

22.03.2025 / Robert Sander

# Der Ceph Orchestrator

## Container für Storage

# Inhalt



- Was ist Ceph?
- Was macht der Orchestrator?
- Hilfe, Container!
- Jetzt auch noch Kubernetes!

# Wer wir sind

# Sichere & freie Kommunikation ist unser Sourcecode



30

Jahre Erfahrung

56

Akademiekurse

6

Linux-Standardwerke

100+

Mitarbeiter

2

Rechenzentren

100%

Berlin

# Die Heinlein-Gruppe: Gemeinsam für digitale Souveränität



## Heinlein Support

- **Akademie:** Für die oberen 10% des Wissens – unsere Linux-Schulungen für IT-Experten.
- **Consulting:** Security- und Strategieberatung, Projektumsetzung und umfassender Support für IT-Administratoren
- **Services:** SLA-Verträge, Hosting und Lizenzen als Unterstützung & Absicherung Ihrer kritischen IT-Infrastruktur

## Weitere Marken

- **mailbox.org:** E-Mail, Online-Office, Cloud-Speicher und Videokonferenzen nach neuesten Sicherheitsstandards und mit grüner Energie.
- **OpenTalk:** Videocalls, wie sie sein sollten – mit unserer sicheren, benutzerfreundlichen und skalierbaren Videokonferenz für Behörden, Provider, Unternehmen und Schulen.
- **OpenCloud:** Zuverlässiges Filesharing & Content Collaboration – gemacht für höchste Enterprise-Ansprüche in puncto Performance, Stabilität und Skalierbarkeit.

# Unser Profil



Sicher und freie  
Kommunikation ist für  
uns keine Floskel – es  
ist unser Source Code  
seit Jahrzehnten

## Robert Sander

- Senior Linux Consultant
  - Ceph
  - CheckMK
  - Samba
  - LDAP
  - Netzwerk, Routing, Firewall
- 2000 Dipl.-Inf.
- 2012 Heinlein Support



Ceph

# Ceph bietet ...



## einen Storage-Cluster

- der sich selbst verwaltet
- der sich selbst heilt
- ohne Engpässe

## drei Schnittstellen

- Objektspeicher (kompatibel zu S3)
- Blockspeicher (für VMs, etc)
- verteiltes Dateisystem



# Ceph Komponenten: Object Storage Daemon

## OSD speichert Daten auf

- HDD
- SSD
- NVMe
- oder was es noch geben wird

## Ein Prozess pro Blockdevice



# Ceph Komponenten:

## Monitore

MONs bilden das Gehirn des Clusters

Quorum für Entscheidungen

Verwaltung Cluster-Map



Ceph Komponenten:

Manager

MGRs unterstützen die Administration

Dashboard

Prometheus-Exporter (oder anderes Monitoring)

REST-API

Orchestrator



Ceph Komponenten:  
Metadata Server  
MDS für Ceph Filesystem

Verwaltung der POSIX-Daten:

Verzeichnisse

Zugriffsrechte, ACLs

Zeitstempel

Locking



Ceph Komponenten:  
Rados-Gateway  
für S3 und Swift  
HTTP-basiertes Object-Storage  
mehrere Instanzen  
"Ingress"-Dienst für Service-IP und Load-Balancing

