



# Operating Elasticsearch in Kubernetes

---

**microxchg** <sup>2019</sup>

MIKKEL LARSEN

[@mikkeloscar](#)

OLIVER TROSIEN

[@otrosien](#)

2019-04-02



# WHO ARE WE?



# WE BRING FASHION TO PEOPLE IN 17 COUNTRIES

**17** markets

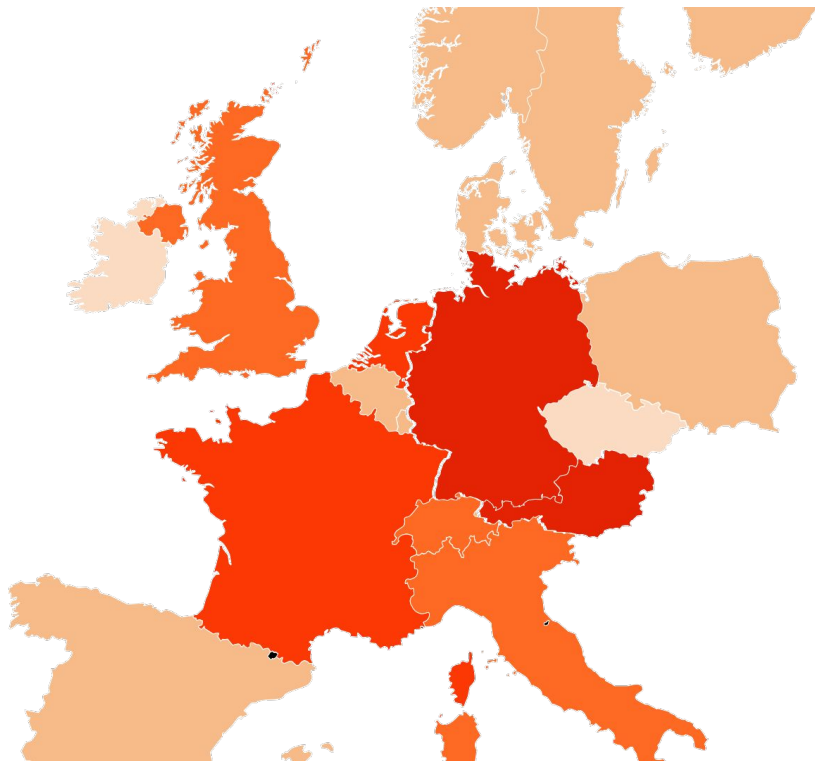
**7** fulfillment centers

**26 million** active customers

**5.4 billion €** revenue 2018

**250 million** visits per month

**15,000** employees in Europe



# SEARCH @ ZALANDO

**300k+**  
Products  
per country

**45%**  
Mobile Traffic

**~2000**  
Brands

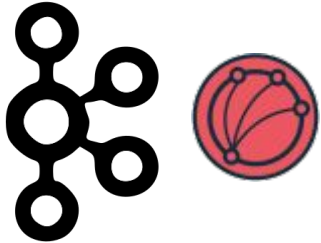
**12K**  
QPS

**~700**  
Categories

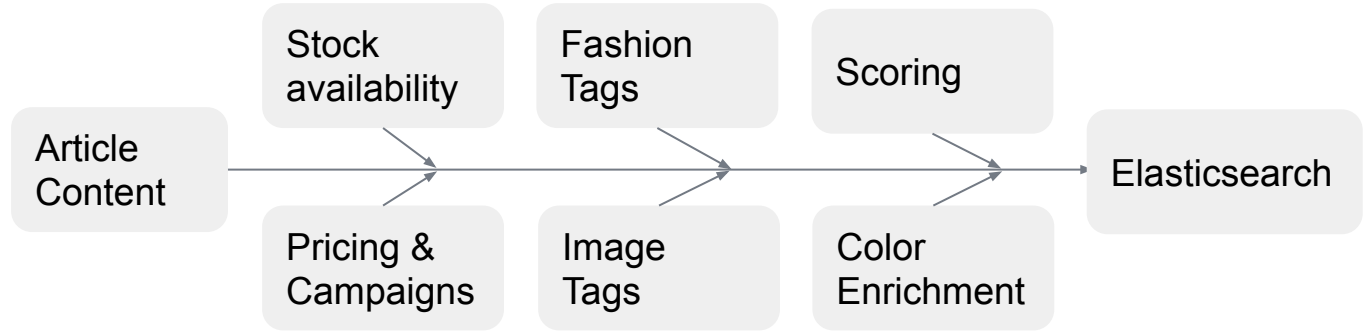
**8K**  
Updates/s



