

Infrastructure for the Future: Public Cloud or Kubernetes?

INNOQ

EBERHARD WOLFF

Fellow at INNOQ Deutschland GmbH

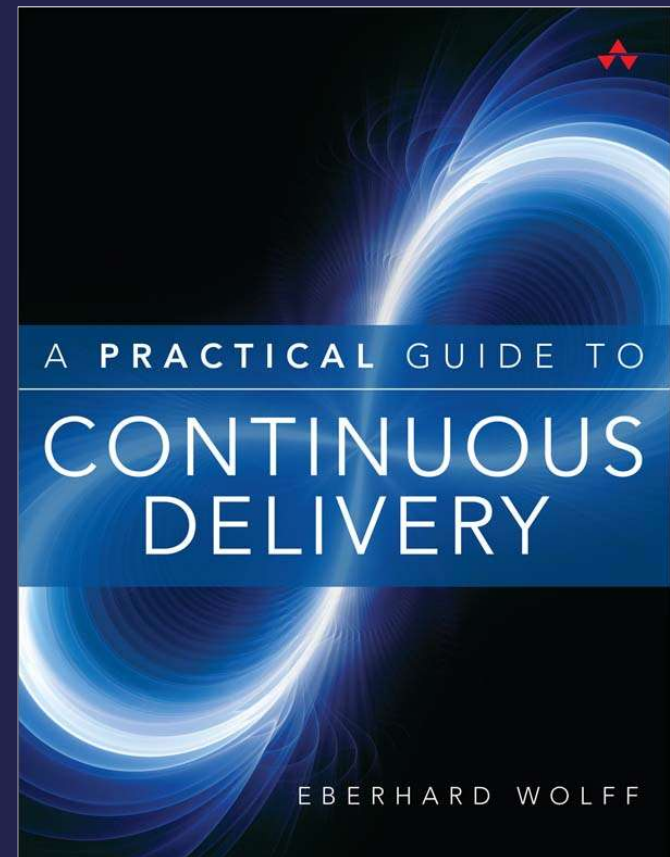
@ewolff

www.ewolff.com



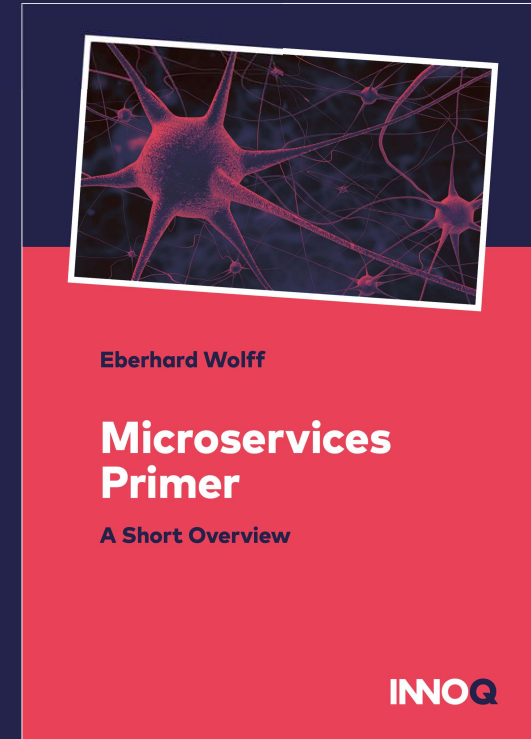
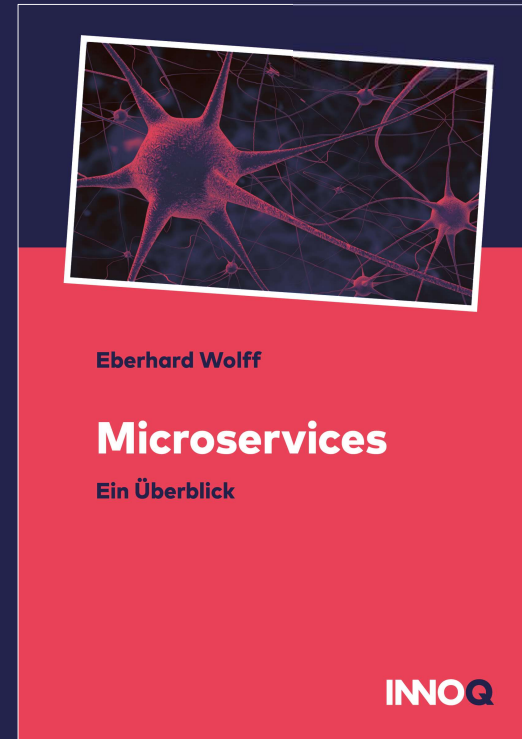
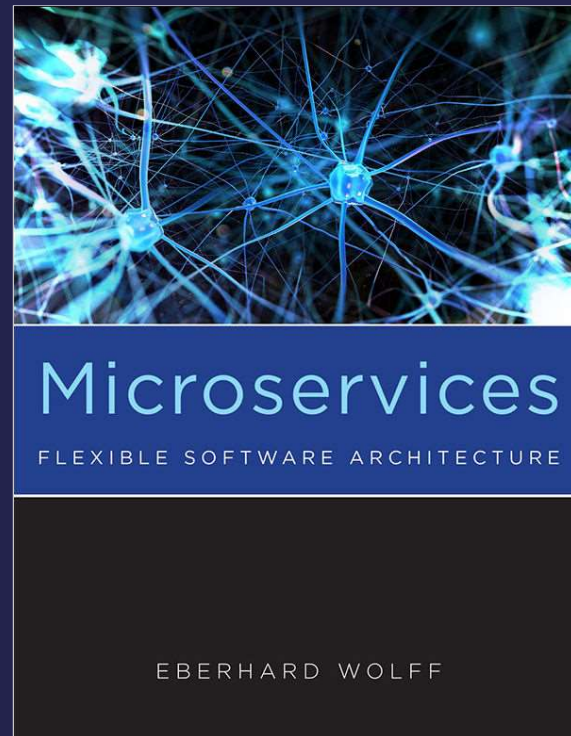
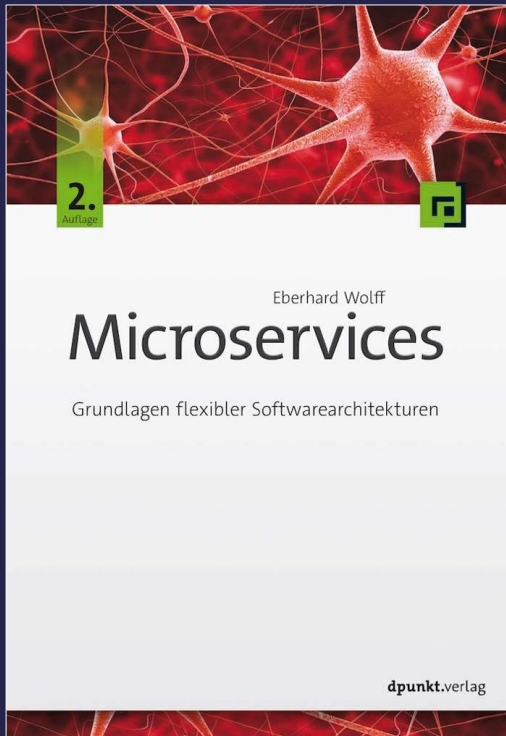