# Ceph Komponenten:



## Ganesha-NFS

Frontend für CephFS

## Samba

Bald auch möglich

## Disaster Recovery

cephfs-mirror

rbd-mirror

## Monitoring-Stack

Prometheus

Grafana

Alert-Manager

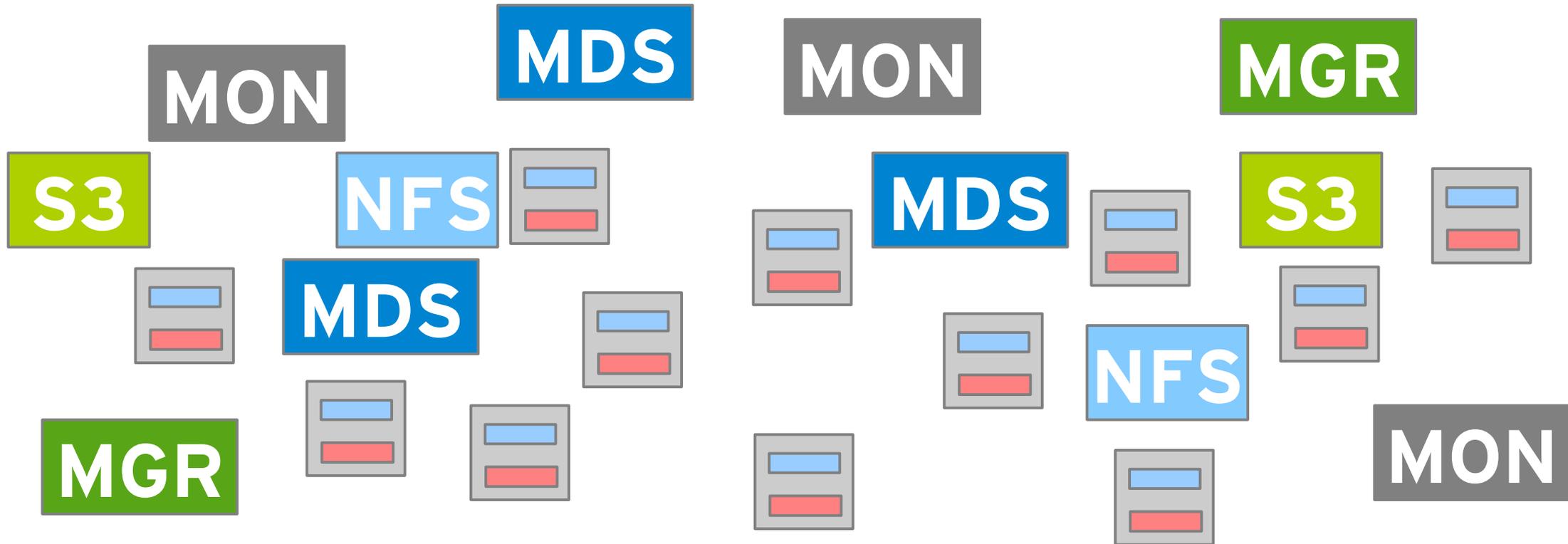
## iSCSI

abgekündigt

## NVMe over Fabric

kommt

# Und schon ist der Ceph-Cluster fertig





# Orchestrator

# Orchestrierung



In system administration, orchestration is the automated configuration, coordination, deployment, development, and management of computer systems and software.

In der Systemadministration ist Orchestrierung die automatisierte Konfiguration, Koordination, Bereitstellung, Entwicklung und Verwaltung von Computersystemen und Software.

[https://en.wikipedia.org/wiki/Orchestration\\_\(computing\)](https://en.wikipedia.org/wiki/Orchestration_(computing))

# Orchestrierung



Ceph ist ein verteiltes System auf vielen Servern

ceph-deploy gab es bis Version 14

Python 2

ceph-ansible Playbooks

"The project is still maintained for the time being but it is encouraged to migrate to cephadm."

seit Version 15: Orchestrator als MGR-Modul



# Orchestrator

## Zwei Implementierungen

cephadm

rook

CLI

Kubernetes

Dashboard-Integration

kubectl

steuert Services und  
Daemons via SSH

Helm-Charts

Beide nutzen Container als "Distributions-Pakete"



cephadm



cephadm

Bootstrapping

Initiale Installation

1 MON

1 MGR

→ funktionierender Ceph-Cluster

<https://docs.ceph.com/en/latest/cephadm/>



# cephadm

## Host-Verwaltung

### Orchestrator hat eigenen SSH-Schlüssel

### Public-Key auf Clusterknoten verteilen

### Clusterknoten als Host hinzufügen

Erzeugt gleich Daemons:

- ceph-crash
- node-exporter
- ceph-exporter
- ggfs weitere



cephadm

Services

Wo läuft was?