# DATA INGESTION

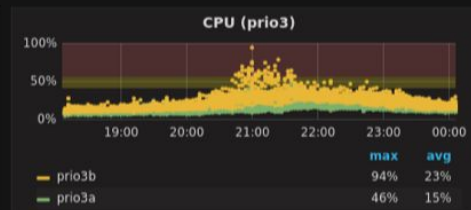
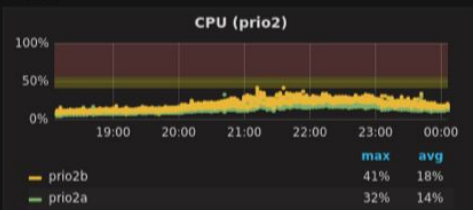
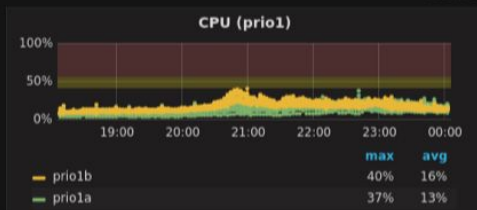
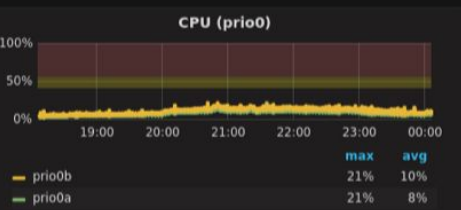
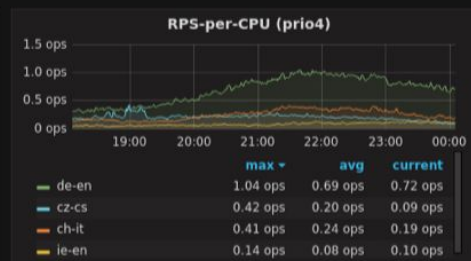
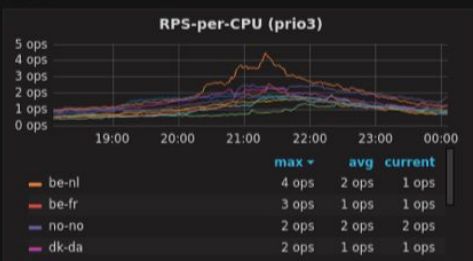
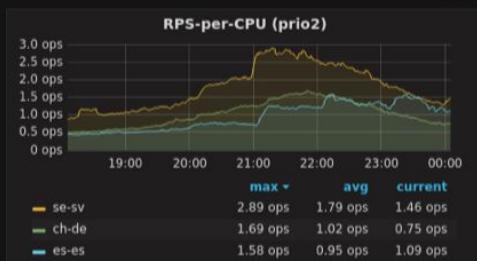
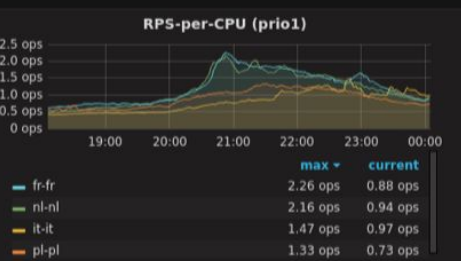
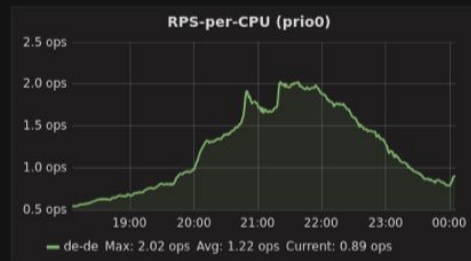
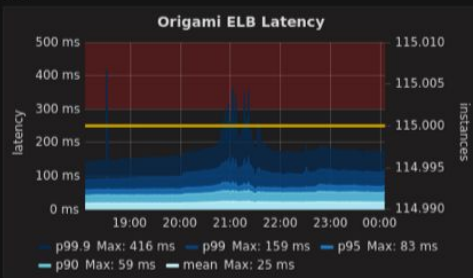
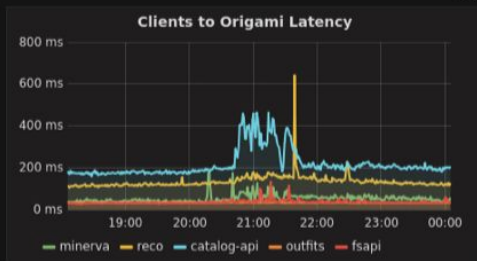
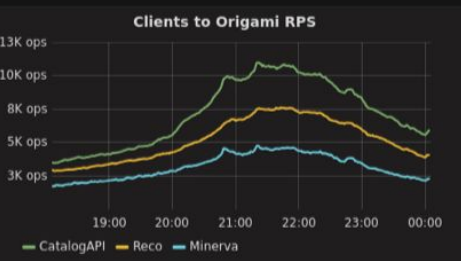
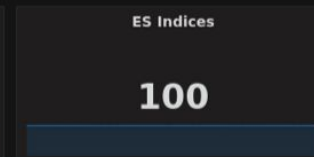
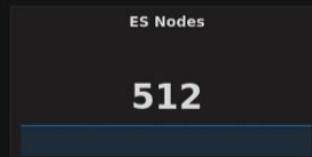
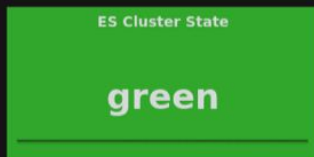
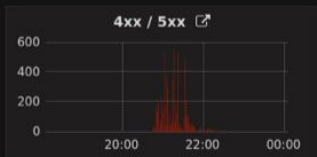
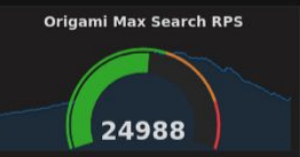


