



# Ablaufplan

19:00 Uhr

Begrüßung

19:15- 19:30 Uhr

Opening InnoMeet Cologne by our sponsor Wolfgang Lang (CEO Crowdfox GmbH)

19:30- 20:00 Uhr

@ocramius Marco Pivetta presents Extremly Defensive PHP

20:15- 20:45 Uhr

The crowdfox **DevOps-Team** presents **Big Data Integration Environments** 

20113 20113 0111

Eat & Meat at Hans im Glück

21:00- Open End





# Opening





Extremely Defensive PHP

Big Data Integration Environments Feedback & Networking





# **Meeting Motivations**

## **New Technologies**



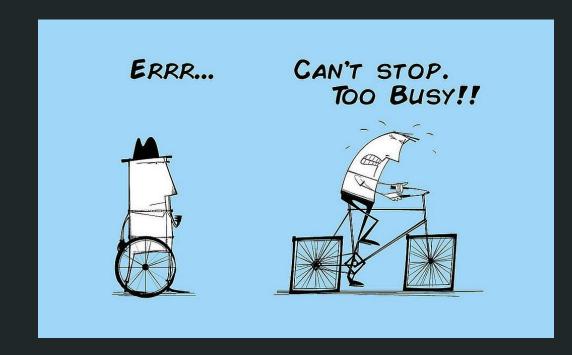




# **Meeting Motivations**

### **Exchange innovative solutions**

#noWarlmmerSo
#noTooBusyToImprove







# **Meeting Motivations**

#### **Not yet another branded Meetup**

#noBuzzwordBingo
#noCultureCult







# Warum dieses Meetup

#### **Real Experiences**

We want:
#realXP
#realMistakes
#realLearnings







# **Sponsor**





# **Extremely Defensive PHP**



Extremely Defensive PHP

Big Data Integration Environments Feedback & Networking





# **Extremely Defesive PHP**

http://ocramius.github.io/extremely-defensive-php/#/



# Big Data Integration Environments



Extremely Defensive PHP

Big Data Integration Environments Feedback & Networking





# **Big Data Integration Environments**







### Intro

#### **Common Problems with Integration Environments**

- You need more than one integration environment #bottleneck
- Budget(Integration) <= Budget(Production)</p>
  - #small datasets
- Data must be anonymized
  - #systemmails
  - #GDPR
- Data must be up to date
  - #migrations
  - #accounting
  - #sliding windows queries and algorithms





### Intro

#### What about "Big Data" Algorithms?

- You need more than one integration environment
  - because the local machine is not big enough
  - local machine is too slow
- Budget(Integration) <= Budget(Production)</p>
  - 🖴 small datasets 😂 😂 😂
- Data must be anonymized
- Data must be up to date
  - meaningful
  - representative





### Intro

#### **Current Integration Environment**

- You need more than one integration environment
  - 9 possible Evironments
- Budget(Integration) <= Budget(Production)</p>
  - 3 Integration Servers vs 20 Production Server
    - 1 x DB
    - 1 x FE
    - 1 x BE
- Data must be anonymized and Data must be up to date
  - Updating all DB Slots takes ~ 6 Days
  - 1 DB server runs 9 DB Instances
  - Every instance has to be updated seperatly
  - Nearly no case specific customization, because this would create side effects





# **Big Data Integration Environments**







#### **Gitlab Review Apps**



#### **Automatic Live Preview**

Code, commit, and preview your branch in a live environment. Review Apps automatically spin up dynamic environments for your merge requests.

#### One-click to Collaborate

Designers and product managers won't need to check out your branch and run it in a staging environment. Simply send the team a link and let them click around.

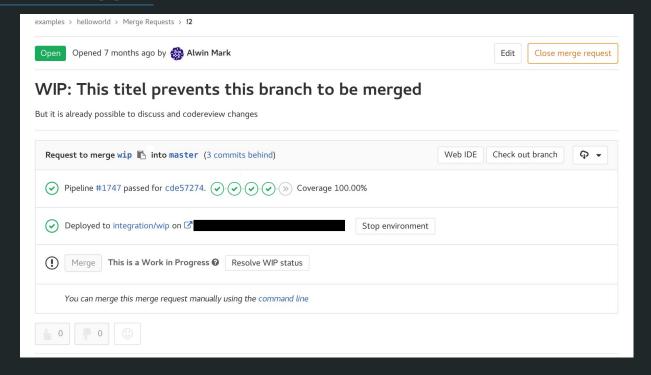
#### Fully-Integrated

With GitLab's code review, built-in CI/CD, and Review Apps, you can speed up your development process with one tool for coding, testing, and previewing your changes.