www.continuous-delivery-buch.de



www.continuous-delivery-buch.de

FREE



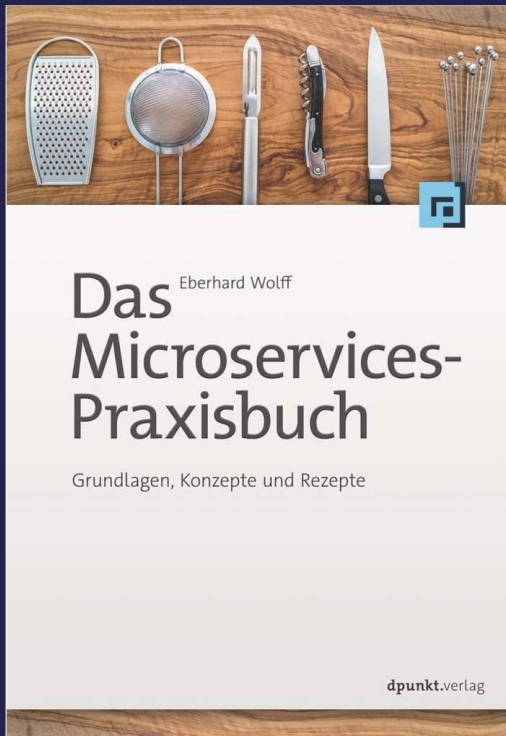
microservices-buch.de

microservices-book.com

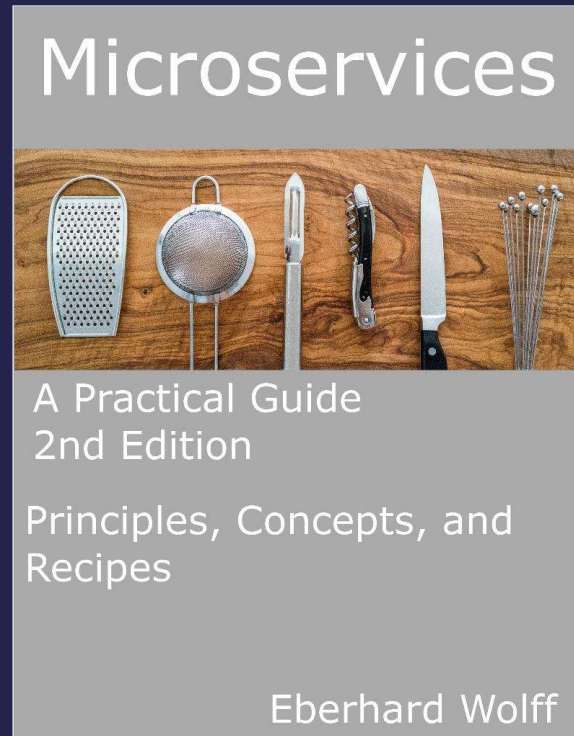
[microservices-buch.de/
ueberblick.html](http://microservices-buch.de/ueberblick.html)

[microservices-book.com/
primer.html](http://microservices-book.com/primer.html)