Kafka / Nakadi \*



\* [nakadi.io](https://nakadi.io)



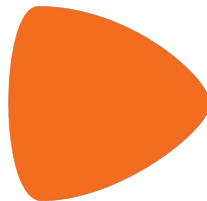


# KUBERNETES @ ZALANDO

**112**  
clusters

**1.3k**  
nodes

Since  
**2016**



zalando

**Default**  
Deployment  
Target

**Node**  
Autoscaling

From **v1.4**  
to **v1.12**

# KUBERNETES CLUSTER SETUP



Master



Master



Config



etcd



Worker



docker



CoreOS



EC2  
Instances



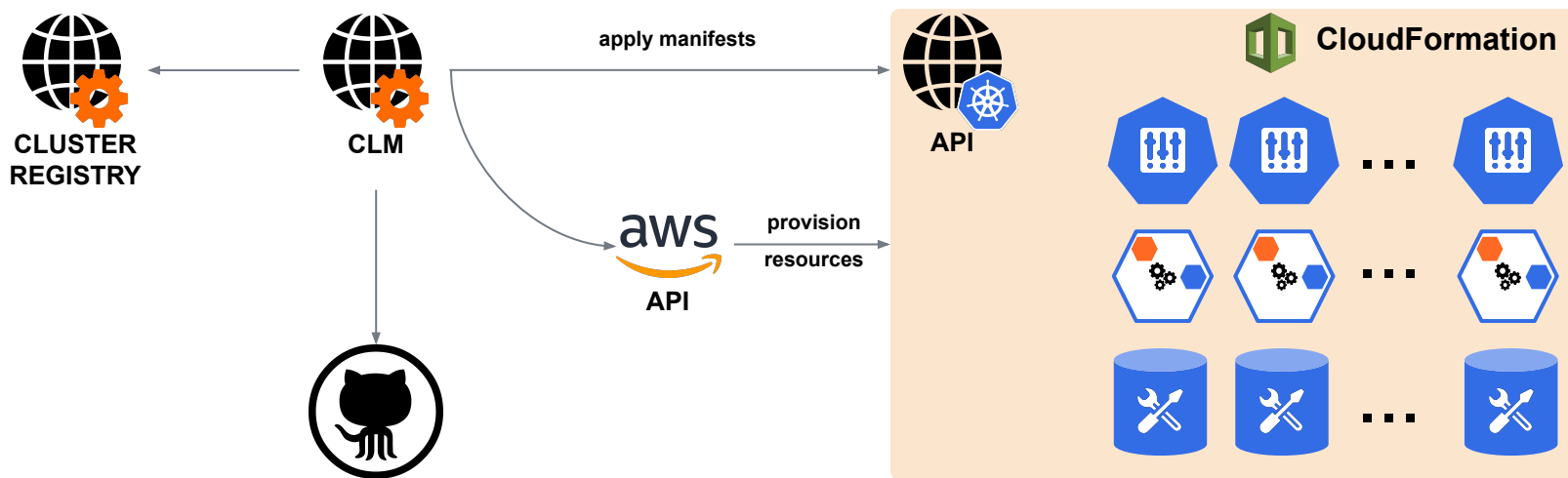
CloudFormation  
Stacks

[github.com/zalando-incubator/kubernetes-on-aws](https://github.com/zalando-incubator/kubernetes-on-aws)



# CLUSTER PROVISIONING

## CLUSTER LIFECYCLE MANAGER (CLM)



[github.com/zalando-incubator/cluster-lifecycle-manager](https://github.com/zalando-incubator/cluster-lifecycle-manager)

