PROXMOX VE OVERVIEW

3 Features
4 Ausbauvariante 1 (Single Node)
5 Clustering Vorteile
6 Ausbauvariante 2 (Cluster ohne HA)
7 Ausbauvariante 3 (async. HA-Cluster mit ZFS)
9 Ausbauvariante 4 (sync. HA-Cluster mit Ceph)
10 Hochverfügbarkeit
11 Ceph Einführung & Features
18 Live-Demo
19 Ceph ausführlichere Informationen
27 Ceph Deployment Varianten
29 Ceph Benchmarks
31 Service & Support
33 Proxmox VE bei der Thomas-Krenn.AG

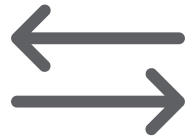
Features

Snapshots

KVM Virtualisierung
(Windows, Linux, BSD)

LXC Container

Live Migration



High-Availability Clustering

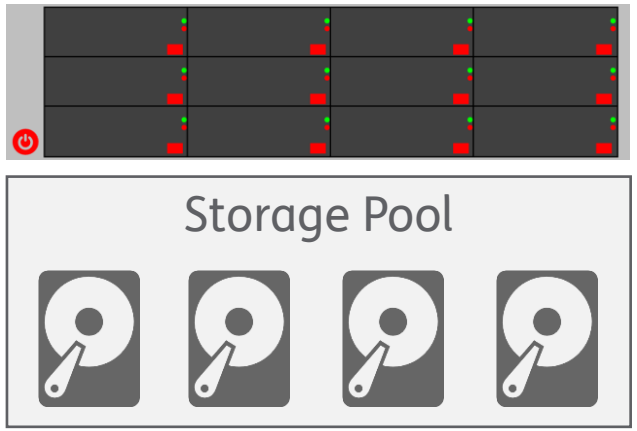
Flexible Storage-Wahlmöglichkeiten

GUI Management

Proxmox Backup Server

Ausbau-Varianten

Single Node



Clustering Vorteile

Live Migration

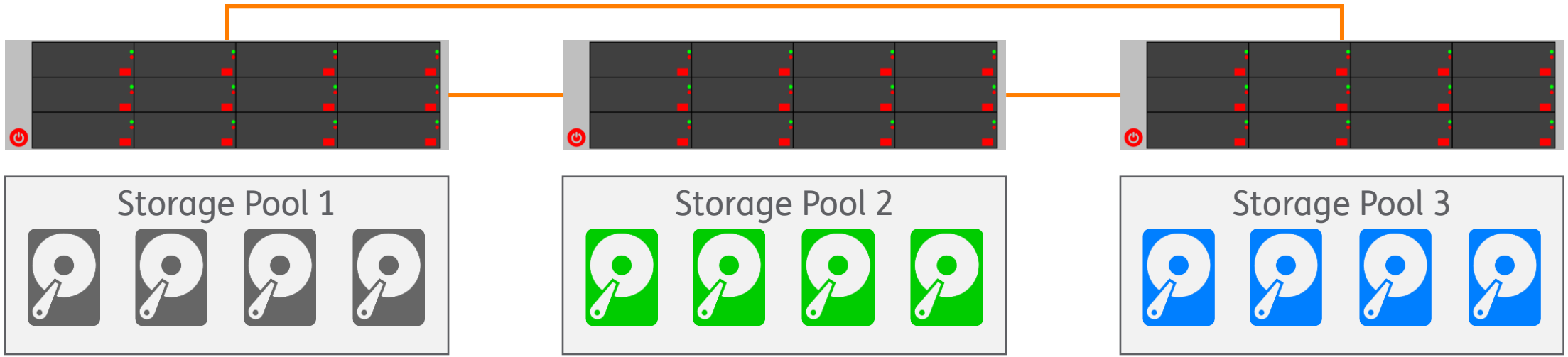
Voraussetzung für Hochverfügbarkeit

Zentrale Verwaltungsoberfläche