Spezifikation: YAML

Service-Typ + ID

Placement über Hostname, Label, Anzahl

Spezifische Attribute für bestimmte Services



cephadm

Standard-Services

mon

mgr

osd

ceph-crash

node-exporter, ceph-exporter

1x Monitoring-Stack

prometheus, grafana, alertmanager



cephadm

weitere Servicetypen

cephfs-mirror

ingress

iscsi

mds

nfs

nvmeof

cephfs-mirror

rbd-mirror

rgw



# cephadm

## Daemons

## Instanzen von Services

## als Prozess in Container ausgerollt

Podman besser als Docker

## Systemd-Unit mit Journal

```
systemctl {status|start|stop|restart} ceph- $\$$ CLUSTERID@daemonname.service
```

```
journalctl -u ceph- $\$$ CLUSTERID@daemonname.service
```

## Host-Networking

```
/var/lib/ceph/ $\$$ CLUSTERID/
```



# cephadm

## Service-Spezifikation

```
service_type: mon
service_name: mon
placement:
  count: 5
  label: mon
```

```
service_type: osd
service_id: SSDs
service_name: osd.SSDs
placement:
  host_pattern: '*'
spec:
  data_devices:
    rotational: false
    size: 2000G:
```



# cephadm

## Container Service

ermöglicht Daemons mit  
eigenen Images

### Beispiel s3manager

```
service_type: container
service_id: s3manager
placement:
  hosts:
    - cephtest20
spec:
  args:
    - --publish=8080:8080/tcp
  envs:
    - ACCESS_KEY_ID=XYZ
    - SECRET_ACCESS_KEY=ABC
    - ENDPOINT=radosgw0.example.net
    - USE_SSL=false
    - PORT=8080
  image: docker.io/cloudlena/s3manager
  ports:
    - 8080
```



# cephadm

## Kommandozeile

```
ceph orch -help
```

```
ceph orch ls [--export]
```

```
ceph orch apply -i service.yml
```

```
ceph orch ps
```

```
ceph orch host {add|ls|maintenance|drain|rm}
```

```
ceph orch device ls
```

```
ceph orch daemon {start|stop|restart}
```

```
ceph orch osd rm {osdid}
```

```
ceph orch upgrade start --ceph_version 19.2.1
```



# Warum Container?

# Ceph in Containern



Immer wieder Probleme mit Distri-Paketen

→ Bibliotheksversionen

→ Version von gcc

Container-Image ist definiertes Binary mit allen Libs,  
lauffähig auf allen aktuellen Kernen

Ceph-Cluster sieht überall "gleich" aus

# Ceph in Containern



## Quellen Container-Images

quay.io

öffentliche Registry von RedHat / IBM

Registry-Proxy

eigene Registry

eigene Container-Images



# Rook für Kubernetes

# rook



Storage Operator für K8s

K8s auf Bare-Metal oder

Externes Ceph für K8s

<https://rook.io/>

CNCF-Projekt

# rook



Container Storage Interface für K8s

CustomResourceDefinitions

Storage aus

- Raw Devices / Partitionen
- Logical Volumes (LVM)
- Persistent Volumes in block mode



rook

## Bootstrapping

operator.yaml anpassen, dann

```
kubectl create -f crds.yaml -f common.yaml \  
-f operator.yaml
```

```
kubectl -n rook-ceph get pod
```

oder per Helm Chart



Demo



# Fragen und Diskussionen

# CompetenceCall



Das Backup für Ihre  
Server-Administration.

Nutzen Sie unsere  
SLA-Verträge und sichern  
Sie sich den 24/7-Support  
unserer Linux-Consultans.

- Kontinuierliche Absicherung mit garantierten Reaktionszeiten und festen SLAs
- Rückendeckung im Notfall: mindestens LPIC-2 zertifizierte Profis mit jahrelanger, täglicher Admin-Erfahrung
- Projektunterstützung: maßgeschneiderte Lösungen, die Flexibilität, Sicherheit, Administrierbarkeit und Hochverfügbarkeit vereinen
- Services: Performanceanalyse, Serverhärtung, Netzwerkanalyse, Konfigurationshilfe, Datenrestaurierung

# Werde Teil des Teams

- Du bist neugierig, voller Tatendrang und überzeugt von Linux, Open Source und sicherer, freier Kommunikation?
- Wir freuen uns über Unterstützung im Team:  
[www.heinlein-support.de/jobs](http://www.heinlein-support.de/jobs)





# Bleiben wir im Kontakt

Robert Sander

Tel. +49 30 40 50 51-40  
[r.sander@heinlein-support.de](mailto:r.sander@heinlein-support.de)

Heinlein Support GmbH  
Schwedter Straße 8/9 | 10119 Berlin  
[www.heinlein-support.de](http://www.heinlein-support.de)