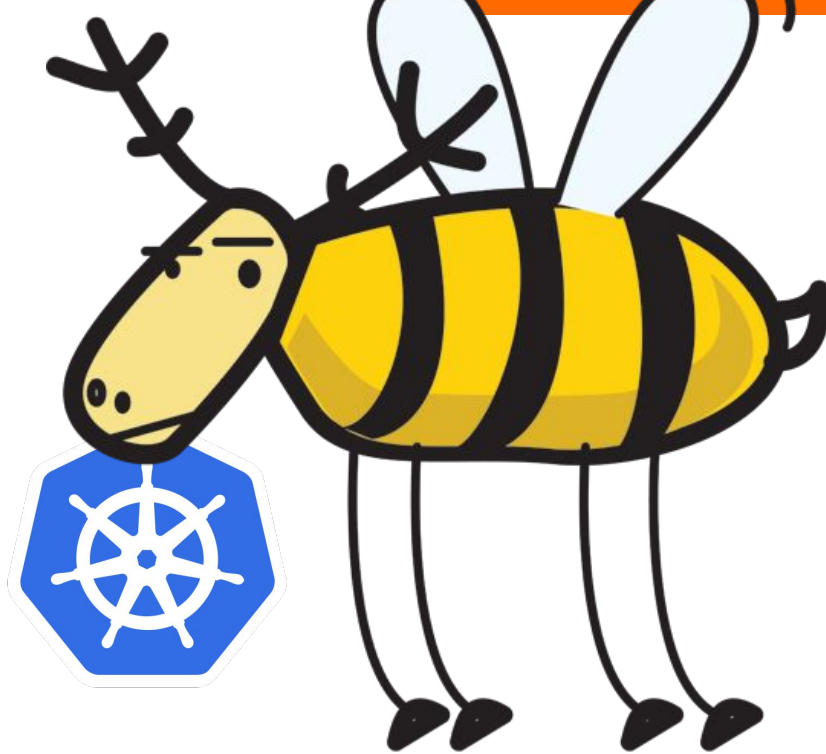
[github.com/zalando-incubator/kubernetes-on-aws](https://github.com/zalando-incubator/kubernetes-on-aws)

## CLUSTER UPDATE POLICY (SLA)

SLA	Production clusters	Test clusters
Minimum lifetime of a Pod before it's force terminated during updates	3 days	8 hours
Minimum amount of time after a node is selected for termination before the infrastructure starts force terminating Pods	6 hours	2 hours
Interval between force terminations of Pods on the same node	5 minutes	5 minutes
Minimum lifetime of ready Pods in the same PDB as the one considered for force termination	1 hour	1 hour
Minimum lifetime of non-ready Pods in the same PDB as the one considered for force termination	6 hours	1 hour

[github.com/zalando-incubator/cluster-lifecycle-manager](https://github.com/zalando-incubator/cluster-lifecycle-manager)

[github.com/zalando-incubator/kubernetes-on-aws](https://github.com/zalando-incubator/kubernetes-on-aws)



Elasticsearch  
2.500 vCPUs  
1 TB RAM

Project “ES Operator”

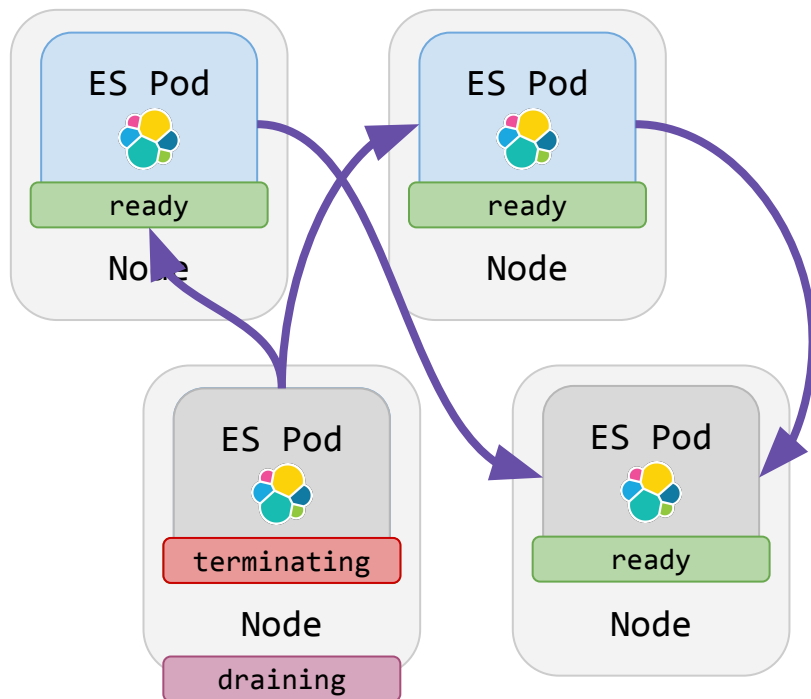
# UPDATING ELASTICSEARCH (STATEFULSET)

