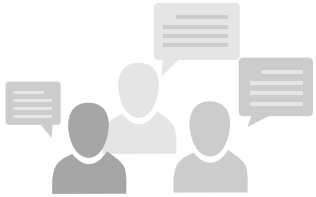
STORAGE REPLICATION MIT ZFS IM DATACENTER

Wolfgang Link, Software Development,
Proxmox Server Solutions GmbH
SLAC | 27-29 Mai 2019 | Berlin

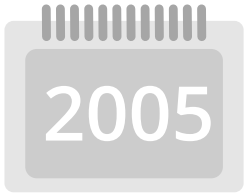
PROXMOX

Herzlich Willkommen!

Proxmox Server Solutions GmbH



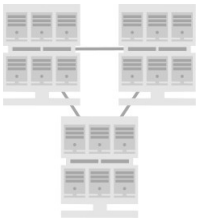
Aktive Community



Proxmox seit 2005 in Wien

Globales Partner-Netz

Proxmox Mail Gateway (AGPL,v3)

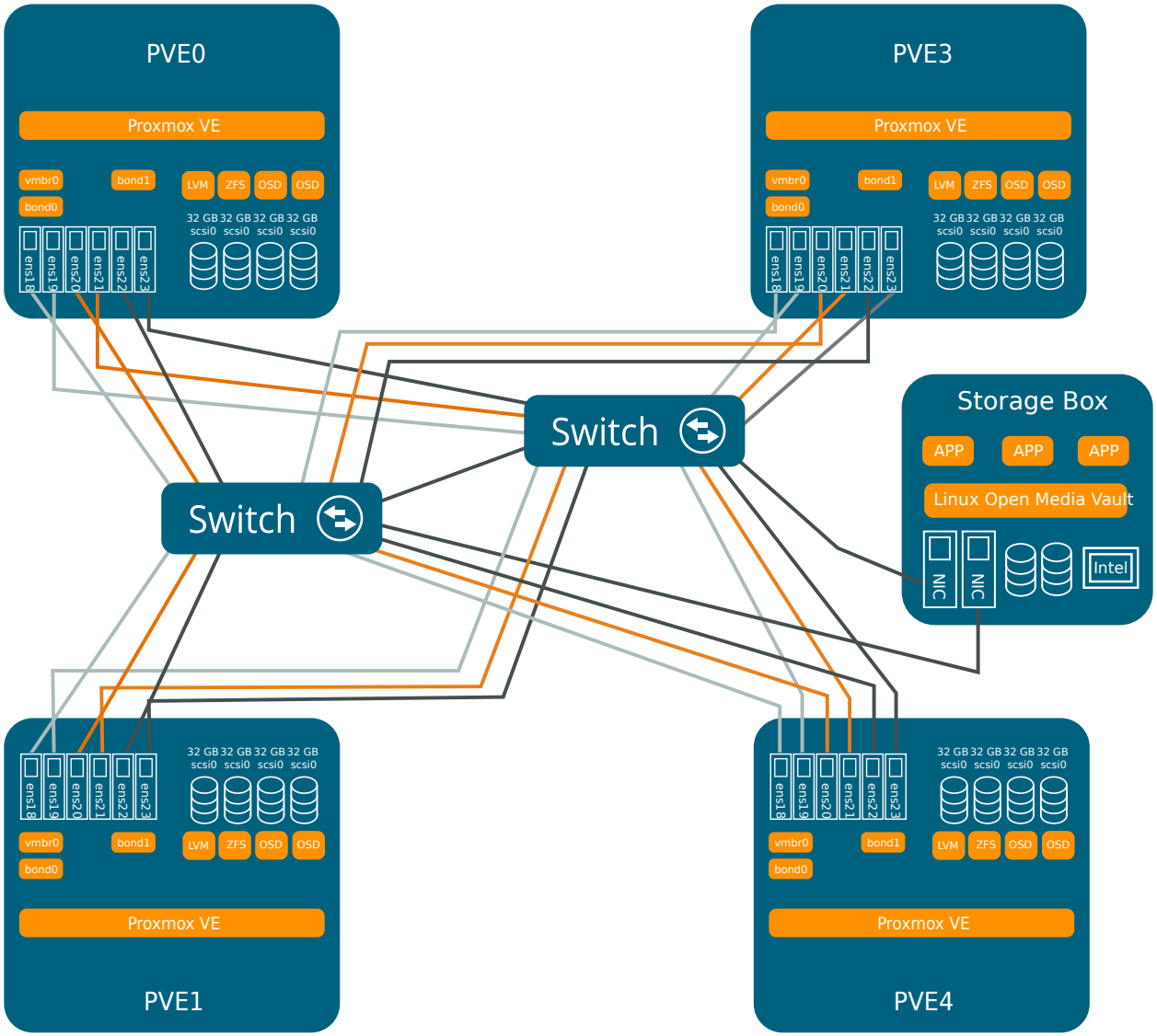


Proxmox Virtual Environment (AGPL,v3)