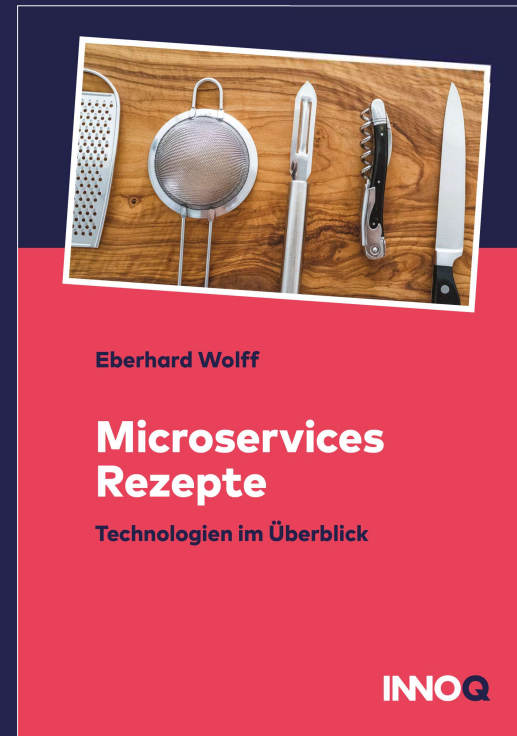
FREE



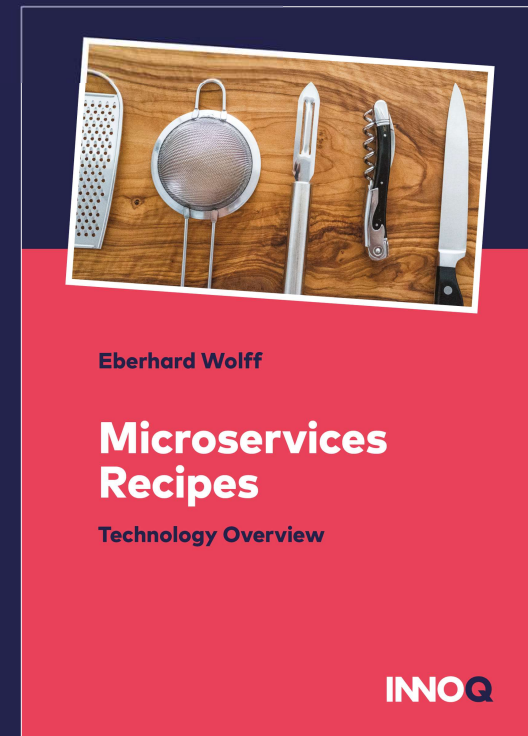
microservices-praxisbuch.de



practical-microservices.com



[microservices-praxisbuch.de/
rezepte.html](https://microservices-praxisbuch.de/rezepte.html)



[practical-microservices.com/
recipes.html](https://practical-microservices.com/recipes.html)

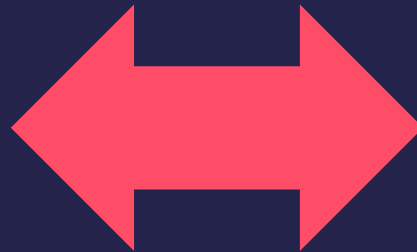
FREE



leanpub.com/service-mesh-primer/

Public Cloud

**Cost
Efficient**



Lock In

Local Kubernetes Cluster



wolff@BLACK-HARDWARE: ~/win/microservice-istio/microservice-istio-demo



```
wolff@BLACK-HARDWARE:~/win/microservice-istio/microservice-istio-demo$ kubectl apply -f infrastructure.yaml
```


Local Kubernetes Cluster



wolff@BLACK-HARDWARE: ~/win/microservice-istio/microservice-istio-demo



```
wolff@BLACK-HARDWARE:~/win/microservice-istio/microservice-istio-demo$ kubectl apply -f infrastructure.yaml
deployment.apps/apache created
deployment.apps/postgres created
service/apache created
service/postgres created
gateway.networking.istio.io/microservice-gateway created
virtualservice.networking.istio.io/apache created
wolff@BLACK-HARDWARE:~/win/microservice-istio/microservice-istio-demo$
```

Kubernetes Cluster Google Cloud

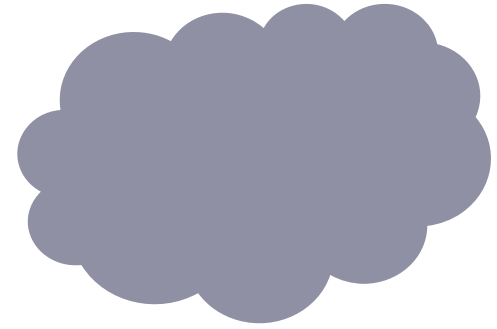


wolff@BLACK-HARDWARE: ~/win/microservice-istio/microservice-istio-demo



```
wolff@BLACK-HARDWARE:~/win/microservice-istio/microservice-istio-demo$ kubectl apply -f infrastructure.yaml
deployment.apps/apache created
deployment.apps/postgres created
service/apache created
service/postgres created
gateway.networking.istio.io/microservice-gateway created
virtualservice.networking.istio.io/apache created
wolff@BLACK-HARDWARE:~/win/microservice-istio/microservice-istio-demo$
```

Public Cloud



- Kubernetes:
Finally the universal abstraction!
- Open market with interchangeable providers.
- More competition
- Cost saving

Kubernetes as Abstraction

Kubernetes: Public Cloud

- Kubernetes as a service
- Amazon Elastic Kubernetes Service (EKS)
- Google Kubernetes Engine (GKE)
- Azure Kubernetes Service (AKS)
- Provide more or less control
- Could roll your own
...much more effort



EKS



GKE



AKS

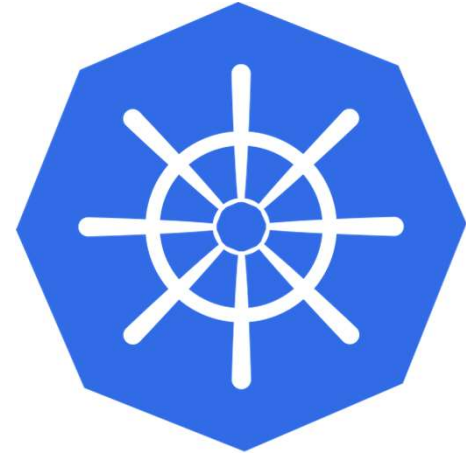
Kubernetes On-Premise

- kubectl 
- kubeadm 
- Gardener (Public IaaS, Open Stack) 
- Rancher (Public IaaS, Hosted, vSphere) 
- Unify public and private Clouds
- ...