## 1) PreStop Hook ([bash script](#))

- Exclude node in ES
- Wait for node to drain (up to 1h)
- Data is moved to existing nodes

## 2) PostStart Hook ([bash script](#))

- Remove all excludes
- Let ES rebalance from existing nodes

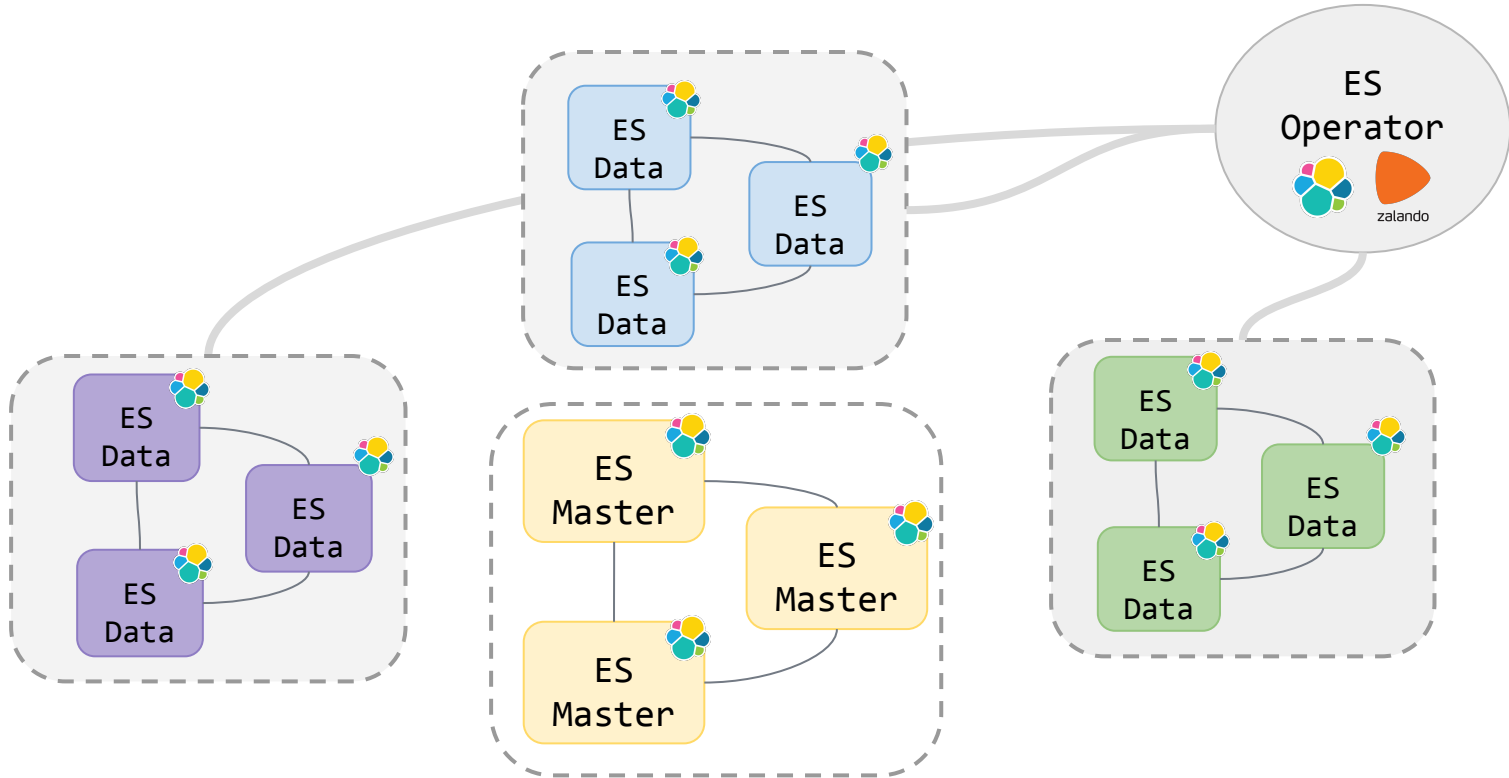


# ELASTICSEARCH DATA SETS

```
apiVersion: zalando.org/v1
kind: ElasticsearchDataSet
metadata:
  name: test-cluster
spec:
  scaling:
    {...}
  replicas: 3
  template: # PodTemplate
    {...}
  volumeClaimTemplates:
    {...}
```