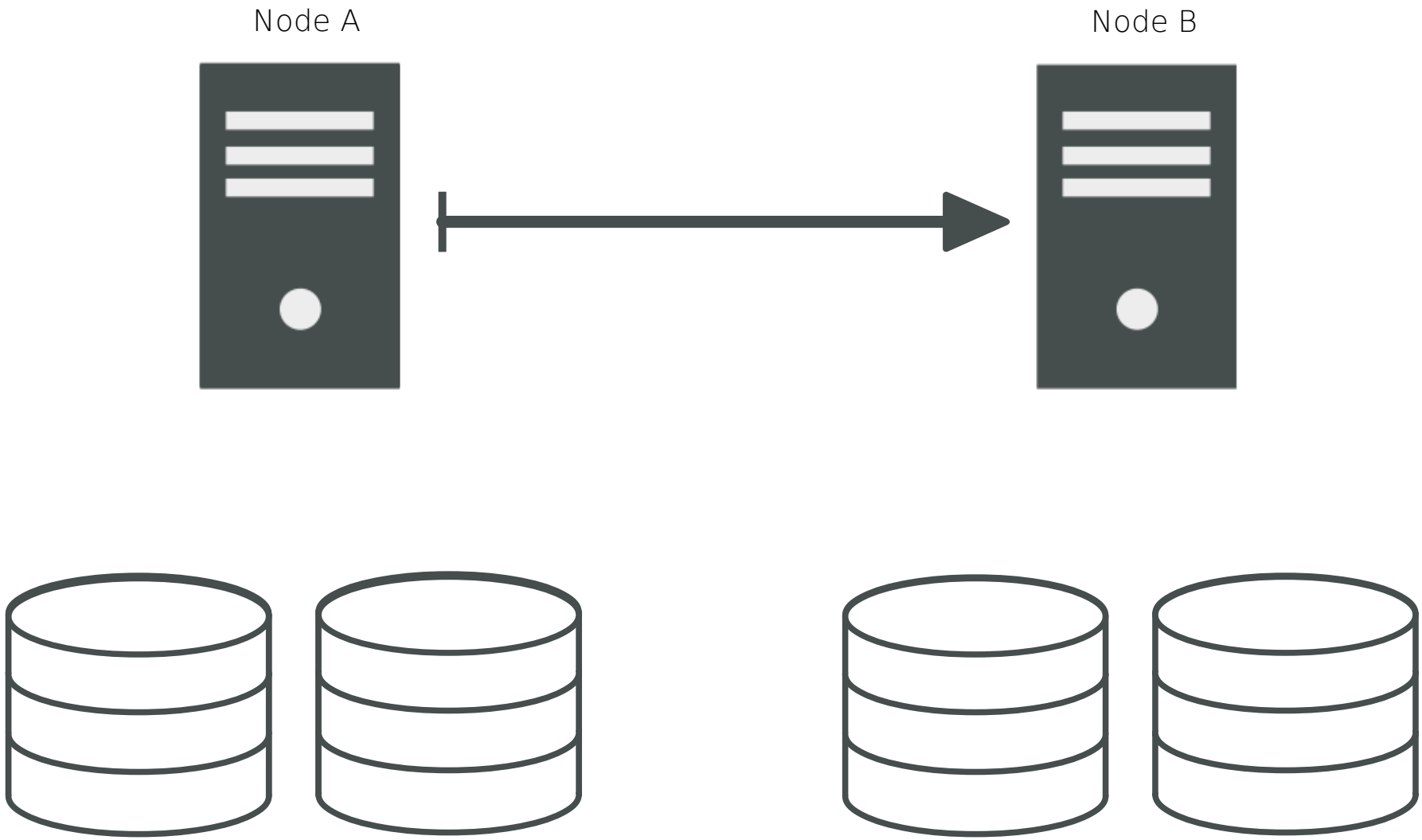
Enterprise Support & Services

PROXMOX

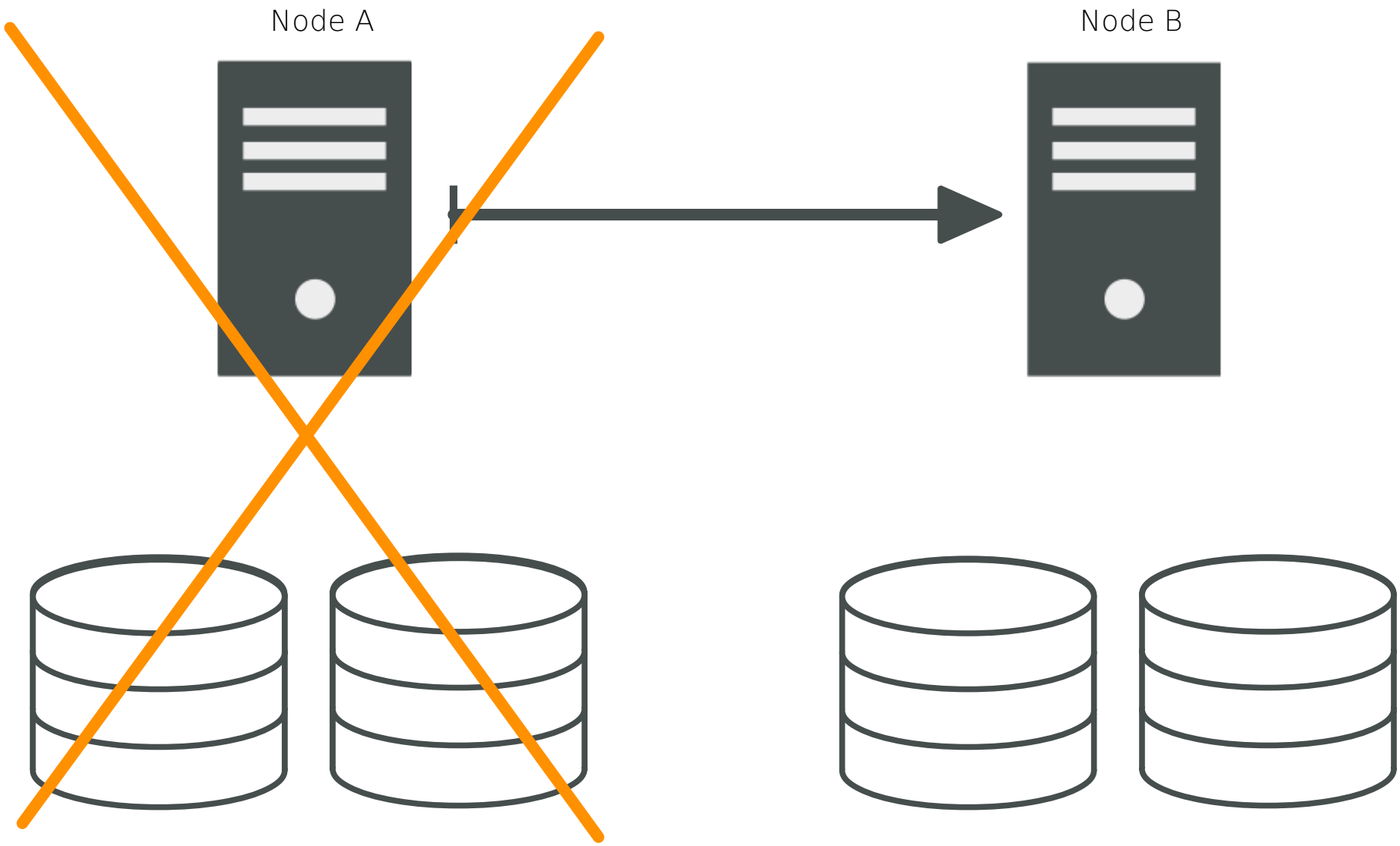
Herkömmlicher Virtualisierungs-Cluster



Was ist Storage Replication?



Warum Storage Replication?



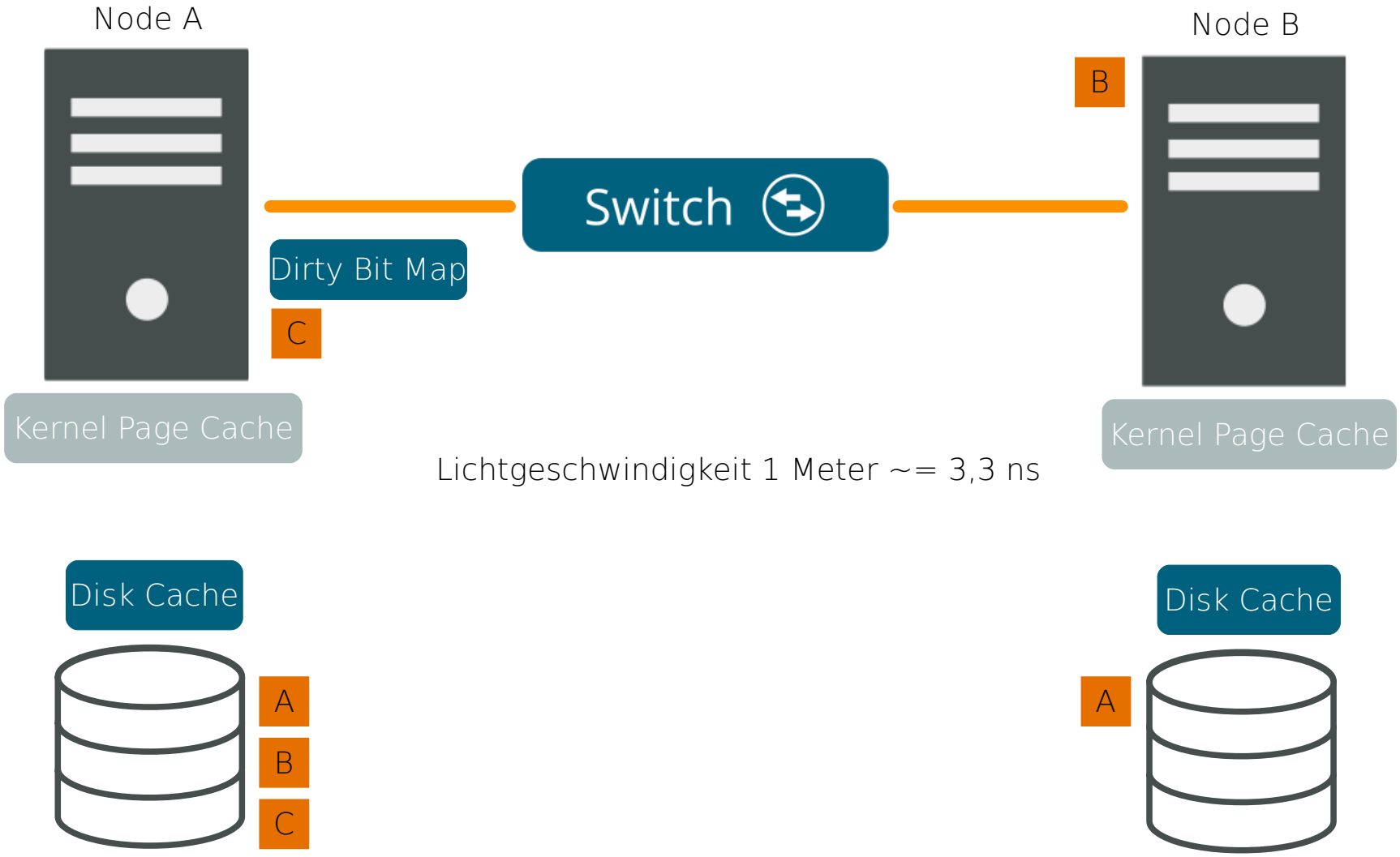
Ist synchron immer synchron?

Storage Replication-Technologien verstehen unter synchron nicht immer das selbe.

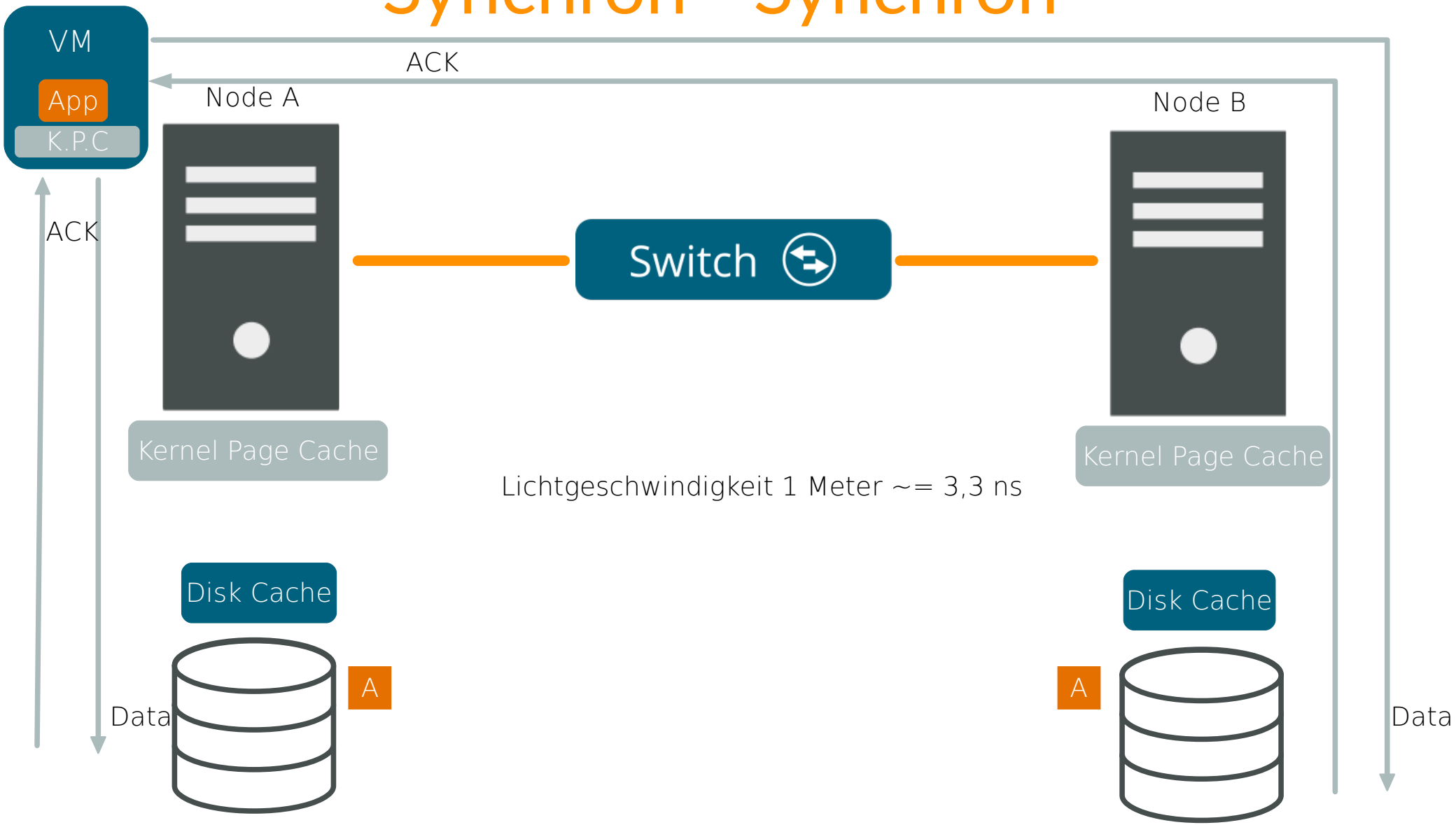
Drei verschiedene Klassen:

- Synchron
- Semi-Synchron
- Asynchron-Synchron

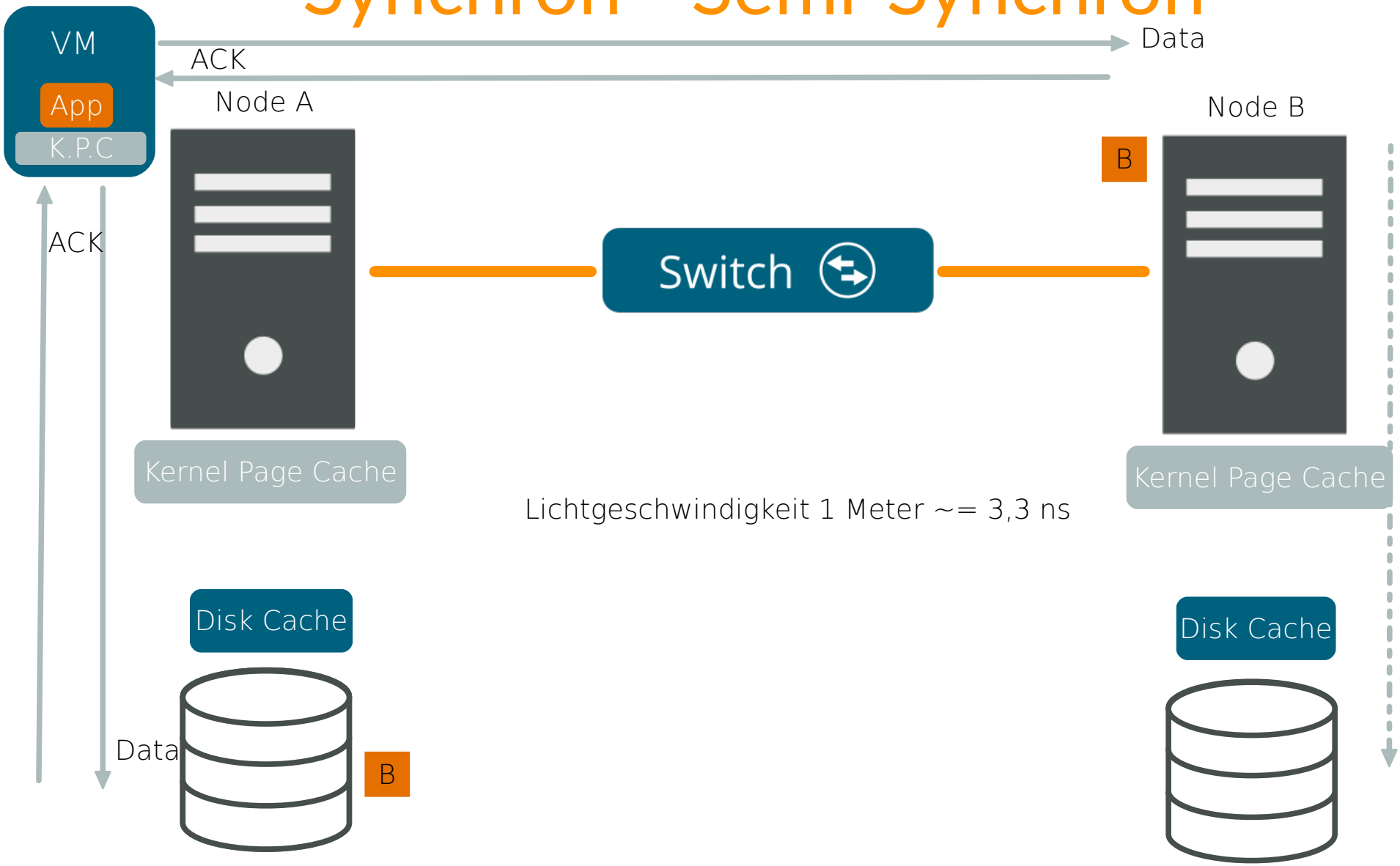
Wann sind Daten geschrieben?