Globales Server-Management

Ausbau-Varianten

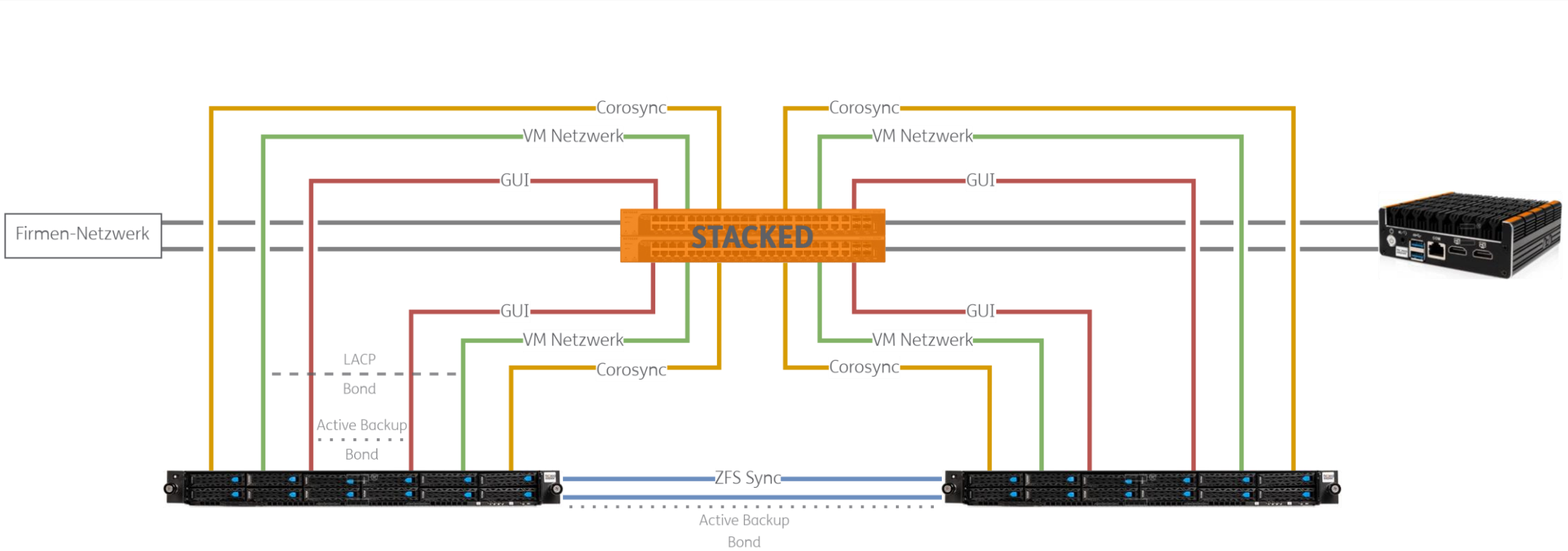
Clustering ohne HA



Ausbau-Varianten

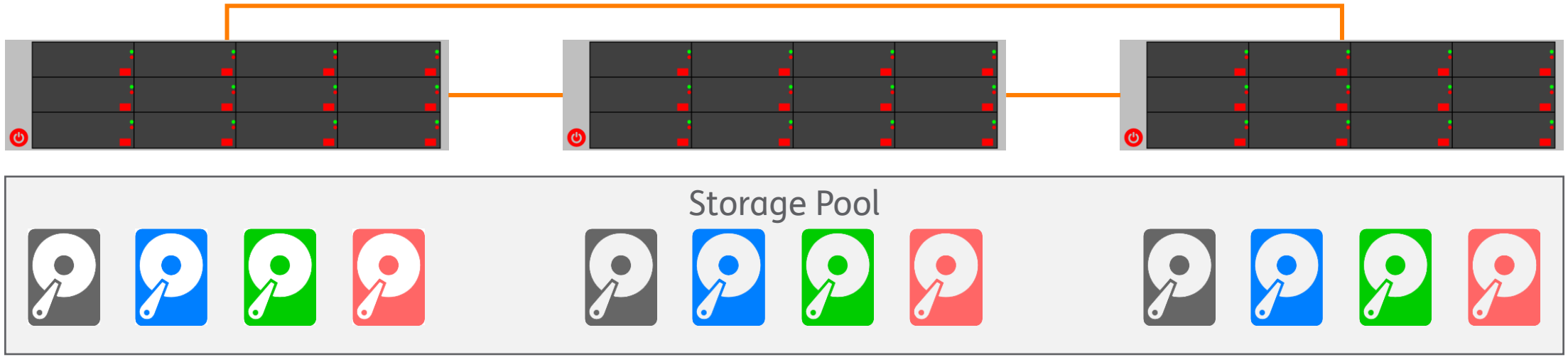
ZFS Cluster mit asynchroner Replikation

ZFS Dual Node Cluster mit Proxmox VE



Ausbau-Varianten

Ceph Cluster



Hochverfügbarkeit

Ausbaustufen mit Proxmox VE

Single Node

- Single Points of Failure (SPoF)
- Redundanz auf Disk-Ebene
- Downtime bei Serverwartung
- Hohe Wahrscheinlichkeit von Datenverlust

Einfache Verfügbarkeit

ZFS Dual Node

- Kein SPoF
- Redundanz auf Hardware-Ebene
- Serverwartung bei Uptime
- Datendelta durch asynchrone Spiegelung

Cluster-Verfügbarkeit

Ceph Cluster

- Kein SPoF
- Redundanz auf allen Ebenen
- Serverwartung bei Uptime
- Kein Datendelta

Hoch-Verfügbarkeit

Was ist Ceph?

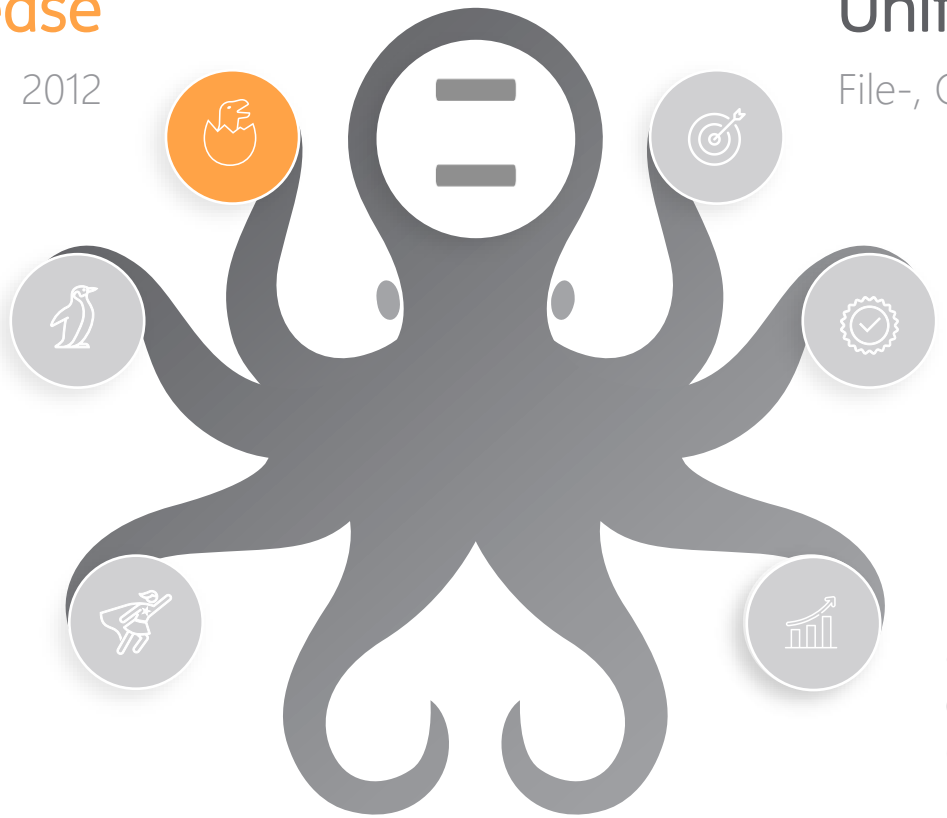
Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage



Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Was ist Ceph?

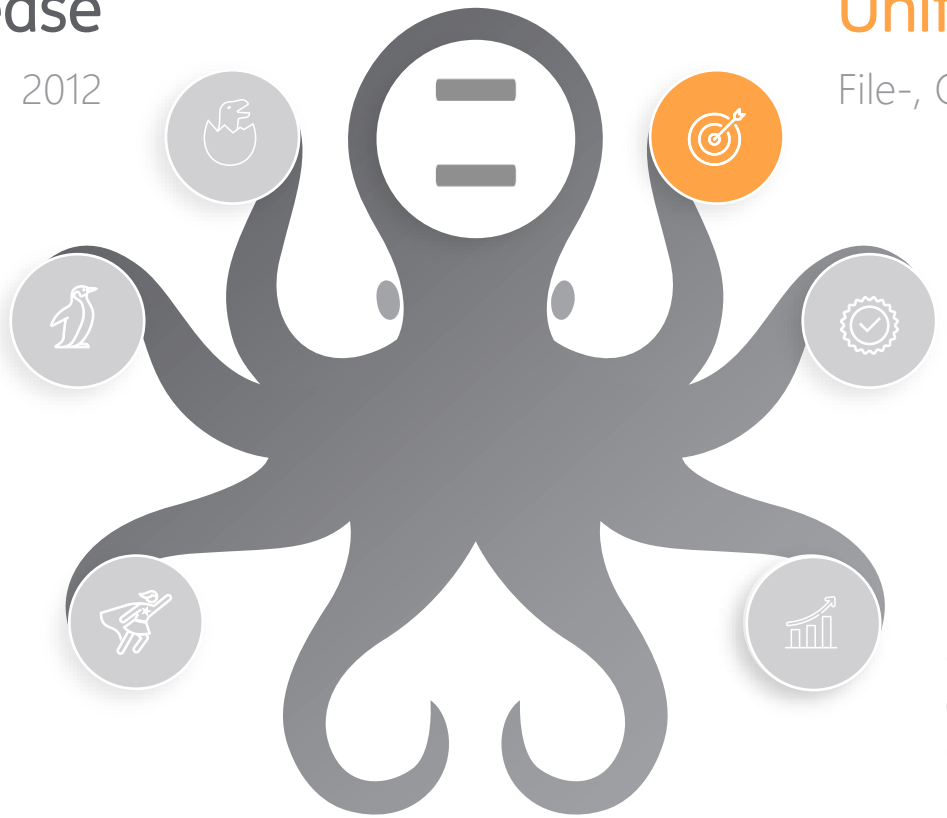
Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage



Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Was ist Ceph?

Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage

Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Was ist Ceph?

Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage

Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Was ist Ceph?

Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage

Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Was ist Ceph?

Selfhealing, Selfmanaging

Release

2012

Unified Storage

File-, Object-, Blockstorage

Open Source

Verteilte Open Source Storage Lösung

Verlässlich

Daten werden mind. 3x gespeichert (SIZE=3)