Kubernetes On-Premise: Complexity?

Kubernetes

- Complex piece of software
 - Solution for clustering
 - Additional networking
 - ...
-
- Why bother with it?



Kubernetes vs. Virtualization

- Virtualization: ubiquitous
- Kubernetes solves many virtualization challenges
 - ...but for containers
- Storage, network are also issues for enterprise virtualization

Kubernetes vs. Virtualization

- VMware Project Pacific:
vSphere with Kubernetes support
- Can run Kubernetes on bare metal
i.e. no virtualization needed
- Complexity *and* power like virtualization
- Will Kubernetes replace virtualization?
- Is Kubernetes more complex than virtualization?

**Kubernetes is a Great
Abstraction – Let's Save
Some Money in the Cloud!**

How Much Cheaper Is Public Cloud?

Price Comparison

- Rough numbers
- Not meant to be a detailed discussion
- Tried to make it as fair as possible

Comparison: Cloud

- Amazon Web Services
- Azure (to have two clouds)

Comparison: On Premise

- Hetzner: A random (cheap) provider
- Private datacenter probably more expensive

	Amazon
RAM	32GB
Threads	8
HDD/ SDD	0
Per month (reserved / on demand)	281.09\$ 107.59\$

	Amazon	Azure
RAM	32GB	32GB
Threads	8	8
HDD/ SDD	0	64GB
Per month (reserved / on demand)	281.09\$ 107.59\$	243.09\$ 91.47\$

	Amazon	Azure	Hetzner
RAM	32GB	32GB	64GB
Threads	8	8	8
HDD/ SDD	0	64GB	2*4TB HDD
Per month (reserved / on demand)	281.09\$ 107.59\$	243.09\$ 91.47\$	51,23\$

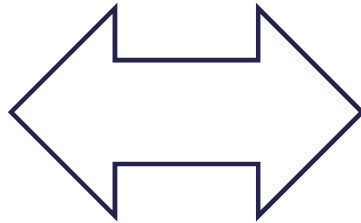
~ 1/2 Amazon
/ Microsoft

**Cloud seems quite
expensive? 🤔**

Elastic Scaling

- One Saturday night
 - Performance tests
 - On 40+ server
 - All with just one guy
-
- Impossible in your data center

Tests:
40+ Server
Cloud
on demand



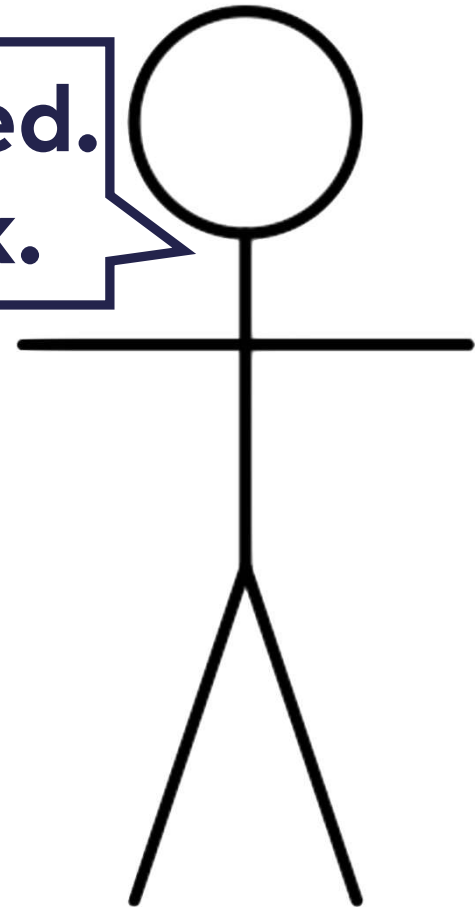
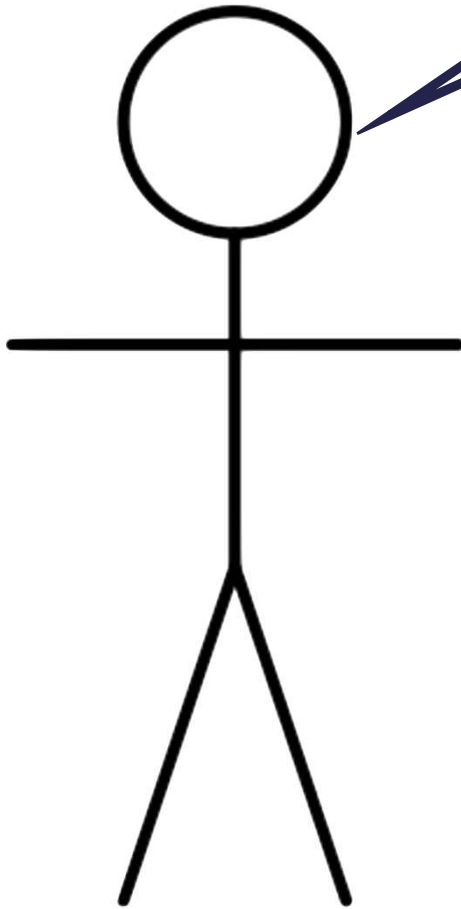
Production:
On premise
Order servers
before
project even
started

