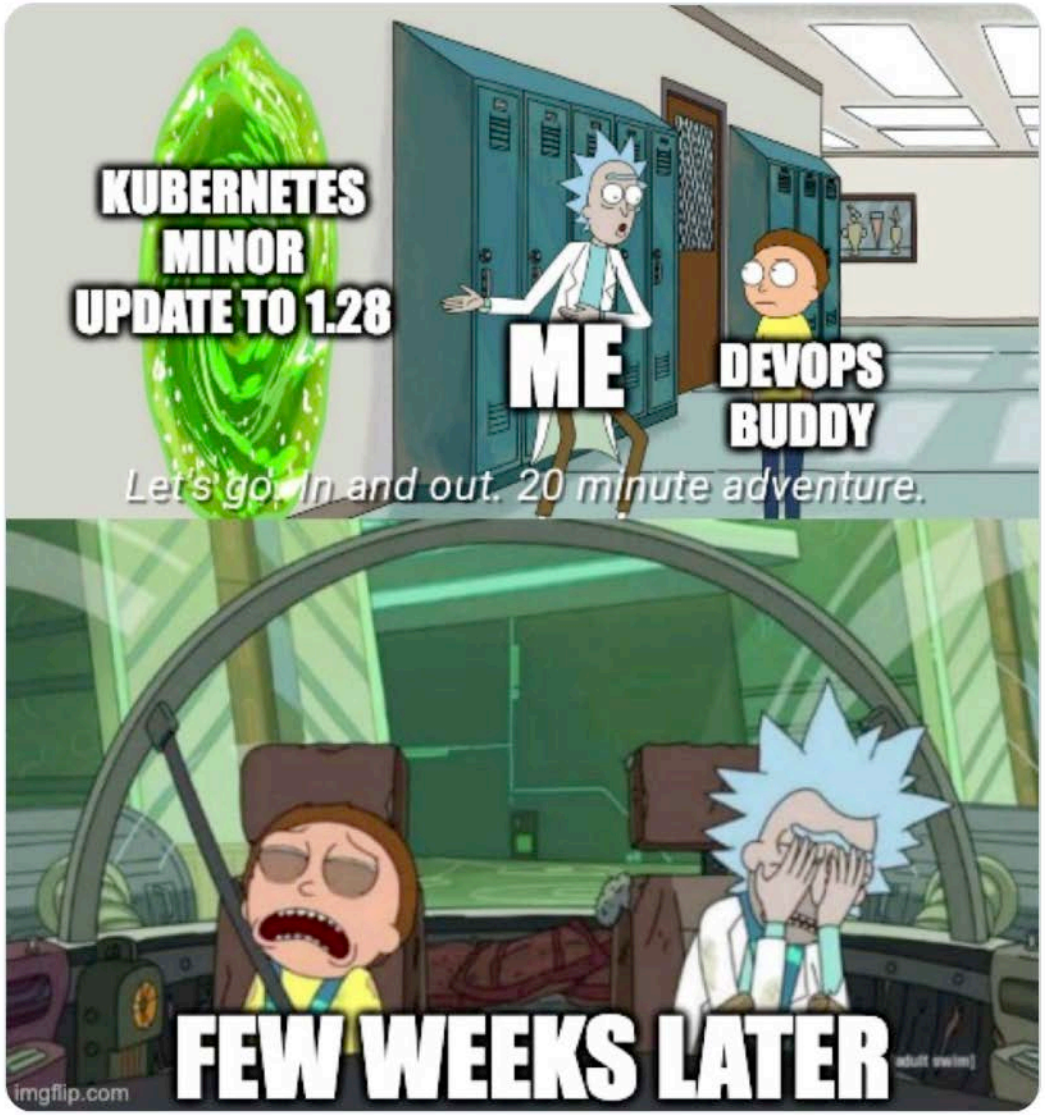

Grafana
OPERATOR FRAMEWORK
Prometheus
kubernetes
PROJECT CALICO
HELM
CoreDNS
containerd
Istio
Gardener
etcd
ceph
OvS
Open vSwitch
openstack
HashiCorp
Terraform
ANSIBLE
GitLab
BARE METAL SERVERS

mittwald.

memenetes
@memenetes

Are you ready for a quick adventure?