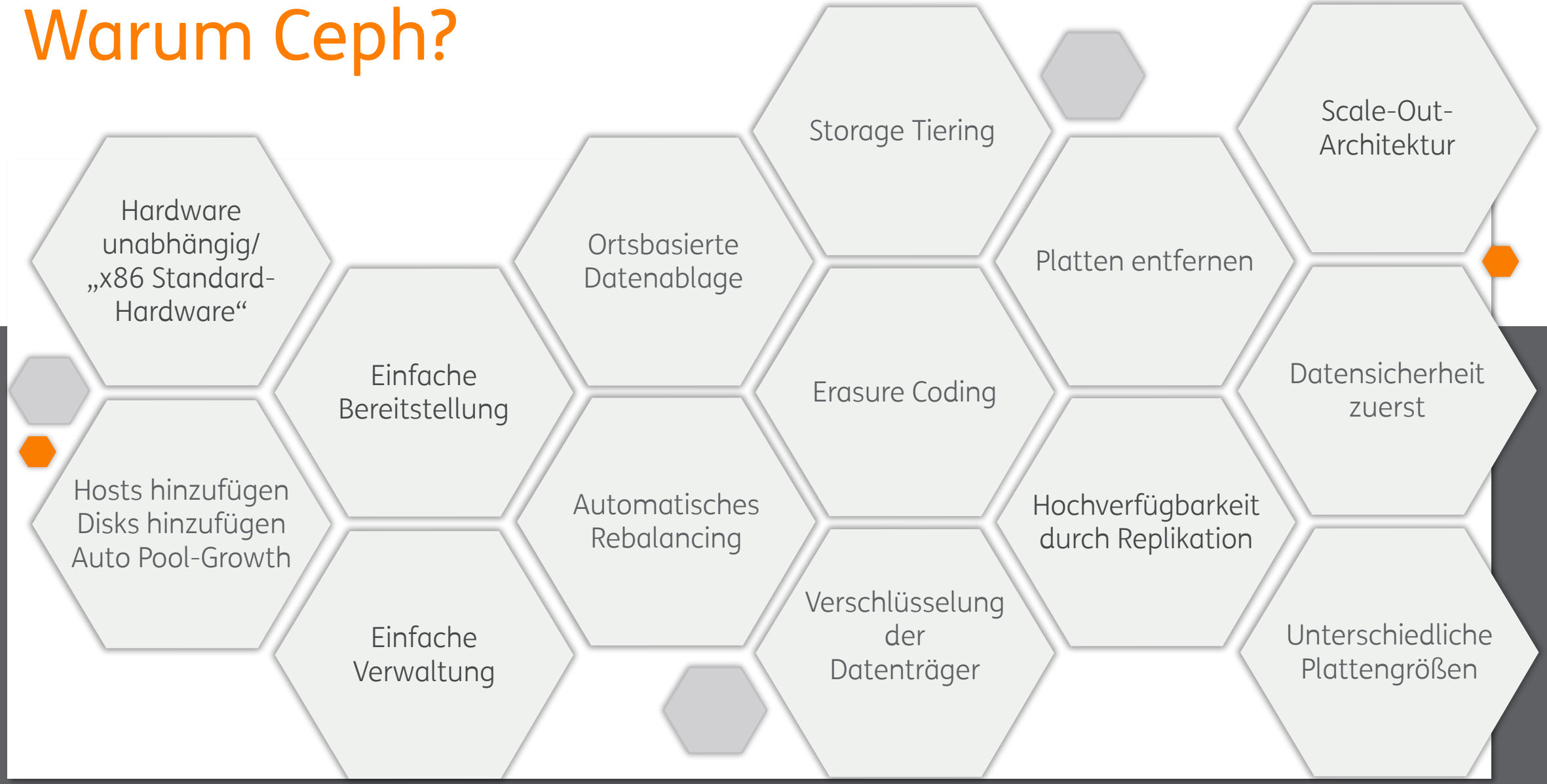
Skalierbar

einfach erweiterbar durch das Hinzufügen von Hosts oder Festplatten

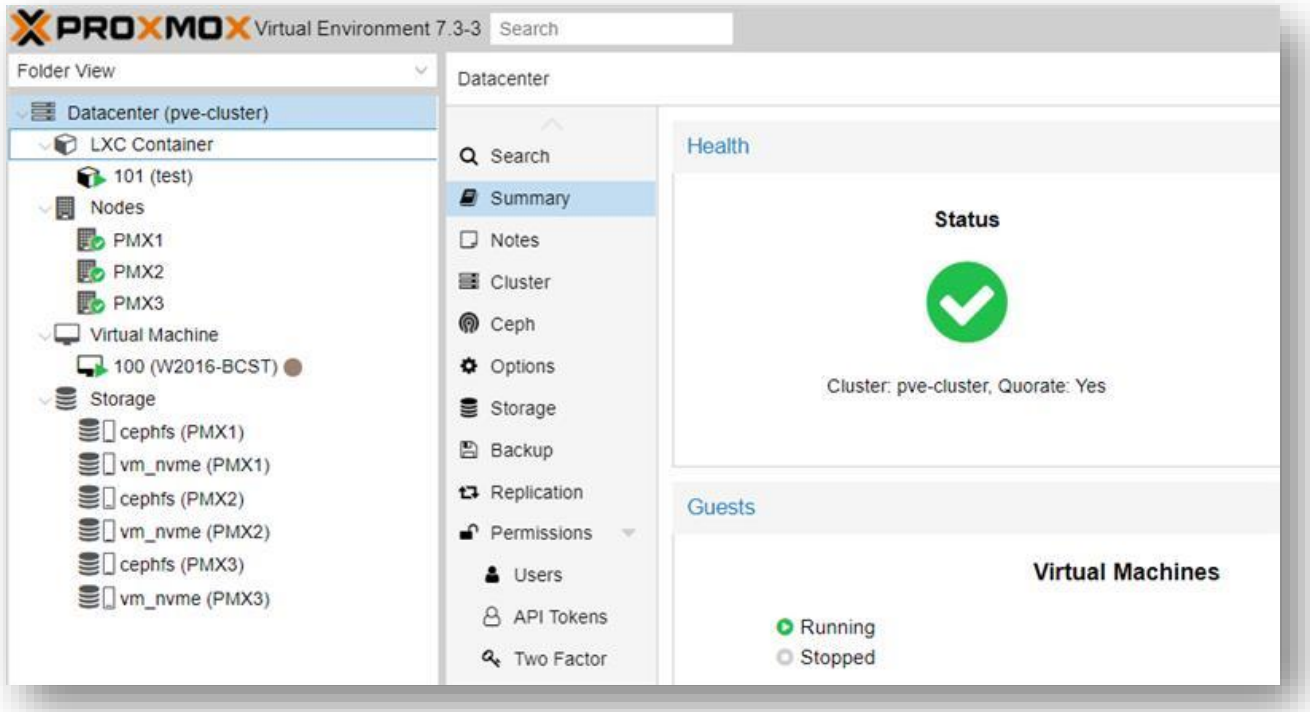
Performant

abhängig vom eingesetzten Netzwerk, CPU und RAM

Warum Ceph?