Testing
Cloud

Here are
my results

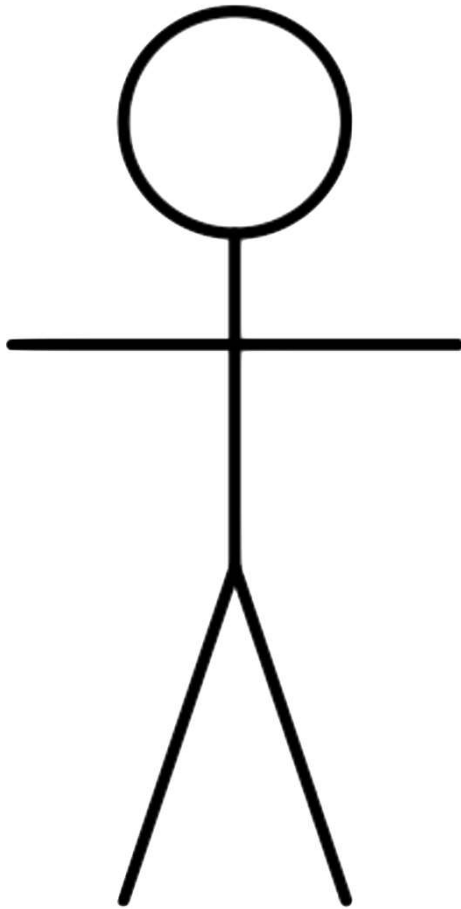
Testing on
premise

Haven't started.
Doesn't work.



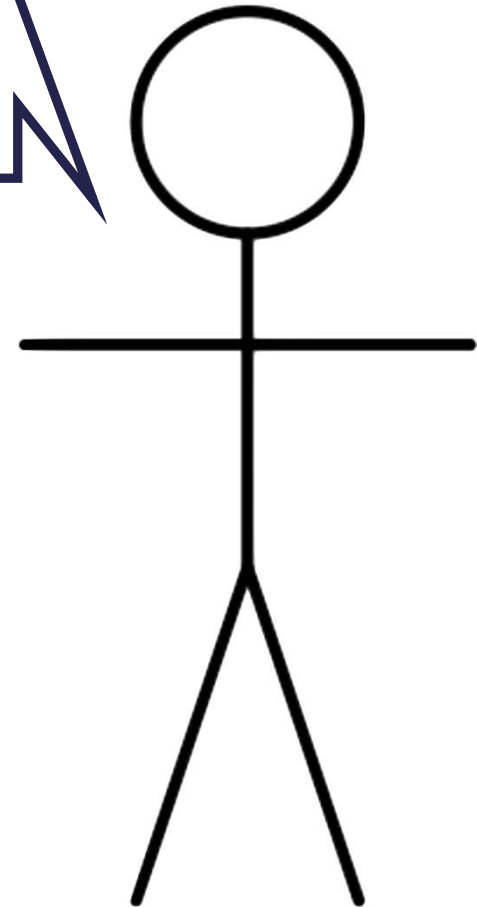
2 Days Later...

**Testing
Cloud**



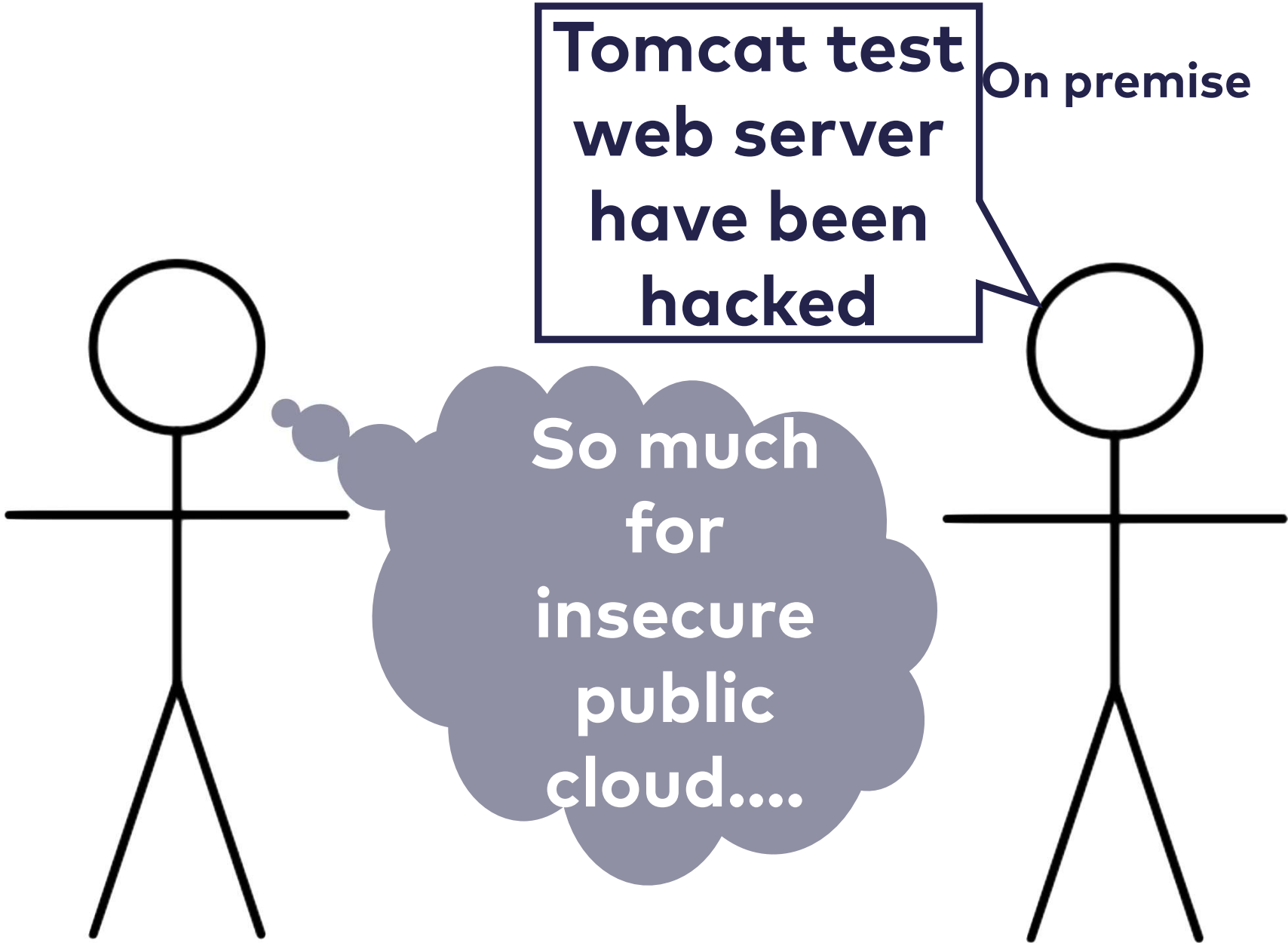
**Finally started.
Firewall was
configured
incorrectly.**