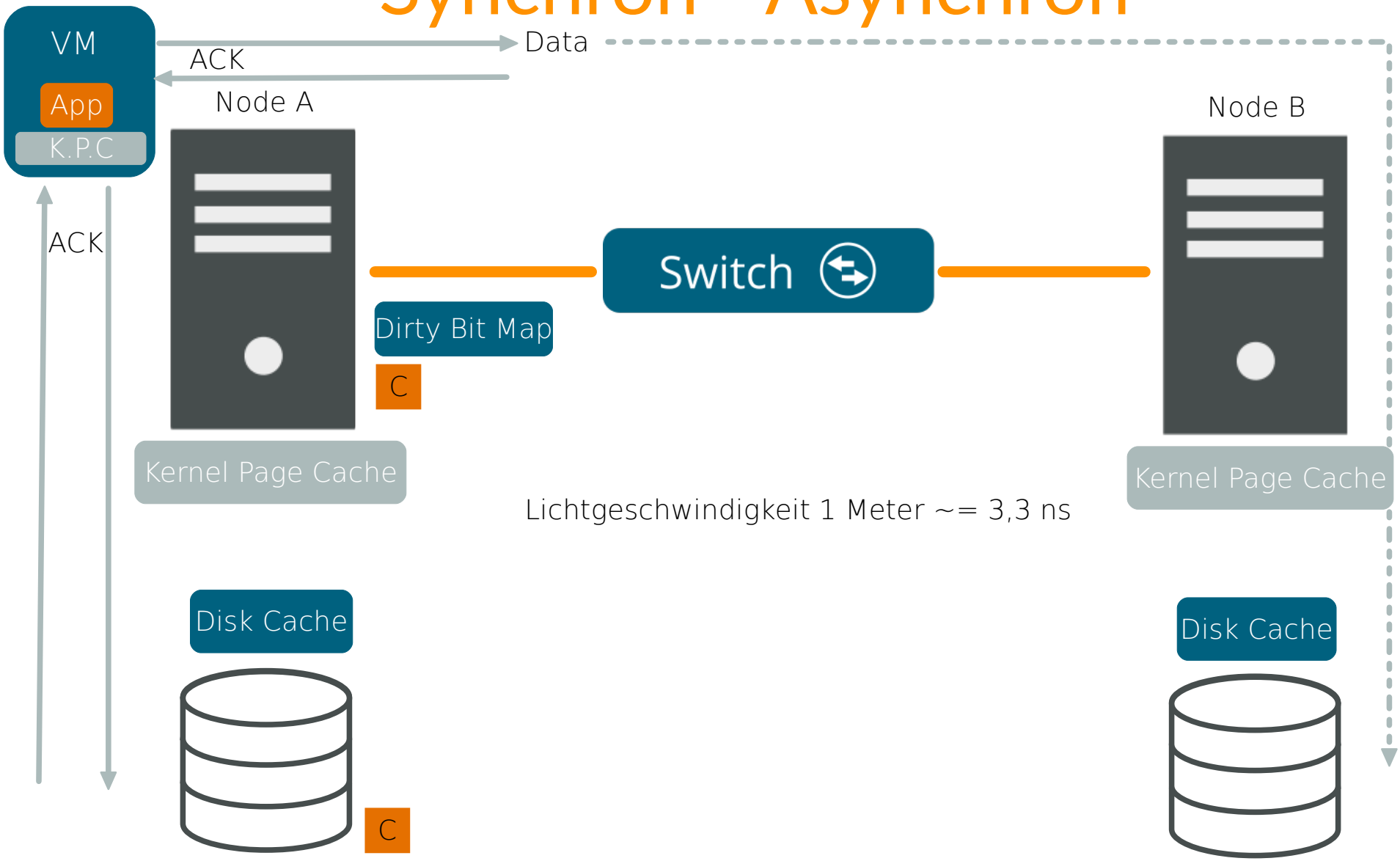
Synchron - Synchron



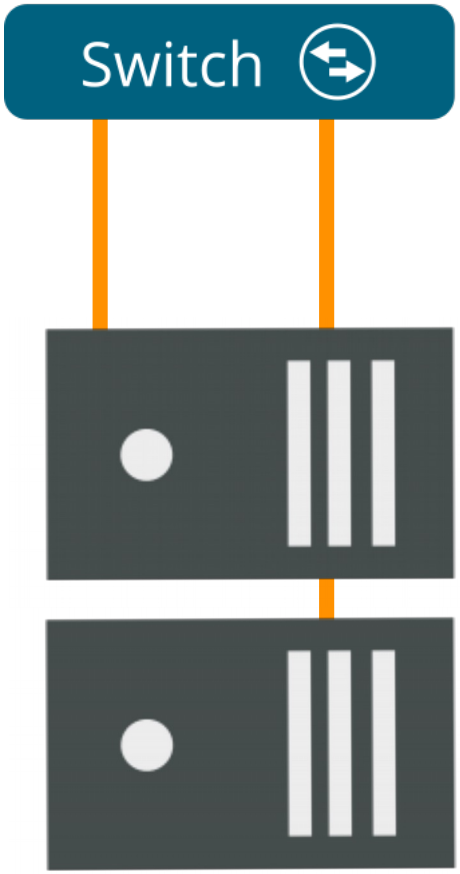
Synchron - Semi-Synchron



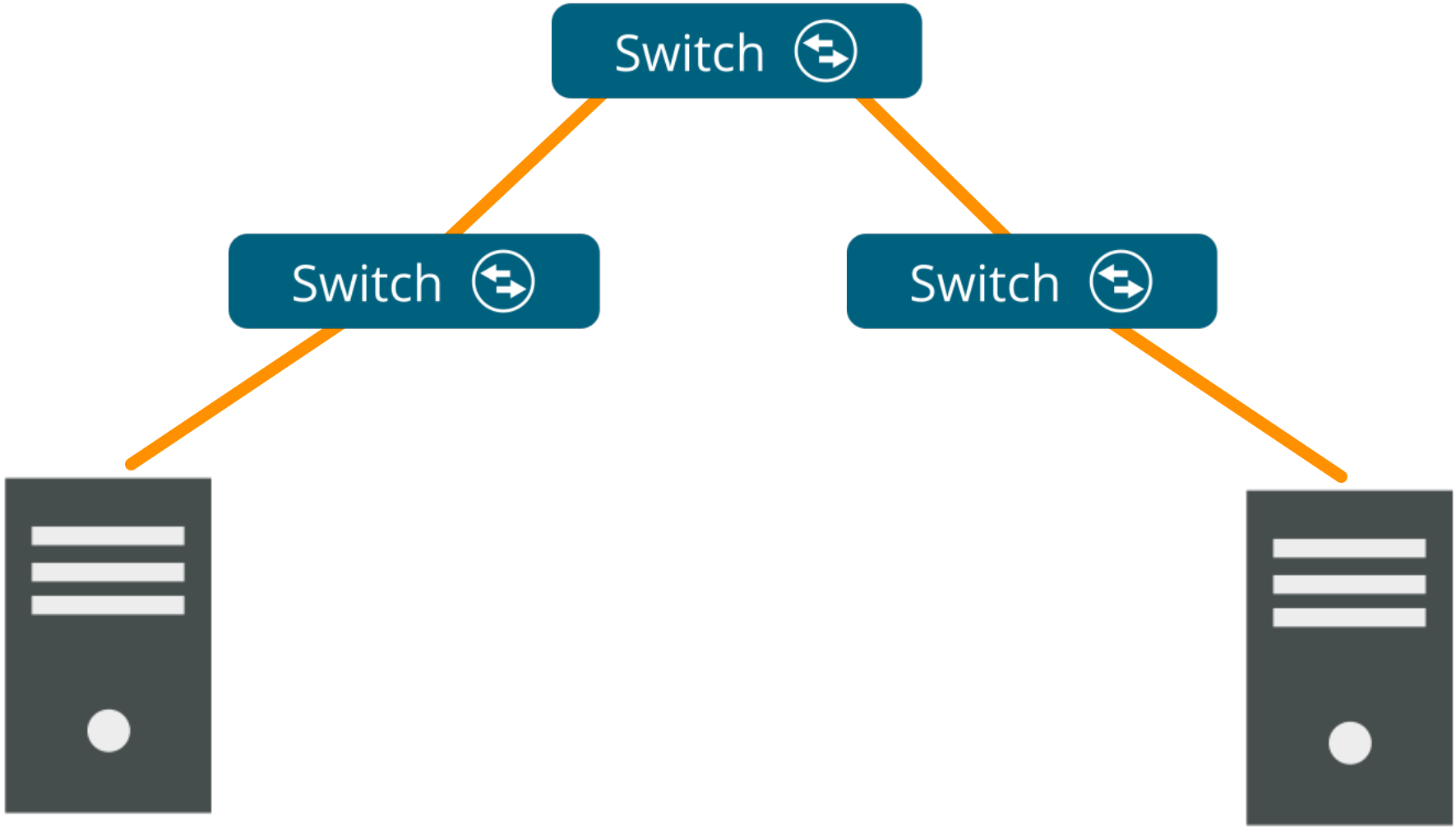
Synchron - Asynchron



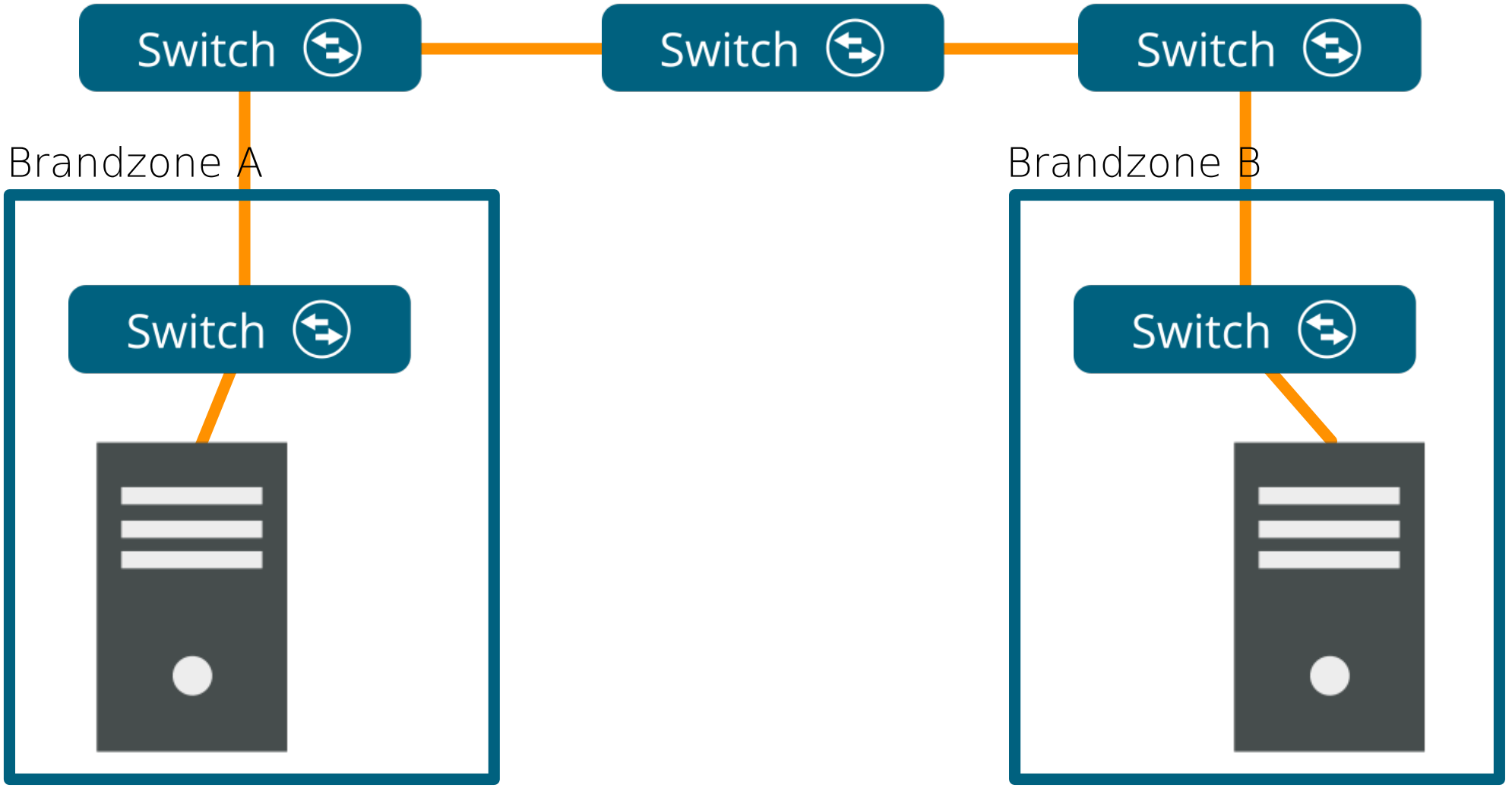
Datacenter Single Rack



Datacenter Multiple Rack



Datacenter Multiple Zonen



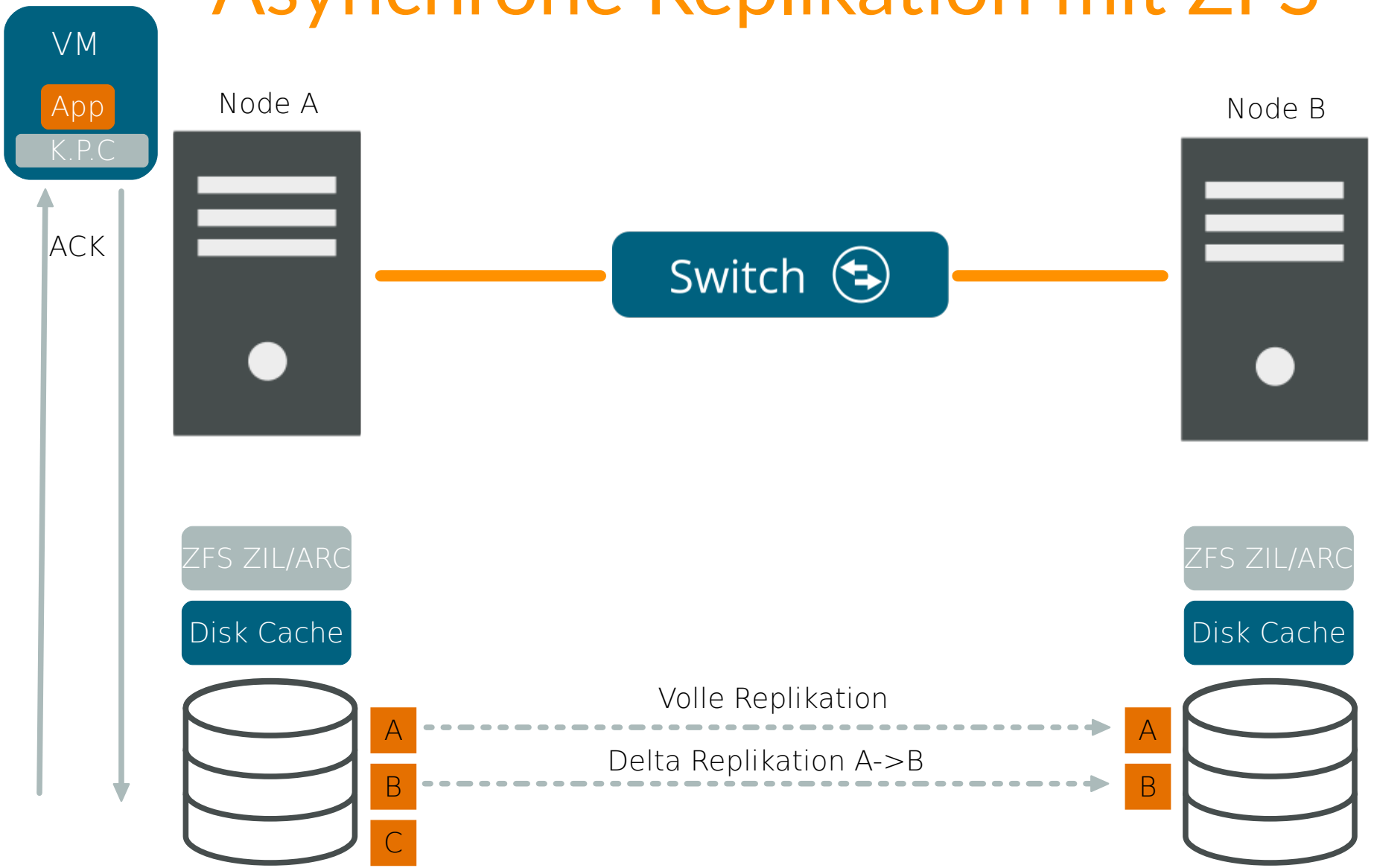
Warum ZFS?

- ZFS vereint:
  - Software RAID
  - Logical Volume Manager
  - Filesystem