Post übersetzen



6:02 nachm. · 24. Aug. 2023 · **43.110** Mal angezeigt

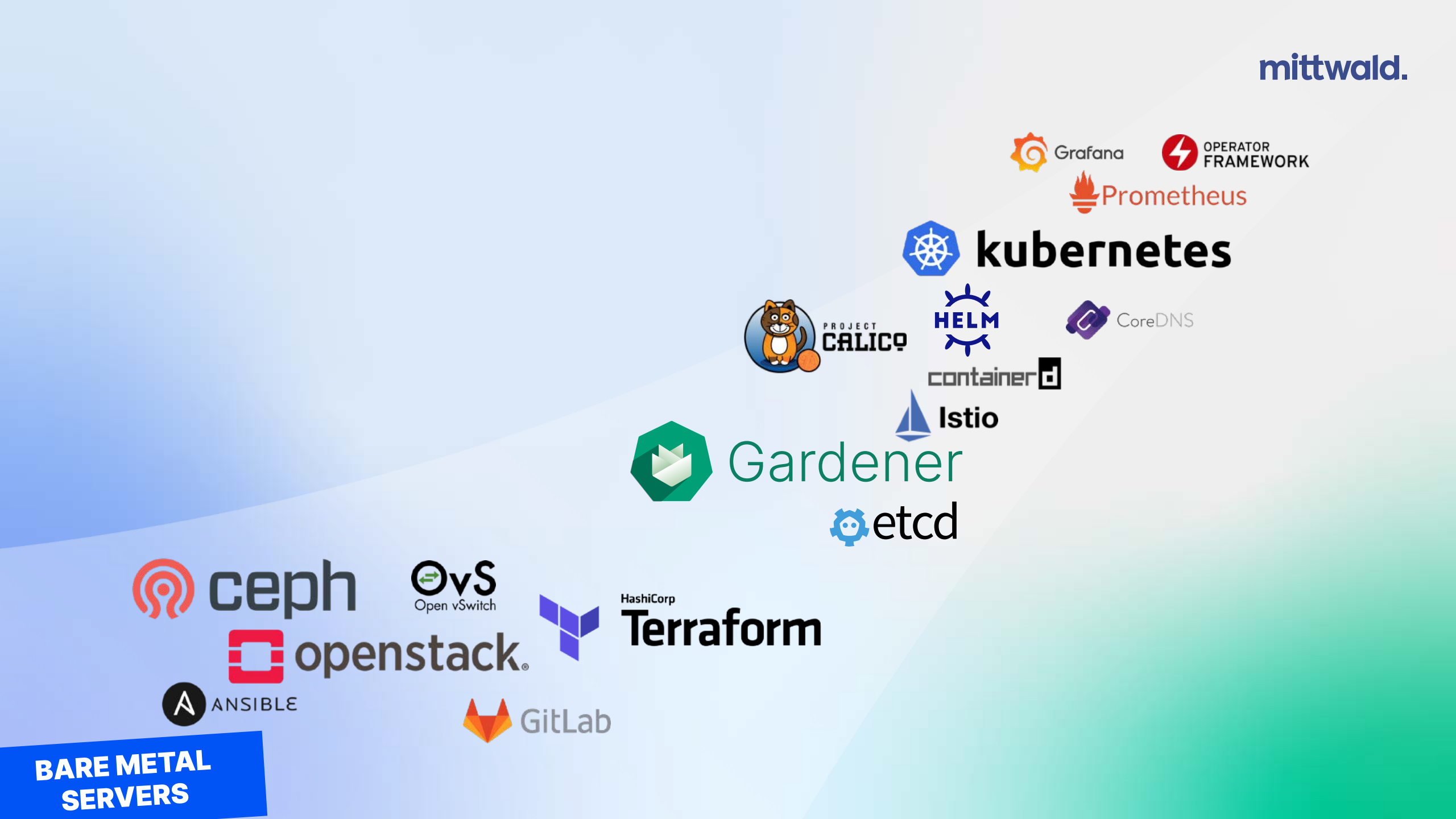
Lukas Tennie
https://unsplash.com/photos/a-close-up-of-a-watch-face-with-the-gears-missing-3dyDozzCORw

mittwald.