#### **Gitlab Review Apps**







### **Gitlab Review Apps**

```
deploy integration:
      stage: deploy
      image: registry.crowdfox.me/devops/docker-rancher-compose:latest
      variables:
        RANCHER URL:
        RANCHER ACCESS KEY:
        RANCHER SECRET KEY:
        RANCHER ENVIRONMENT:
      script:
54
        - rancher --env $RANCHER ENVIRONMENT rm -s helloworld/service-$CI COMMIT REF NAME && echo "wait until container is down" && sleep 10 || true
        - rancher --env $RANCHER ENVIRONMENT run --name helloworld/service-$CI COMMIT REF NAME -p 80 $CI REGISTRY IMAGE/temporary:$CI COMMIT SHA
      environment:
        name: integration/$CI COMMIT REF NAME
        on stop: stop integration
      only:

    branches

      except:
        - master
    stop integration:
      stage: cleanup
      image: registry.crowdfox.me/devops/docker-rancher-compose:latest
      variables:
        RANCHER URL: 'http://94.130.22.27:8080/'
        RANCHER ACCESS KEY: '37F2CCE9C7A5A589F907'
        RANCHER SECRET KEY: $RANCHER API SECRET KEY INTEGRATION
        RANCHER ENVIRONMENT: 1a16
      script:
      script:
        - rancher --env $RANCHER ENVIRONMENT rm -s helloworld/service-$CI COMMIT REF NAME
      when: manual
      environment:
78
        name: integration/$CI_COMMIT_REF_NAME
        action: stop
```





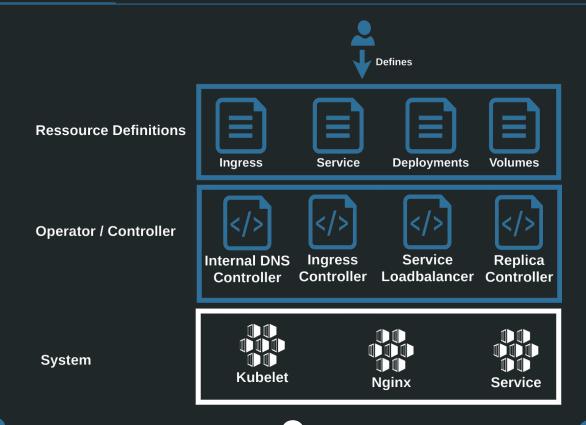
### **Kubernetes Operators and Controller Logic**







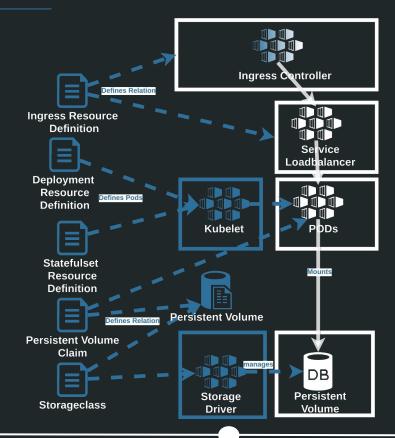
## Kubernetes - High level architecture







#### **Kubernetes - Common resources**







#### Rook

- 鲎 Hyper-scale or Hyper-converged
  - Dedicated Rook/Storage Kubernetes Cluster
  - Single "Hyper-converged" Kubernetes Cluster
- Works as an Kubernetes Operator for Ceph Cluster 鲎
  - Manages small and medium failures
  - Cares about basic configurations
  - Is able to create Pools

Also because it just manages Ceph:

- Pretty stable 鲎
- Scales at will 鲎 ▲ I mean with your Network connection. (Recommended 10 Gb/s)

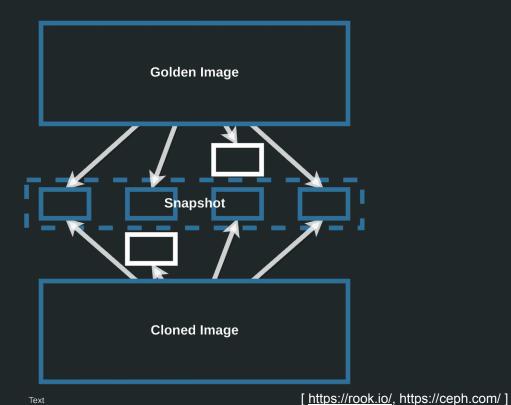






#### **Block Storage**

- Only changed Blocks will be copied #copyOnWrite
- Common Blocks are shared #deduplication







# **Big Data Integration Environments**







# **New Integration Environment**

Demo

https://gitlab.com/innomeat.cologne/bigdata-integration-environment







## **New Integration Environment**

#### What we have achieved

- You need more than one integration environment
  - As many Environments as there is spare CPU and RAM (and disk space for the deltas)
- Budget(Integration) <= Budget(Production)</p>
  - ■3 Node K8s App Cluster + 1 Node DataNode <= 3K8s App Nodes + 3 Node Mysql Replication
- Data must be anonymized and Data must be up to date
  - Updating all DB Slots takes < 2h (But during this time we can not deploy new Environments)
  - Application and Data is separated and can be scaled separately
  - One update Process for all
  - "Pheonix setup" allows specific customization





# **Questions**



Join us on Slack: <a href="https://innomeet.cologne/slack.html">https://innomeet.cologne/slack.html</a>





### **Questions**

#### **More often than Daily**

- 1. Breeder creates temporary image, snapshot and protect
- 2. Clone temporary image to golden image
- 3. RBD "flatten" golden image
- 4. create new snapshot on the golden image and protect it
- 5. Use this snapshot for environments
- 6. unprotect temporary breeder image and delete it
- 7. repeat





## **Questions**

#### What is Hyperconverged

Managed all by the same Hypervisor

In Rook context it means, that one Kubernetes Cluster is managing:

- Applications
- Network
- Storage