Hardware unabhängig/ „x86 Standard-Hardware“
Einfache Bereitstellung
Hosts hinzufügen
Disks hinzufügen
Auto Pool-Growth
Einfache Verwaltung
Ortsbasierte Datenablage
Automatisches Rebalancing
Storage Tiering
Erasure Coding
Verschlüsselung der Datenträger
Platten entfernen
Hochverfügbarkeit durch Replikation
Scale-Out-Architektur
Datensicherheit zuerst
Unterschiedliche Plattengrößen

Live Demo


PROXMOX Virtual Environment 7.3-3
Search
Folder View
Datacenter (pve-cluster)
LXC Container
101 (test)
Nodes
PMX1
PMX2
PMX3
Virtual Machine
100 (W2016-BCST)
Storage
cephfs (PMX1)
vm_nvme (PMX1)
cephfs (PMX2)
vm_nvme (PMX2)
cephfs (PMX3)
vm_nvme (PMX3)
Datacenter
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Health
Status
Cluster: pve-cluster, Quorate: Yes
Guests
Virtual Machines
Running
Stopped

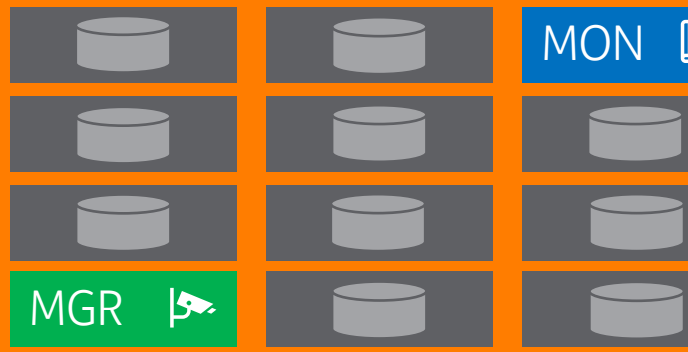
RADOS Gateway

RADOS Block Device

CephFS

LIBRADOS

RADOS



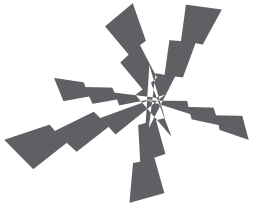
Ceph Monitor

Manager Map

Quorum

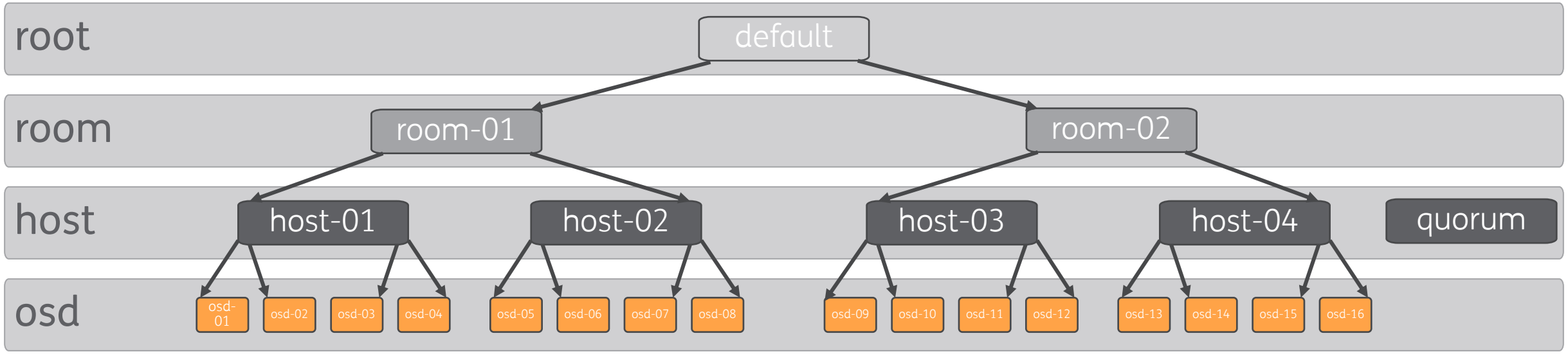
OSD Map

Monitor Map



CRUSH Map

CRUSH Map


root
default
room
room-01
room-02
host
host-01
host-02
host-03
host-04
quorum
osd
osd-01
osd-02
osd-03
osd-04
osd-05
osd-06
osd-07
osd-08
osd-09
osd-10
osd-11
osd-12
osd-13
osd-14
osd-15
osd-16