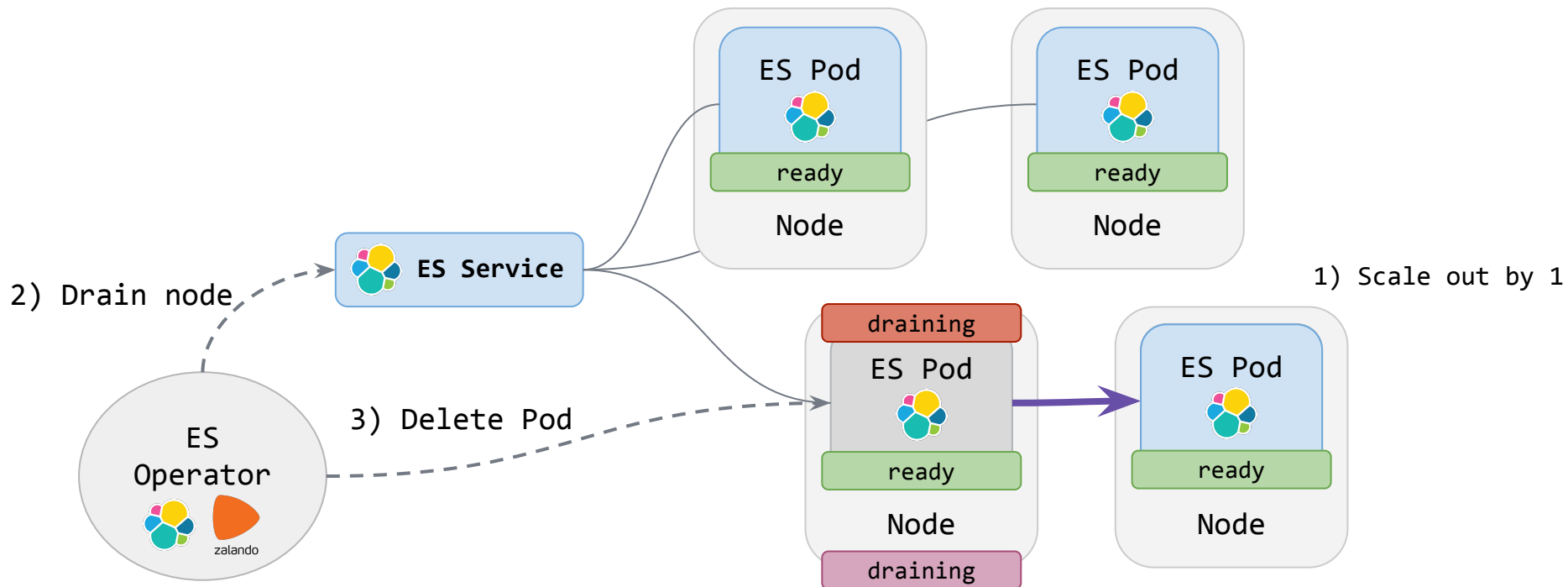
[github.com/zalando-incubator/es-operator](https://github.com/zalando-incubator/es-operator)

# ELASTICSEARCH DATA SETS



[github.com/zalando-incubator/es-operator](https://github.com/zalando-incubator/es-operator)

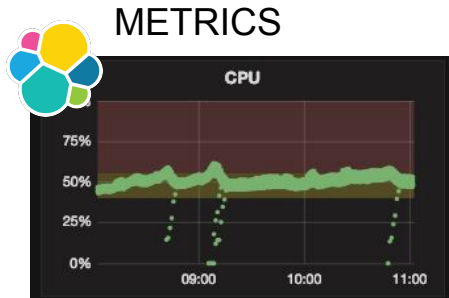
# UPDATING ELASTICSEARCH (OPERATOR)



[github.com/zalando-incubator/es-operator](https://github.com/zalando-incubator/es-operator)