- ZFS bietet *Bit Rot Protektion*
- ZFS hat *write Cache*
- ZFS ist transaktional
- Seit ZFS 0.8.0 hat es native Verschlüsselung

Asynchrone Replikation mit ZFS

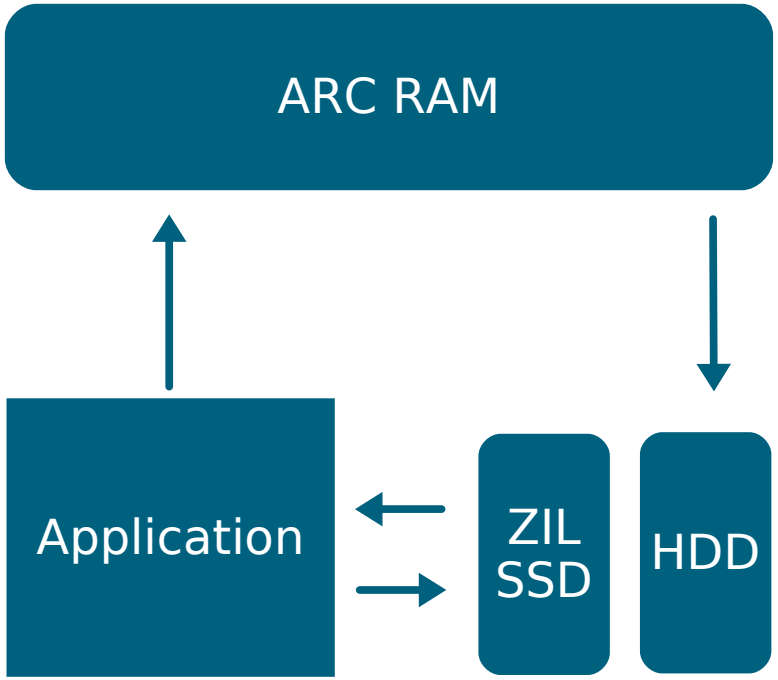
- Daten sind immer konsistent.
- Daten werden beim Senden komprimiert.
- Daten sind durch *checksumming* beim Transport zusätzlich geschützt.
- Abgebrochene Streams können wieder aufgenommen werden.
- Sehr geringe Netzwerk-Nutzung, da nur das Delta gesendet wird.