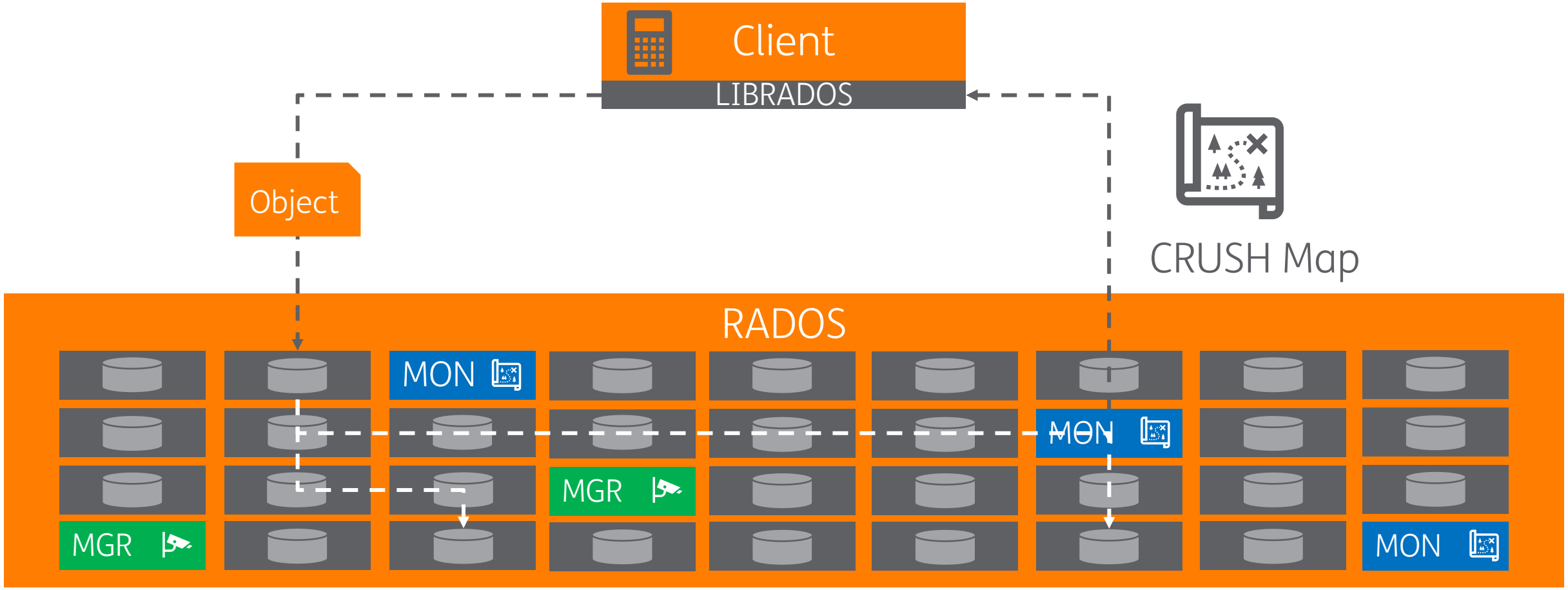
Ceph Manager

- Metriken des Clusters für Dashboard
- Vorhalten von Monitoring-Daten
- Ermöglicht Einbindung von zusätzlichen und externen Monitoring-Tools
- Notwendig zum Betrieb eines Dashboards

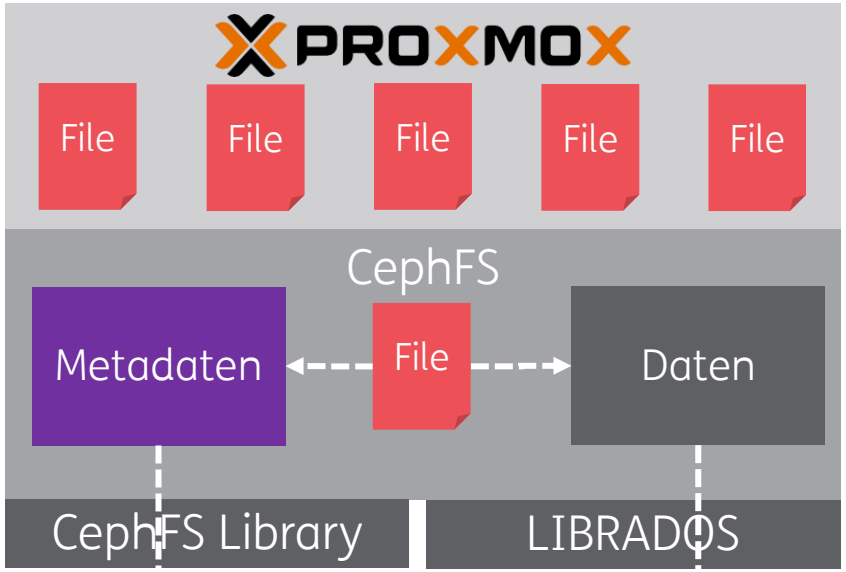
Object Storage Daemon

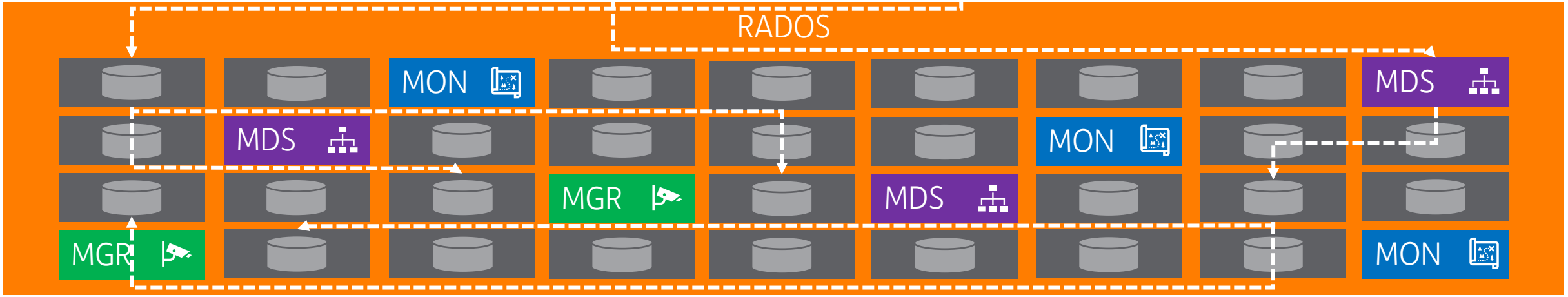
- Dienst zur Abstraktion der physikalischen Datenträger
- Schnittstelle zum Client
- Schreibt Daten und Replikate auf Datenträger
- Eigenschaften: Größe, Device Klasse, Gewichtung, Auslastung, Status