**Testing on
premise**



**One engineer
two days
cost?**

**Loosing 2 days when
working weekends 😞**



**Tomcat test
web server
have been
hacked**

On premise

**So much
for
insecure
public
cloud....**

Firewalls on Amazon

- Build-in
- On the network level
- Easy to configure
- Many more security features
- Physical access, PCI, BSI C5 compliance...

Instagram Pre-Acquisition



- Global photo sharing
- S3 storage
- Cloudfront Content Delivery Network
- Custom code
- 14 mio user
- Several 100 mio photos
- 3 engineers (Dev and Ops)

Cloud = Components

- Databases + backup, disaster recovery...
- Content delivery network
- Object storage
- ...

**Public Cloud is not about
costs
...but components**

Amazon's Service Catalog



Analytics



Application Integration



AR & VR



AWS Cost Management



Blockchain



Business Applications



Compute



Customer Engagement



Database



Developer Tools



End User Computing



Game Tech



Internet of Things



Machine Learning



Management & Governance



Media Services



Migration & Transfer



Mobile



Networking & Content
Delivery



Robotics



Satellite



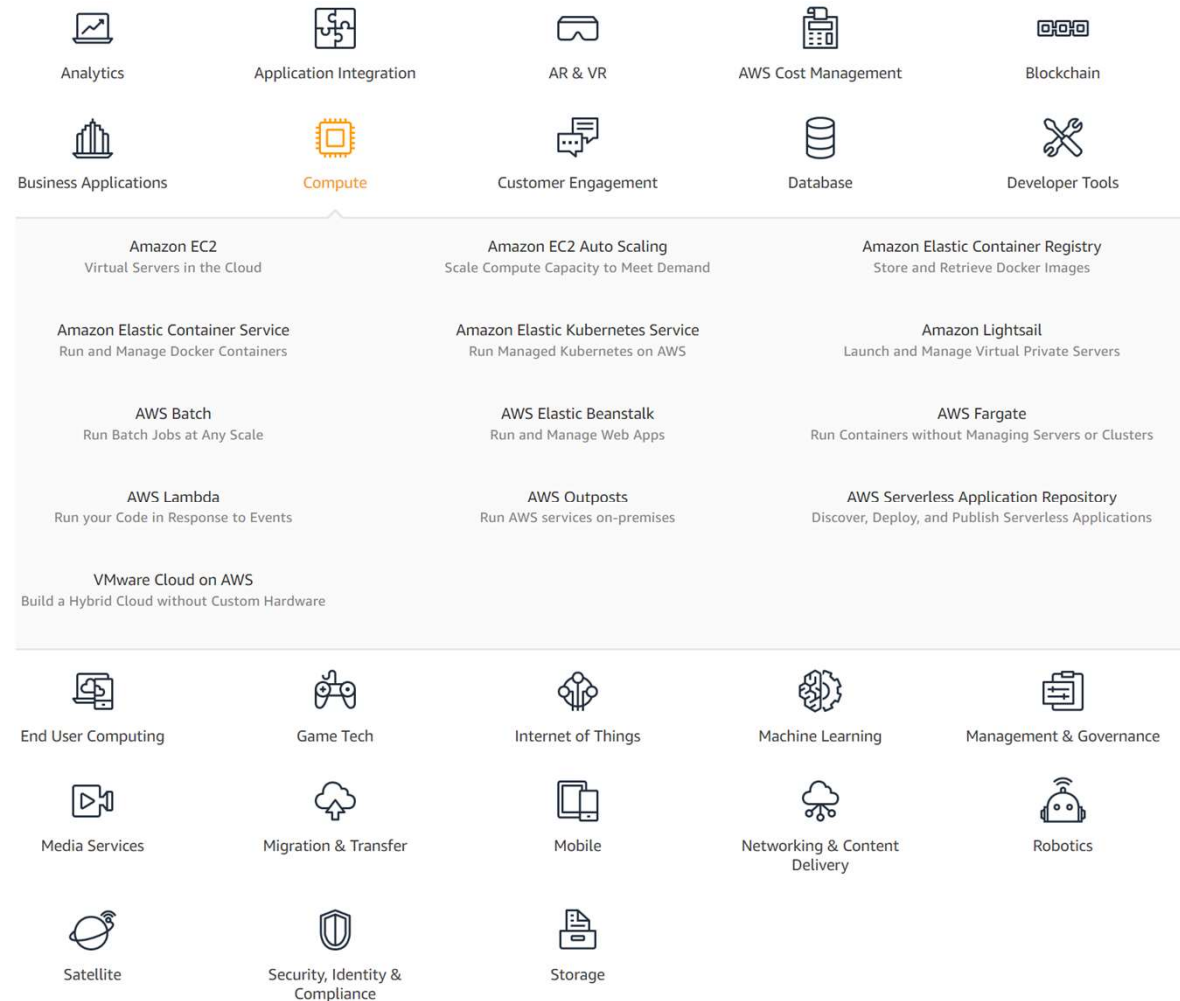
Security, Identity &
Compliance



Storage

Amazon's Service Catalog

- Kubernetes is an abstraction just for Compute



**Can We Have All the
Components on Kubernetes?**

PaaS / Serverless

Kubernetes

- Good foundation
- Provides:
 - Deployment,
 - service discovery,
 - load balancing
- Service meshes solve additional challenges (security, observability, canary deployment ...)

Alternative: PaaS / Serverless

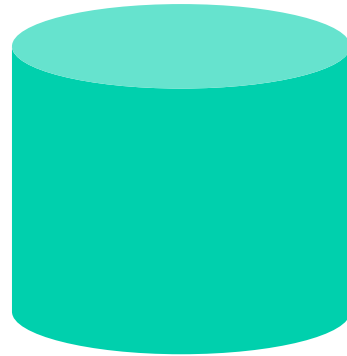
- Higher level of abstraction
- `cf push` instead of `Dockerfile` + `YAML`
- On premise: PaaS hard to operate
- Public cloud: operations done for you