Asynchrone Replikation mit ZFS

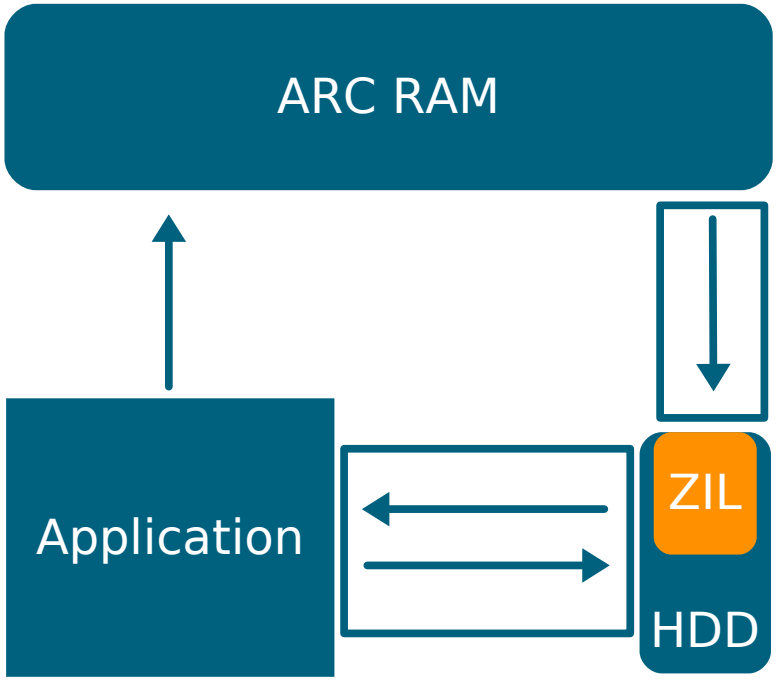


ZFS ARC, L2ARC and ZIL

With ZIL

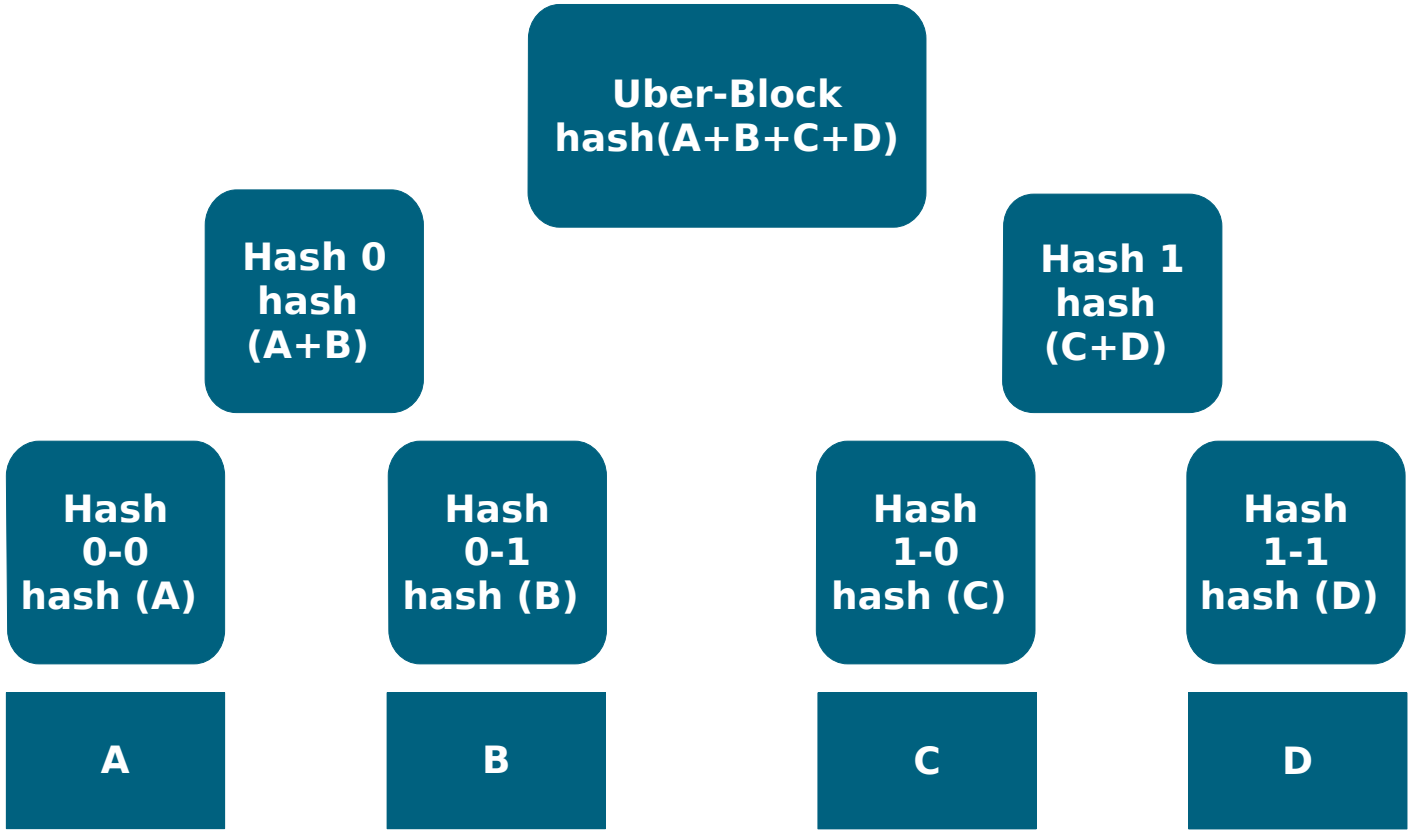


Without ZIL

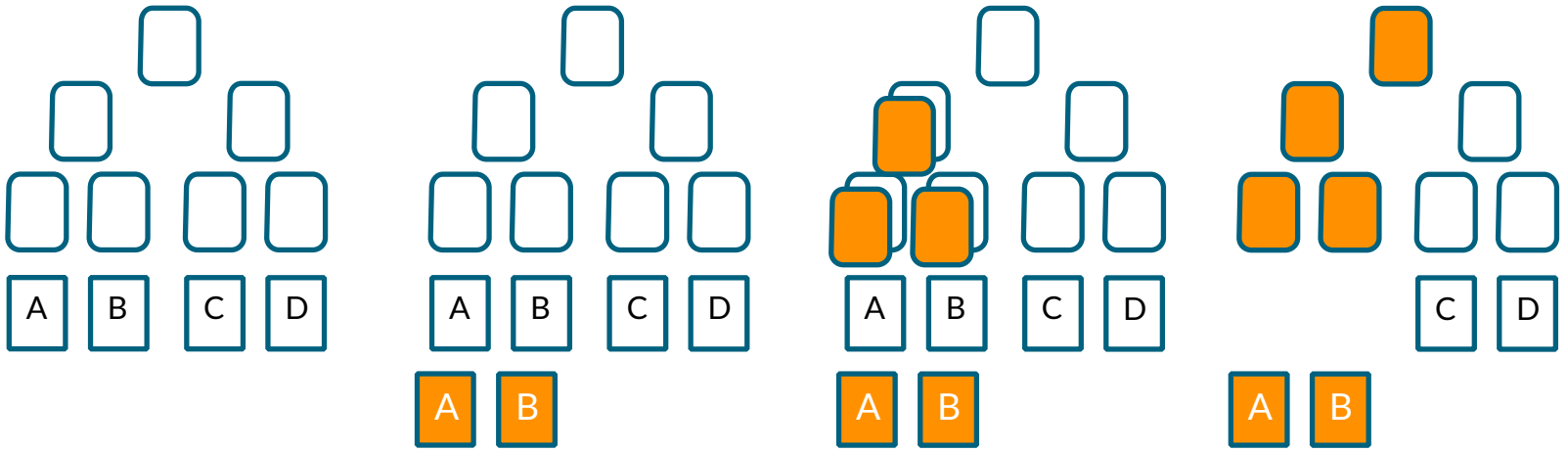


ZFS Tree

- Markle-Tree (Hash-Tree)



Copy on write on ZFS



Proxmox VE mit integriertem Storage Replication-Framework

- Komplettes Handling der Snapshots und Konfiguration.
- Über das Webinterface (GUI) konfigurierbar.
- Automatische Benachrichtigung im Fehlerfall.
- HA der VM/CT mit Replikation möglich.
- Komplette Integration in das Proxmox VE Framework.