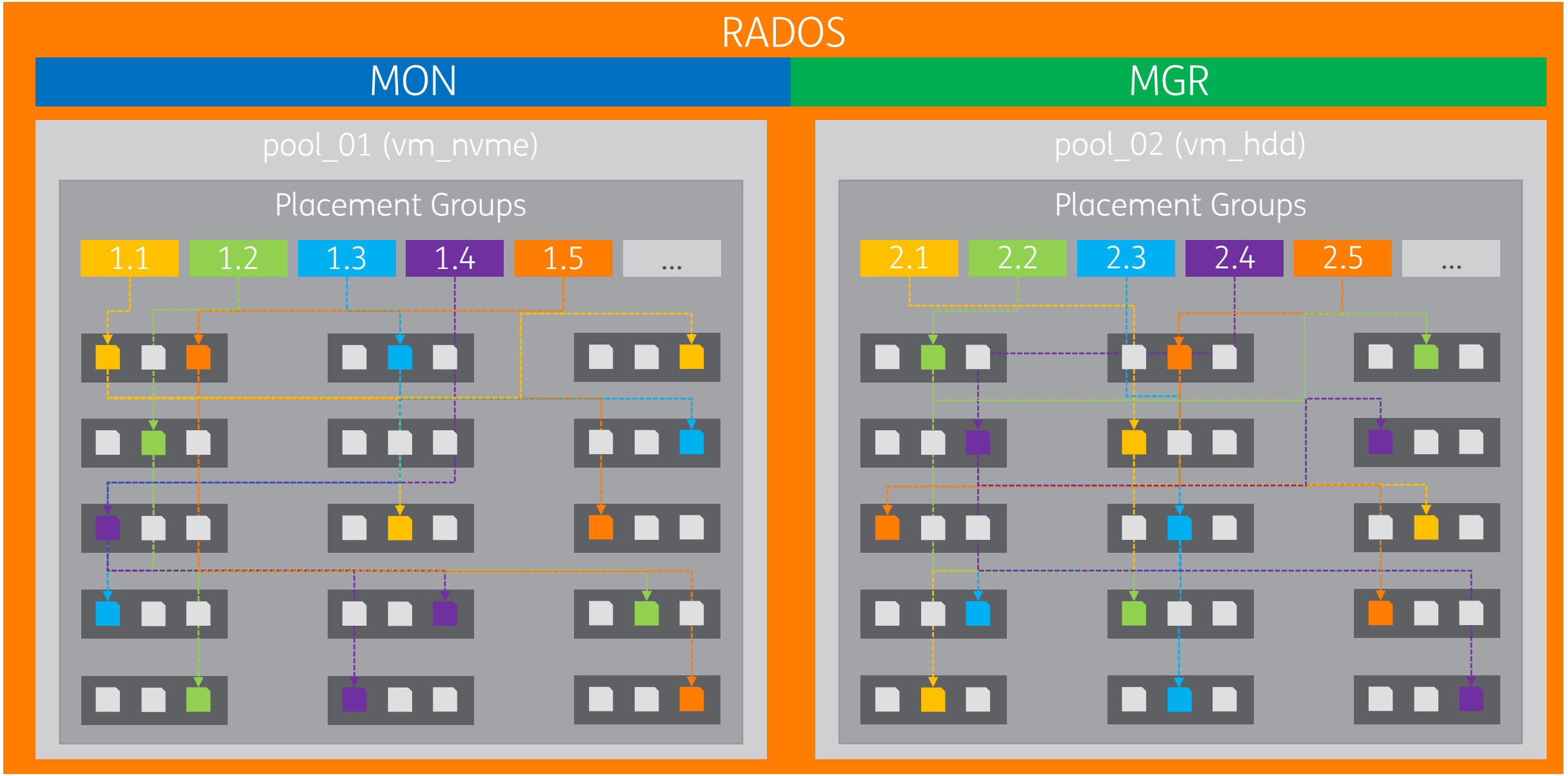
Datenspeicherung



CephFS


PROXMOX
File
File
File
File
File
CephFS
Metadaten
File
Daten
CephFS Library
LIBRADOS