# SCALING UP ELASTICSEARCH (1)

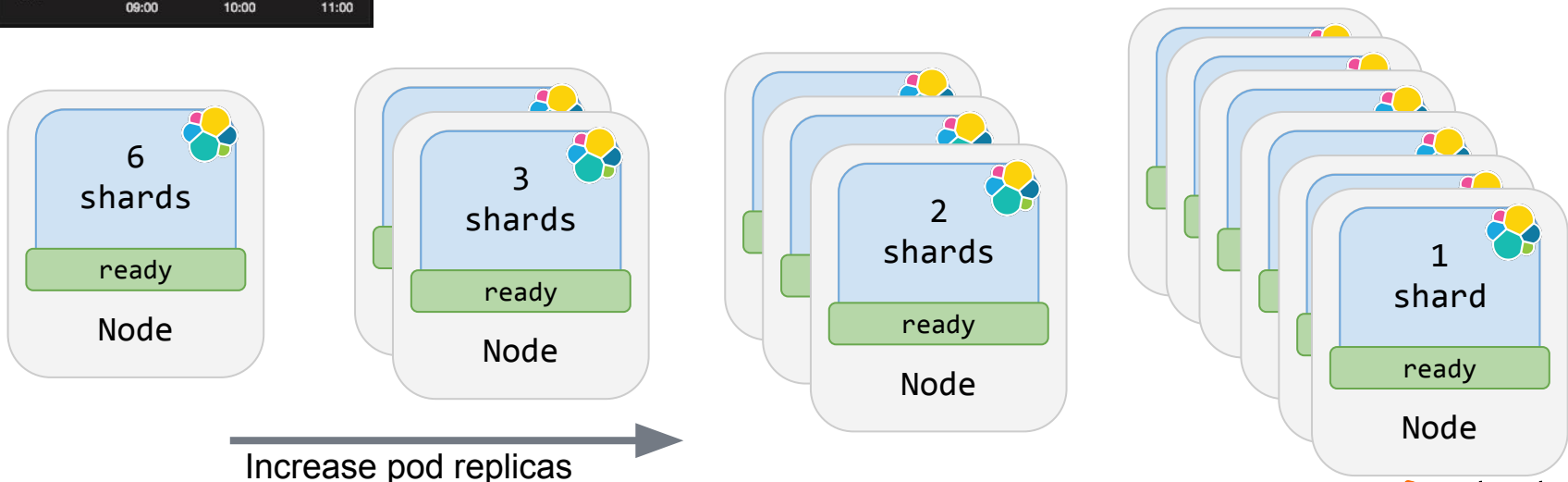


## Thresholds

- CPU
- Duration
- Cooldown

## Boundaries

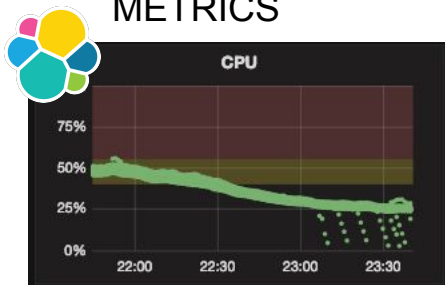
- Max # Pod replicas



# SCALING DOWN ELASTICSEARCH

**DON'T OPERATE  
WHEN CLUSTER  
IS NOT GREEN!**

## METRICS

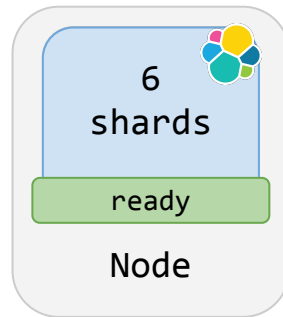
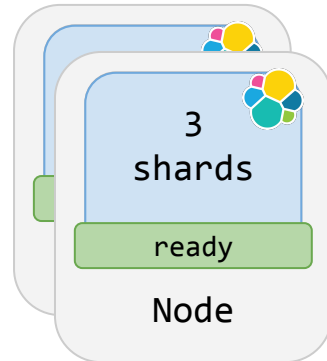
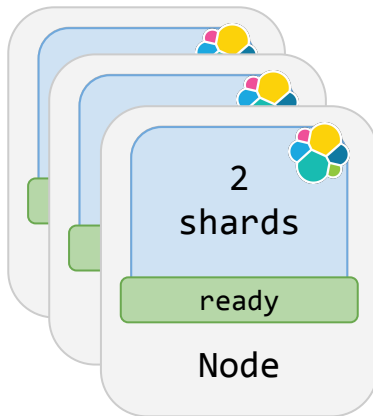
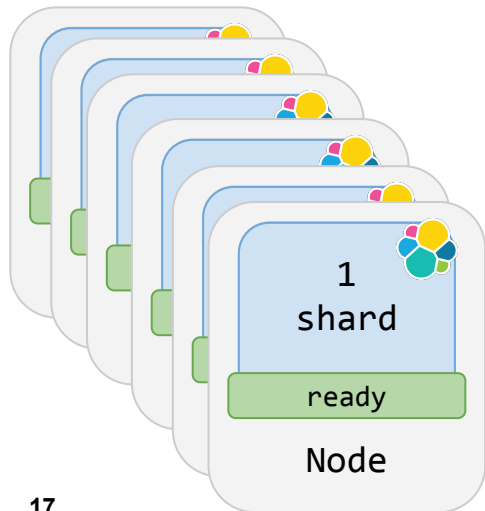