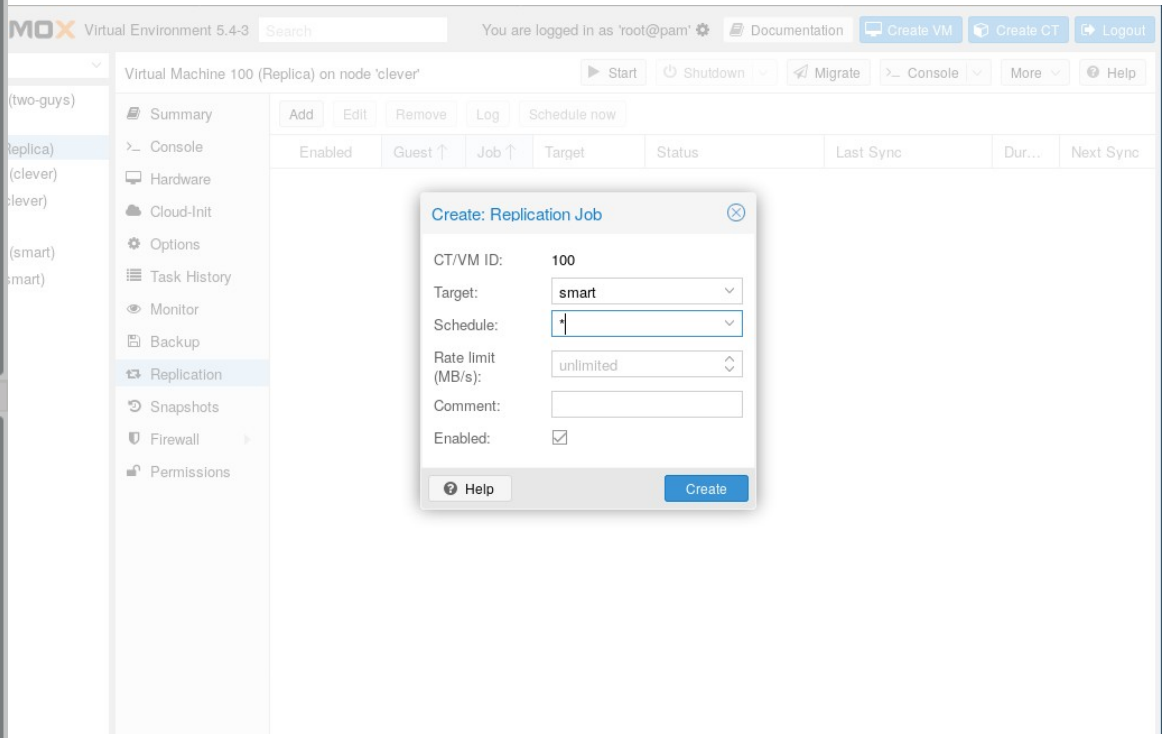
ZFS Replikations-Struktur

wlink@ella:/home/wlink 60x23

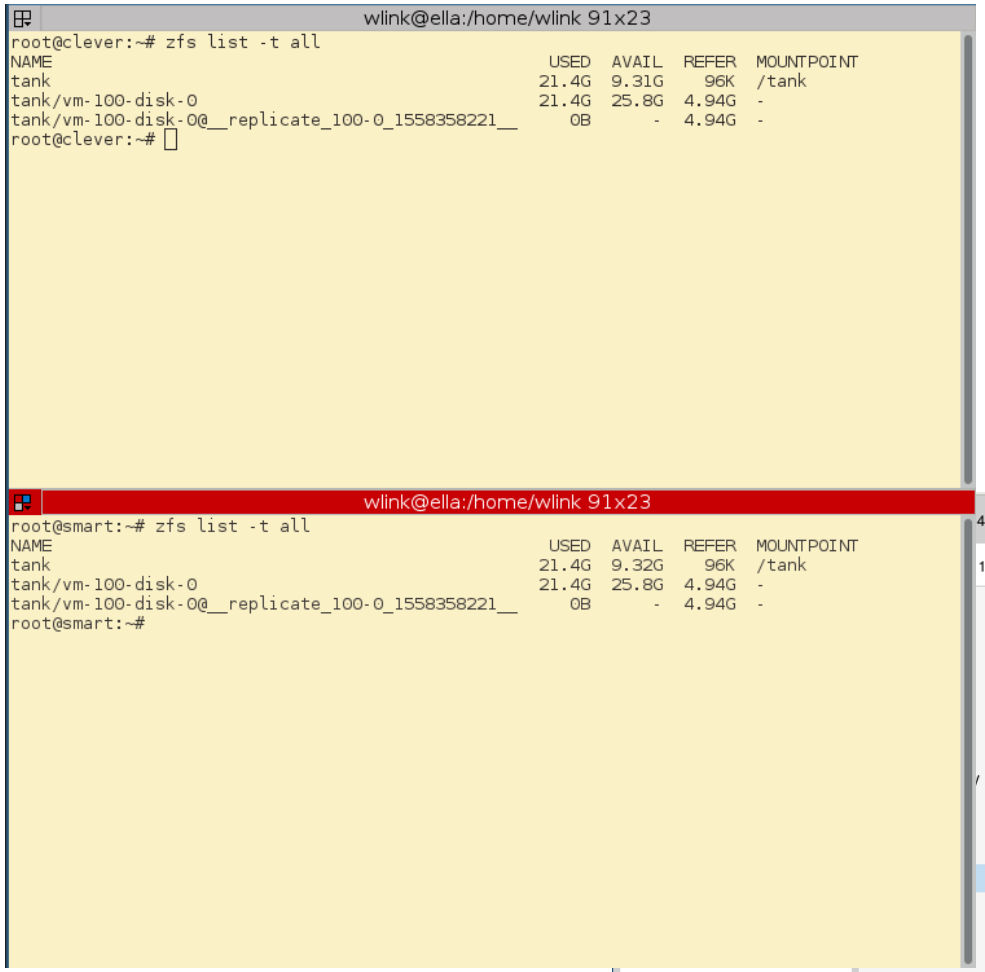
```
root@clever:~# zfs list -t all
NAME                  USED  AVAIL  REFER  MOUNTPOINT
tank                 21.4G  9.31G    96K  /tank
tank/vm-100-disk-0   21.4G  25.8G  4.94G  -
root@clever:~#
```

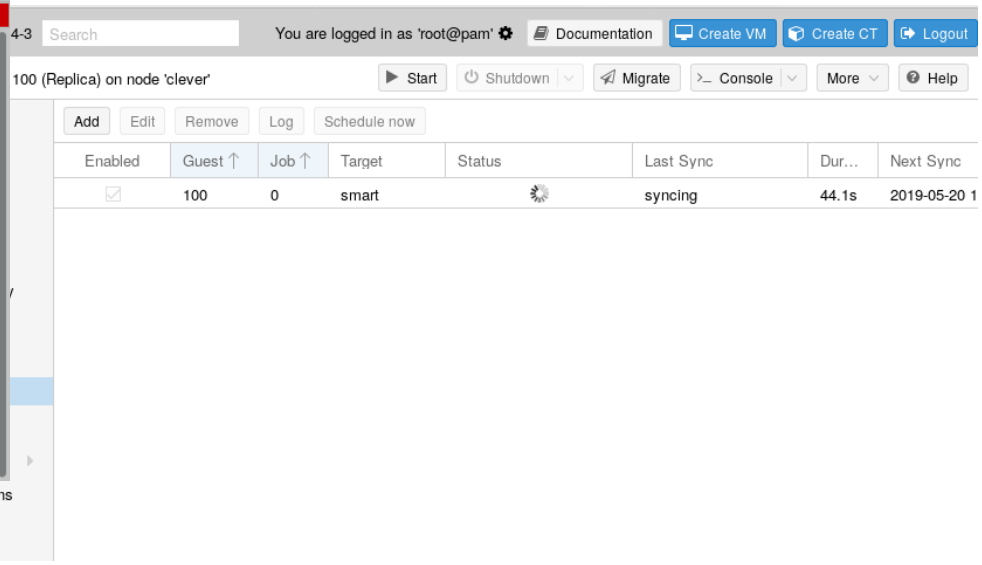
wlink@ella:/home/wlink 60x23

```
root@smart:~# zfs list
NAME   USED  AVAIL  REFER  MOUNTPOINT
tank  4.23M  30.8G    96K  /tank
root@smart:~#
```



ZFS Replikations-Struktur





ZFS Replikations-Struktur

wlink@ella:/home/wlink 91x23

```
root@clever:~# zfs list -t all
NAME                                                  USED  AVAIL  REFER  MOUNTPOINT
tank                                                 21.4G  9.31G    96K  /tank
tank/vm-100-disk-0                                   21.4G  24.6G  4.94G  -
tank/vm-100-disk-0@__replicate_100-0_1558358341__    1.18G      -  4.94G  -
root@clever:~#
```

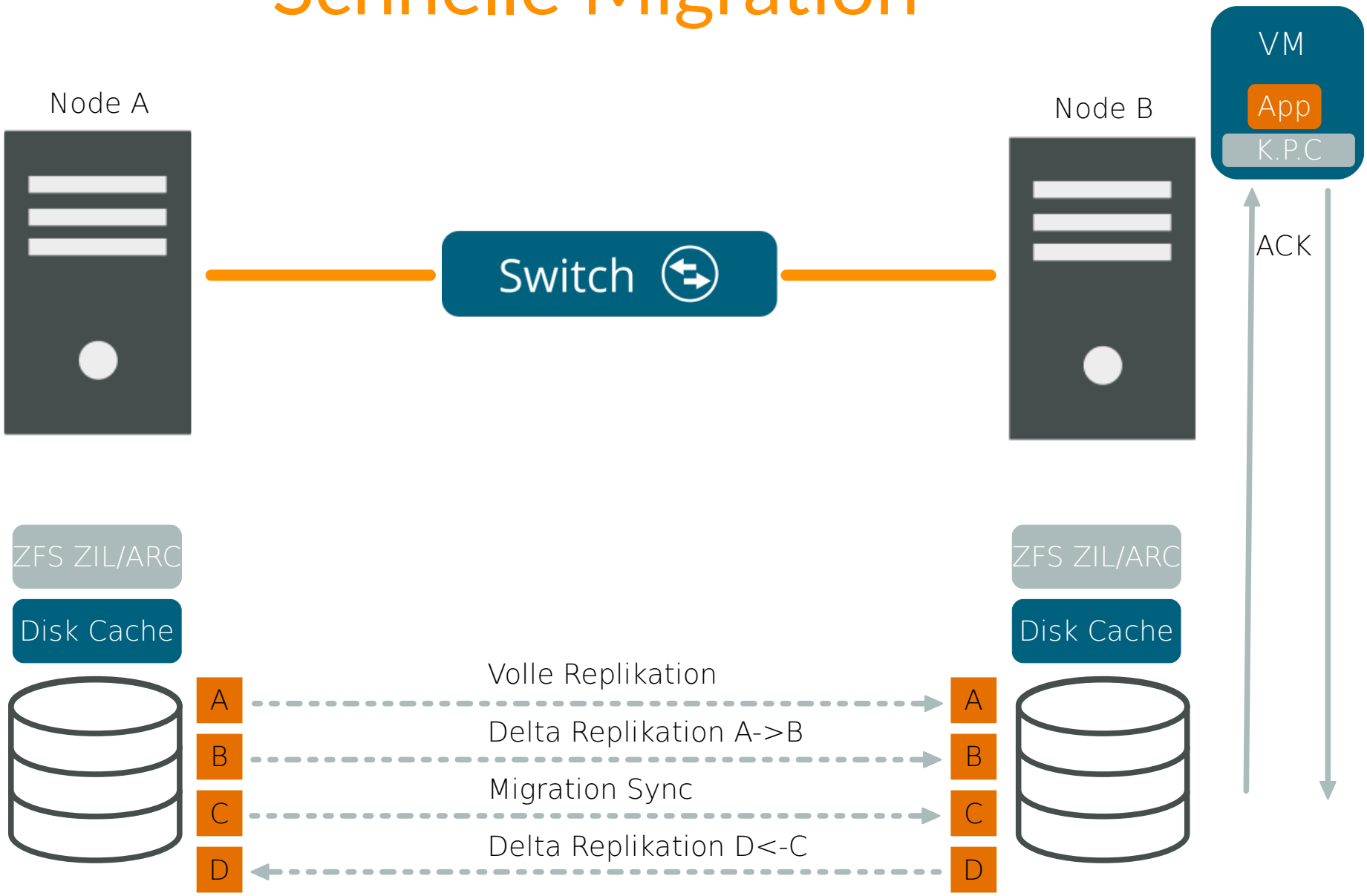
wlink@ella:/home/wlink 91x23

```
root@smart:~# zfs list -t all
NAME                                                  USED  AVAIL  REFER  MOUNTPOINT
tank                                                 21.4G  9.32G    96K  /tank
tank/vm-100-disk-0                                   21.4G  25.8G  4.94G  -
tank/vm-100-disk-0@__replicate_100-0_1558358341__       0B      -  4.94G  -
root@smart:~#
```