- Devs interested in Kubernetes is a success of DevOps

Alternative: PaaS

- PaaS will use Kubernetes as foundation
- Serverless uses Kubernetes as foundation
- Can set up these services on Kubernetes
...but overhead for operations

**Operations overhead for
other components?**

State in the Cloud



Amazon SLA

- Single EC2 instance: 90% per hour
otherwise: no charge
- 99,99% availability
- Failure:
all running instances not accessible
in >1 data center (availability zone)
- i.e. it is fine if a data center fails

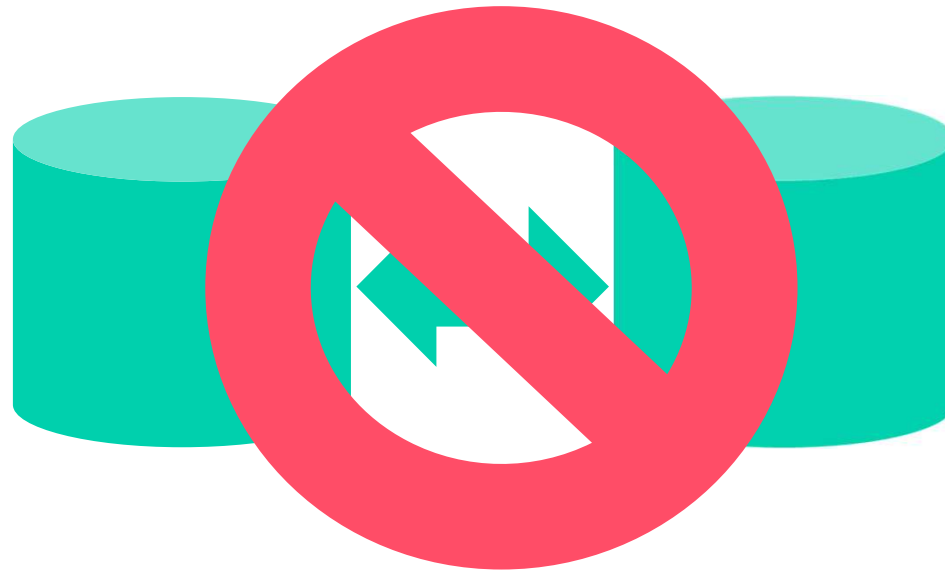
State in the Cloud



Cloud systems:
not highly available

State in the Cloud

Replicate data!



Network might fail

Consistency?

Resilience?

State is Hard

- Let's take an arbitrary example
- Relational database: probably worse
...haven't been built with cloud in mind.

Running Stateful Systems: An Example



Elasticsearch

- Search engine
- Works with structure JSON
- Text, geo ...
- No good fit for update heavy use-cases
- A great piece of software!
- I happen to know it in enough detail.

Elasticsearch Cluster



- Master node: coordinate, create index
- Data node: Store data
- Ingest node: Pre-processing pipelines
- Data sharded
- Shards are replicated



Elasticsearch Cluster

- Kubernetes and Elasticsearch both handle failures of pods
- Kubernetes: Restart pod
- Elasticsearch: Route traffic
- That might be a problem

Kubernetes Operator Pattern

- Advanced automation to
 - Deploy application
 - Take and restore backups
 - Upgrade application code
 - ...

Kubernetes Operator Pattern

- Use Custom Resource Definition (CRD)
- CRD extend Kubernetes configuration
- Controller makes reality match the desired state
- Operator hides pods, storage, stateful sets, backups, upgrades...
- Still IaaS or PaaS?