### Thresholds

- CPU
- Duration
- Cooldown

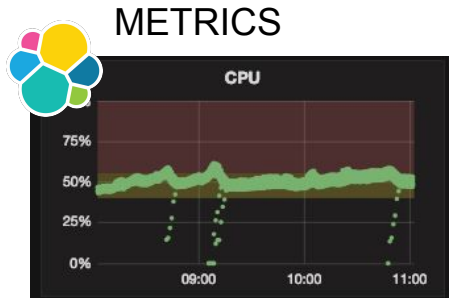
### Boundaries

- Min # Replica
- Max # Shards per node
- Max disk usage (%)



Decrease Pod replicas →

# SCALING UP ELASTICSEARCH (2)

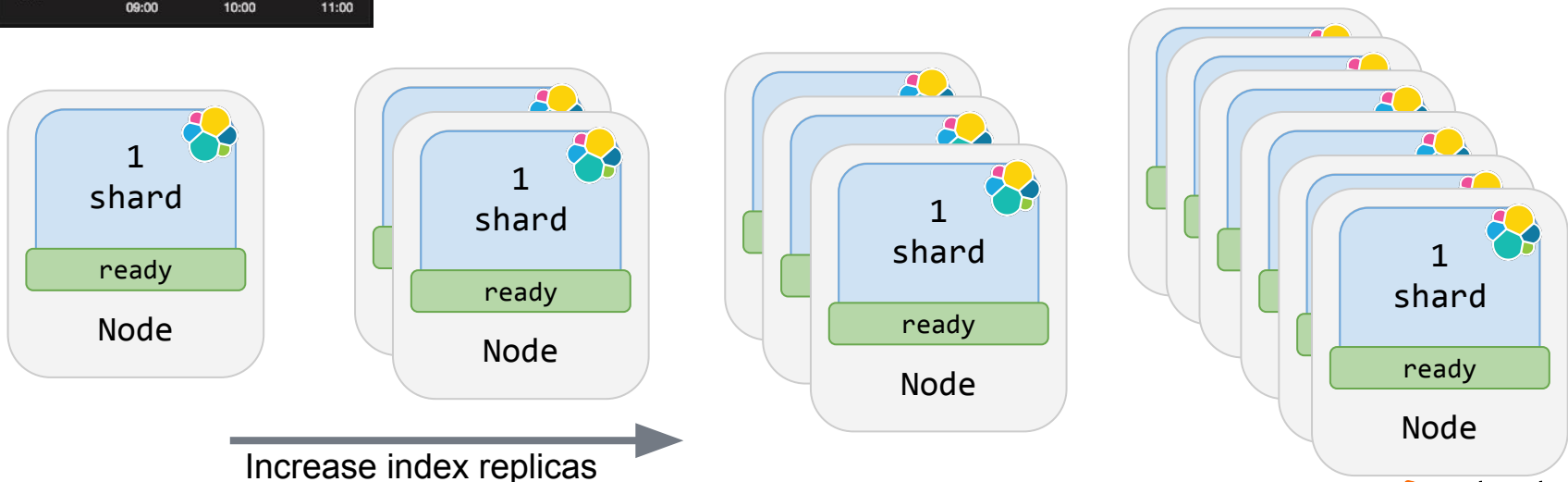


## Thresholds

- CPU
- Duration
- Cooldown

## Boundaries

- Min # Shards per node



DEMO



# OPEN SOURCE

## **ES Operator**

[github.com/zalando-incubator/es-operator](https://github.com/zalando-incubator/es-operator)

## **Cluster Lifecycle Manager**

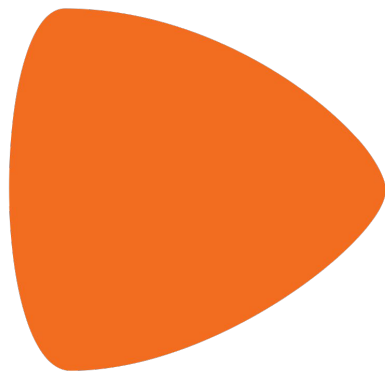
[github.com/zalando-incubator/cluster-lifecycle-manager](https://github.com/zalando-incubator/cluster-lifecycle-manager)

## **Kubernetes on AWS**

[github.com/zalando-incubator/kubernetes-on-aws](https://github.com/zalando-incubator/kubernetes-on-aws)

## **microXchg demo**

[github.com/otrosien/microxchg19-demo](https://github.com/otrosien/microxchg19-demo)



**zalando**





**MIKKEL LARSEN**

[mikkel.larsen@zalando.de](mailto:mikkel.larsen@zalando.de)

[@mikkeloscar](https://www.instagram.com/mikkeloscar)



**OLIVER TROSIEN**

[oliver.trosien@zalando.de](mailto:oliver.trosien@zalando.de)

[@otrosien](https://www.instagram.com/otrosien)



2019-04-02