RADOS
MON
MDS
MDS
MON
MGR
MDS
MGR
MON

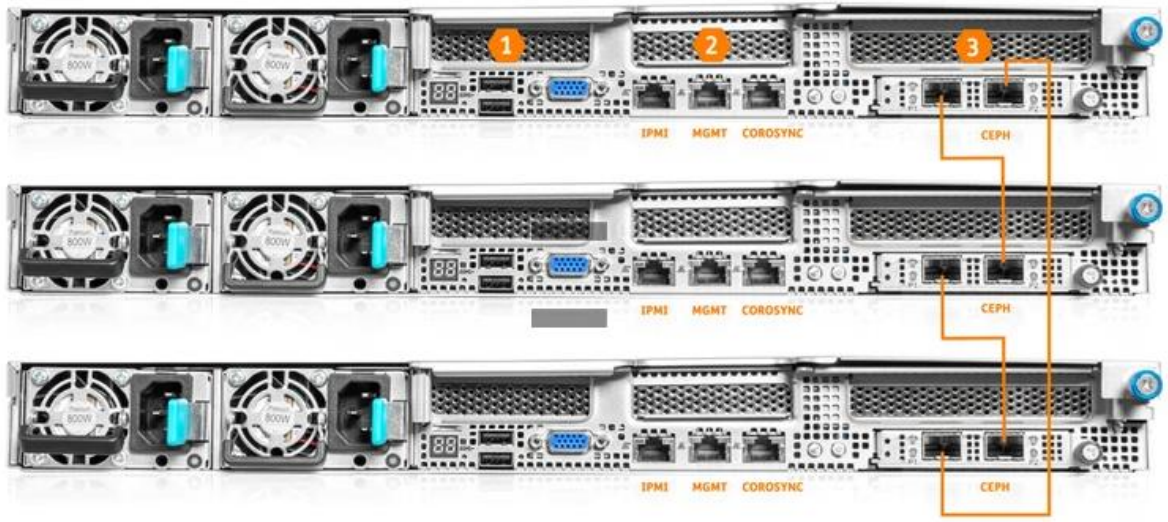

RADOS
MON
MGR
pool_01 (vm_nvme)
Placement Groups
1.1
1.2
1.3
1.4
1.5
...
pool_02 (vm_hdd)
Placement Groups
2.1
2.2
2.3
2.4
2.5
...

Unsere Ceph Deployments

in der Einsteiger-Variante

- 3 Nodes
- keine Storage Switches
- **Proxmox VE + Ceph** (RBD & CephFS)
- Full Meshed Network
  - Bonded RR (manual routing)

    oder
  - RSTP mit open-vSwitch



Unsere Ceph Deployments

im Multi-Room Setup

- MIN: 5 Nodes
- Room 1: 2 Nodes
- Room 2: 2 Nodes
- Room 3: Quorum Node
- SIZE = 4 | MIN-SIZE = 2
- **Proxmox VE + Ceph** (RBD & CephFS)

Benchmarks

- 1 Pool | SIZE 3 | 18 OSDs / Cluster
- Meshed Round Robin (4x 25 Gbit/s)
  - RADOS Bench Write: 8004 MB/s
  - RADOS Bench Read: 23894 MB/s

mit 6x Intel D7-P5520 je Node (3 Nodes)

- 1 Pool | SIZE 4 | 20 OSDs / Cluster
- Switched Setup (2x 25 Gbit/s)
  - RADOS Bench Write: 3279 MB/s
  - RADOS Bench Read: 14971MB/s

mit 5x Intel D7-P5520 je Node (4 Nodes)

Nützliche Tools zur Analyse

- FIO
- IPERF/IPERF3
- RADOS BENCH
- NLOAD
- TMUX
- CEPH COMMANDS

Support Subscriptions

Basic

1 CPU pro Jahr / 2 CPU pro Jahr

- Zugriff auf Enterprise-Repositorys
- Stabile Software-Updates
- Support über das Kundenportal
- Drei Support-Tickets
- Reaktionszeit: 1 Werktag

Standard

1 CPU pro Jahr / 2 CPU pro Jahr

- Alle Basic Leistungen

Zusätzlich:

- 7 Support-Tickets (Gesamt 10)
- Remote support (via SSH)
- Offline Subscription Key Aktivierung

Premium

1 CPU pro Jahr / 2 CPU pro Jahr

- Alle Standard Leistungen

Zusätzlich:

- Unbegrenzte Support-Tickets

Voller Funktionsumfang auch ohne Subscriptions

Proxmox VE & Ceph bei der Thomas-Krenn.AG

- Wie können wir unterstützen?
  - Exzellente Beratung, Installation und Einweisung
  - Singlehost bis Multi-Room
- Welche Expertise haben wir?
  - Proxmox Gold & Solution Partner
- Was geben wir an die Community?
  - Forum Aktivität (Support)
  - Online-Webinare Proxmox VE & Ceph
  - Bugreports & Feature Requests (Bugzilla)
  - Wiki-Artikel & Know How im Thomas Krenn Wiki

Proxmox VE Appliances & Server

- Aktuelle Systeme
- All NVMe v3 (Genoa)
- Proxmox Performance Paper All NVMe (Intel)
- Individual-Systeme
- LES, Tower, 1-4 HE
- Industrie-Server

Weiterführende Links

- Proxmox Admin Guide: https://pve.proxmox.com/pve-docs/pve-admin-guide.html
  - Einrichtung und Verwaltung von Proxmox
  - Dokumentation
- Proxmox Wiki: https://pve.proxmox.com/wiki/Main_Page
  - Anleitungen
- Forum: https://forum.proxmox.com/
  - Community Austausch
- Bugzilla: https://bugzilla.proxmox.com/
  - Bug-Reports
  - Feature Requests
- Thomas-Krenn-Wiki: https://www.thomas-krenn.com/de/wiki/Kategorie:Proxmox
  - Anleitungen
  - Troubleshooting

THOMAS KRENN®

jsterr@thomas-krenn.com

Jonas Sterr

THOMAS KRENN®

Vielen Dank für Ihre Aufmerksamkeit!