Run Elasticsearch on Kubernetes



- Use Elastic's Kubernetes operator
- Basic version free
<https://www.elastic.co/subscriptions/enterprise>
- Define #node, master, data, ingest etc
- Quite easy to get started

Elasticsearch Operator / CRDS

Elasticsearch version

of nodes

Might be master

Can store data

Can index document

```
apiVersion: elasticsearch.k8s.elastic.co/v1alpha1
kind: Elasticsearch
metadata:
  name: quickstart
spec:
  version: 7.2.0
  nodes:
    - nodeCount: 1
    config:
      node.master: true
      node.data: true
      node.ingest: true
```

Alternatives

- <https://kubedb.com/> diverse databases
- <https://github.com/upmc-enterprises/elasticsearch-operator>
- <https://www.kubestack.com/catalog/elasticsearch-v0.3.0-kbst.0/>
- ...

**Running Elasticsearch on
Kubernetes?
Easy!**

**We Can Have All the Cloud
Components on Kubernetes!**



Elasticsearch

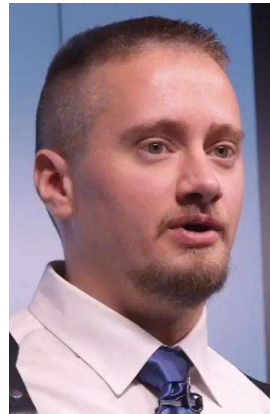


<https://aphyr.com/posts/323-jepsen-elasticsearch-1-5-0> (2015)

Current status – many problems fixed:
<https://www.elastic.co/guide/en/elasticsearch/resiliency/current/index.html>

15 pages if printed

Still: cluster fail in interesting ways



Challenges: Clustered Elasticsearch

- Network partitions
- Backup
- Disaster recovery
- Diverse failure scenarios
- Really reliable distributed systems
- Hard

Elasticsearch Operator

- Installation: easy
- Operation issues: Limited help
- Still need to understand many details

How do Elasticsearch clusters really work?

How do Kubernetes and Elasticsearch deal with clusters?

Other Challenges

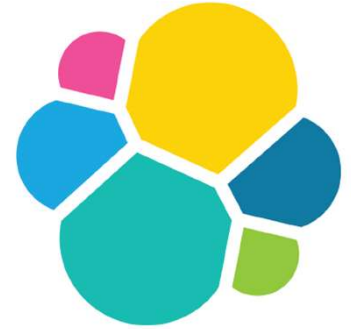
- Patches
 - Upgrades
 - Security updates
-
- Ideally without interrupting service
 - Ideally before exploits appear

Past Security Problems Elasticsearch



- Remote code execution
CVE-2015-5377
- Read arbitrary files
CVE-2015-5531 CVE-2015-3337
- Modify files of other applications
CVE-2015-4165
- Azure credentials might be logged
CVE-2018-3827

Public Cloud: Elastic



- Hosted Elasticsearch on Amazon, Google, and Azure
- Incl. (security) patches, upgrades, snapshots...
- High availability across availability zones
- Support
- <https://www.elastic.co/products/elasticsearch/service>

Public Cloud: Amazon



- Backup, patches
- High availability across zones
- Support
- <https://aws.amazon.com/elasticsearch-service/>

Public Cloud: Amazon



- Licensing issues with Elastic
- Result: Open Distro for Elasticsearch
<https://opendistro.github.io/for-elasticsearch/>
- Amazon is so dedicated to a product that they create their own distribution.

Kubernetes Operators: Sum Up

- Installing a database is easy
- Kubernetes operator
- Running a database reliably is hard
- Operators help
 - ...but you still need to fix issues
- Hosted services do the heavy lifting

Kubernetes Operators: Sum Up

- Do you want to compete with Amazon / Elastic?
- Did you test database disaster recovery, backup?
- Did you test your DB against network partitioning?
- Can you keep your database patching up-to-date?
- Do you dare to update the operator?
- Hosted services might have more people working on Elasticsearch than you have in your IT organization

Kubernetes Operators: Sum Up

- This is just Elasticsearch
- What about supporting
Kubernetes,
Kafka,
relational database, ...

amazon



- <https://fortune.com/2019/03/27/volkswagen-amazon-industrial-cloud/>
- One of Germany's largest companies partners with the Cloud
- Can your IT do better than Volkswagen?

Cloud (In)dependence

Cloud Independence Up Front: Cost



Porting to Different Cloud: Cost



Cloud Independence

- Code should not depend on Cloud solution
- Operations will depend on cloud anyway
- Operations of cloud-independent solution harder
- Can still migrate to an independent solution later.
- Why accept additional costs / risk from the beginning?

Conclusion



Christian Reilly

@reillyusa



Remember a decade ago when people in public cloud used to argue that on prem datacenters were dead? Wonder what they are doing now. Oh yeah, on prem datacenters.



James Watters  @wattersjames · Sep 17

This reminds me of how Sun Microsystems responded to Linux
twitter.com/QuinnyPig/stat...

8:21 AM · Sep 17, 2019 · [Twitter for iPhone](#)

Conclusion

- Kubernetes provides an IaaS Abstraction
- PaaS might be a better fit for microservices
- Kubernetes operator simplify installation of e.g. databases
- ...but operating databases remains complicated

Conclusion

- Public cloud feature: components not cost
- Use the components!
- Invest in abstraction only if needed!
- Avoid unneeded dependencies in the code!

Options

	VM	Kubernetes	Managed
On Premise	Hard to get right	Operator: simplification	Not available
Cloud	VM less reliable	Kubernetes cluster handles unreliable hardware	Component Everything handled for you