ORGANIZATIONAL BOUNDARY
Grafana
OPERATOR FRAMEWORK
Prometheus
kubernetes
PROJECT CALICO
HELM
CoreDNS
containerd
Istio
Gardener
etcd
ceph
OvS
Open vSwitch
HashiCorp
Terraform
openstack
ANSIBLE
GitLab
BARE METAL SERVERS

Erika Wittlieb
https://pixabay.com/photos/lemonade-stand-lemonade-summer-2483297/

mittwald.


Grafana
OPERATOR FRAMEWORK
Prometheus
kubernetes
PROJECT CALICO
HELM
CoreDNS
containerd
Istio
Gardener
etcd
ceph
OvS
Open vSwitch
openstack
HashiCorp
Terraform
ANSIBLE
GitLab
BARE METAL SERVERS

mittwald.

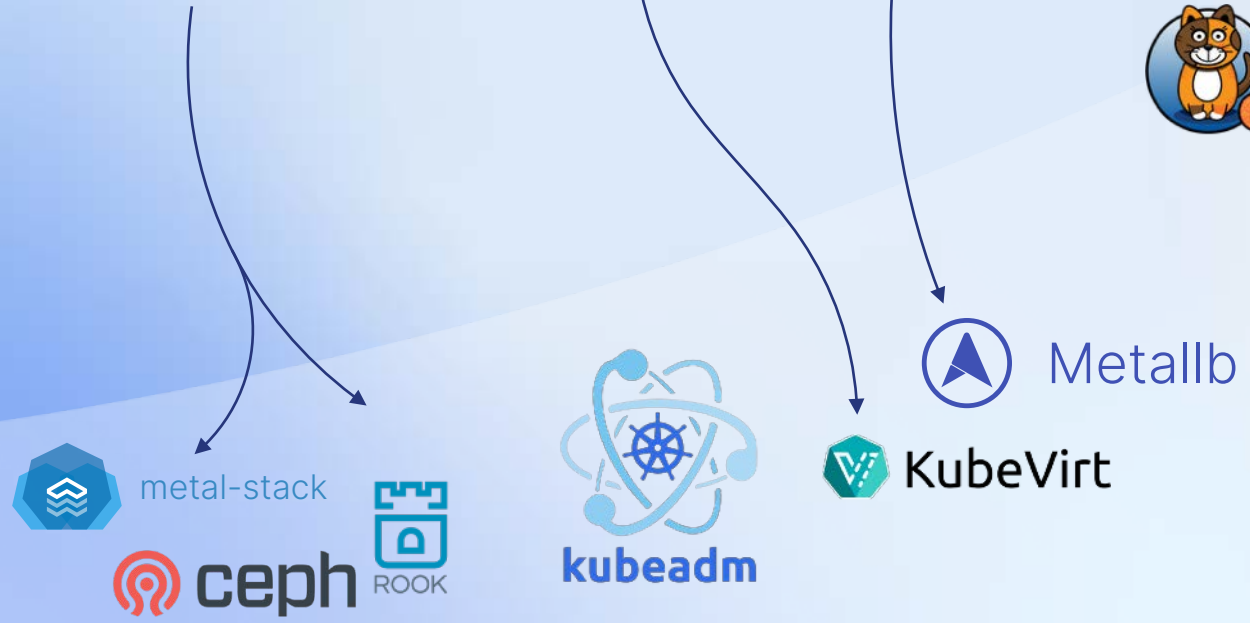
OUR APPROACH:

KUBERNETES-NATIVE EVERYTHING

STORAGE

VIRTUALIZATION

NETWORKING



Grafana

OPERATOR FRAMEWORK

Prometheus

kubernetes

PROJECT CALICO

HELM

CoreDNS

containerd

BARE METAL SERVERS

REMINDER:
YMMV

mittwald.

https://metal-stack.io/

https://github.com/onmetal

OUTLOOK

mittwald.

ACCEPT (AND MANAGE ACCORDINGLY) COMPLEXITY WHERE YOU NEED IT

KNOW YOUR PRODUCT

KEEP IT SIMPLE WHERE YOU DON'T

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

https://github.com/martin-helmich

https://www.mittwald.de
https://www.martin-helmich.de