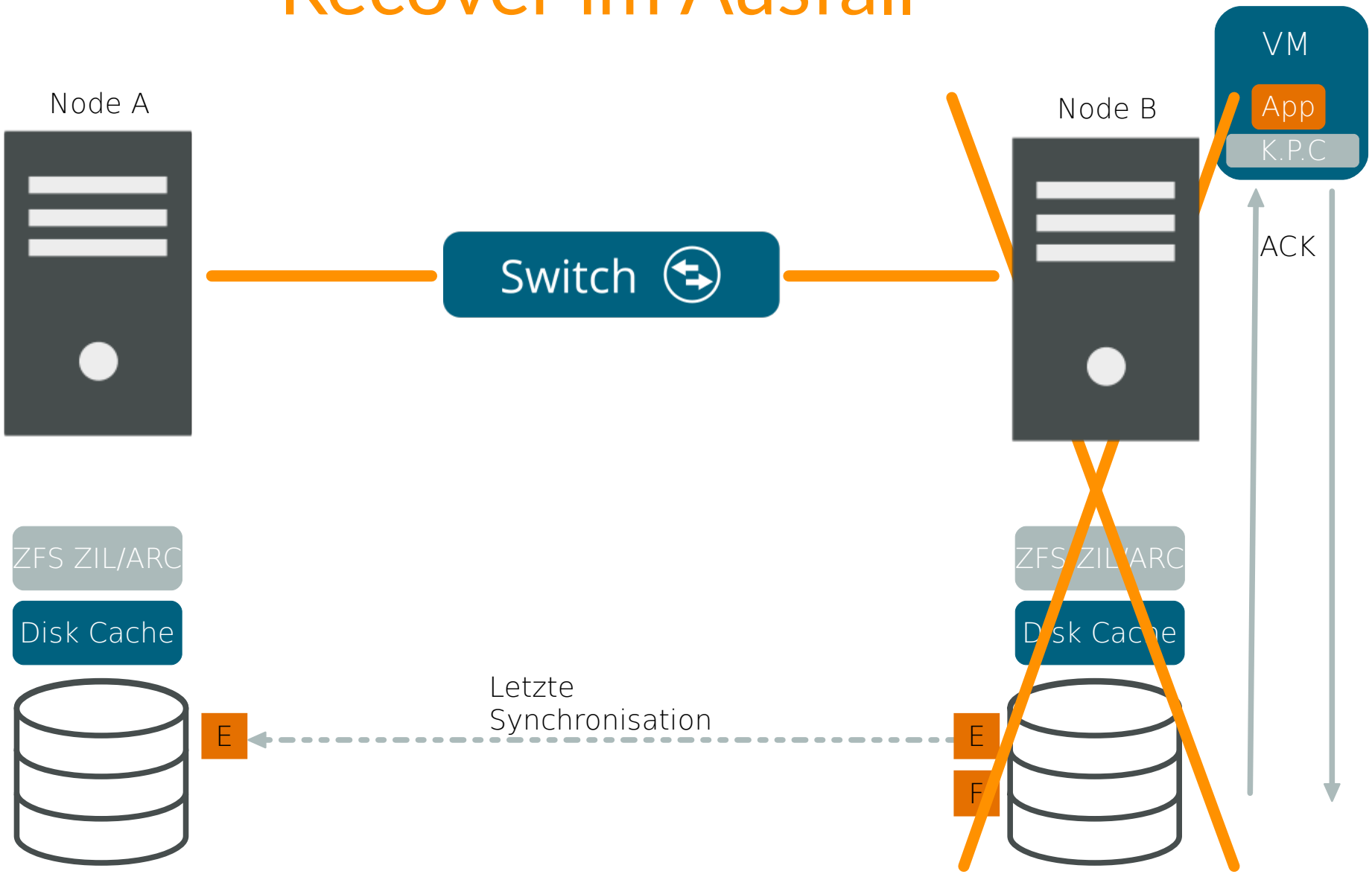
Transport mit ssh

- Sehr sicher.
- Hardware beschleunigt.
- Gut maintained durch OpenSSL.
- Standard -> Jeder kennt sich aus.

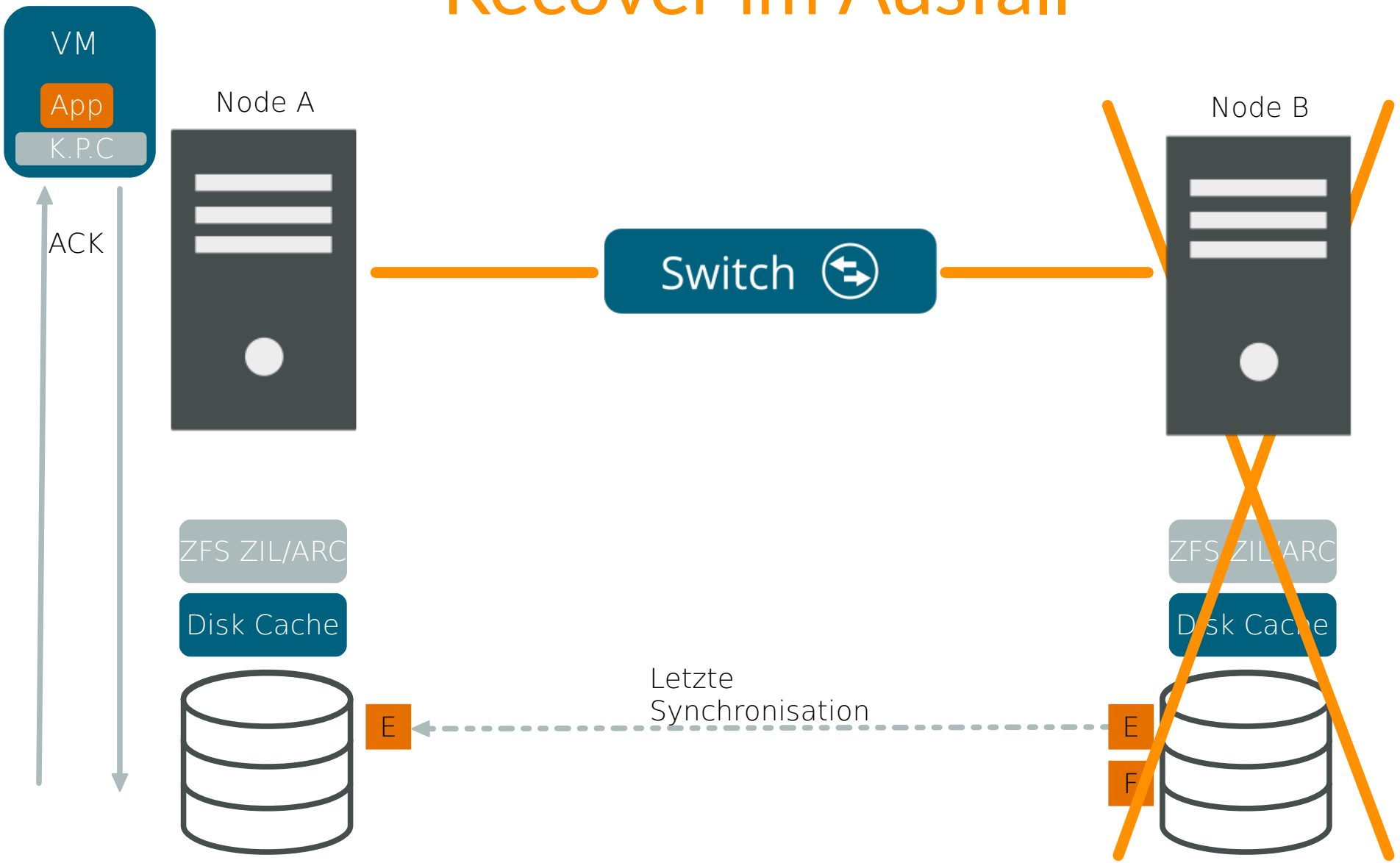
Schnelle Migration



Recover im Ausfall



Recover im Ausfall



Synchrone oder Asynchrone Replikation?

- Synchrone Replikation
  - Pro: Daten sind immer auf dem gleichen Stand.
  - Contra: sehr Ressourcen intensiv. Endet meist in langsamen Systemen.
  - Contra: hohe Lernkurve, da synchrone Speicher sehr komplex werden.

Synchrone oder Asynchrone Replikation?

- Asynchrone Replikation
  - Pro: verbraucht weniger Ressourcen.
  - Pro: Einfache Konfiguration, da die Architektur einfacher ist.
  - Contra: Verlust von nicht synchronisierten Daten im Desaster-Fall.

Links und weitere Infos

- https://www.proxmox.com
- https://pve.proxmox.com/wiki/\Storage_Replication
- https://zfsonlinux.org/
- https://pve.proxmox.com/wiki/PVE-zsync

DANKE FÜR IHRE AUFMERKSAMKEIT!

Proxmox Server Solutions GmbH
Bräuhausgasse 37
1050 Vienna
Austria

office@proxmox.com
www.proxmox.com

© 2019 Proxmox. All Rights Reserved. All brand names, product names, or trademarks belong to their respective holders.