

Der unendlich große Fileserver mit Ceph und Samba

Motivation

- „Unstrukturierte Daten“
- Arbeitsplatzrechner, Laptops, Laborgeräte, Fertigungsmaschinen
- Archivsystem
- Niemand wird je wieder eine Datei löschen
- Regulatorische Anforderungen / Compliance Daten für 10 Jahre aufheben
- Budget

- Speicherkapazität, Wachstum und Zuverlässigkeit wichtiger als Geschwindigkeit

→ **Heinlein Support**

- IT-Consulting und 24/7 Linux-Support mit ~28 Mitarbeitern
- Eigener Betrieb eines ISPs seit 1992
- Täglich tiefe Einblicke in die Herzen der IT aller Unternehmensgrößen

→ **24/7-Notfall-Hotline: 030 / 40 50 5 - 110**

- 28 Spezialisten mit LPIC-2 und LPIC-3
- Für alles rund um Linux & Server & DMZ
- Akutes: Downtimes, Performanceprobleme, Hackereinbrüche, Datenverlust
- Strategisches: Revision, Planung, Beratung, Konfigurationshilfe

→ **Robert Sander**

- Senior Linux Consultant
 - Ceph
 - CheckMK
 - Samba
 - LDAP
 - Netzwerk, Routing, Firewall
- 2000 Dipl.-Inf.
- 2012 Heinlein Support
- 2014 <https://www.meetup.com/de-DE/Ceph-Berlin/>

Lösung

- Samba-Cluster mit CTDB
- mehrere Samba-Knoten vor einem Ceph-Cluster
- CephFS für die Share-Daten und den CTDB
- Samba-Konfiguration verteilt in der CTDB-Registry

- Nur für File-Shares
 - nicht für ein Samba Active Directory Setup

- NFS-Exporte auch möglich, mit NFS-Ganesha
 - separates Setup per Ceph Orchestrator

CephFS

- POSIX-kompatibles Netzwerkdateisystem
- Speichert Dateidaten als Ceph-Objekte
- Ceph-Metadatenserver verwalten Ordnerstruktur
 - Verzeichnisse
 - Zugriffsrechte (auch ACLs)
 - Zeitstempel

- Client nur für Linux verfügbar

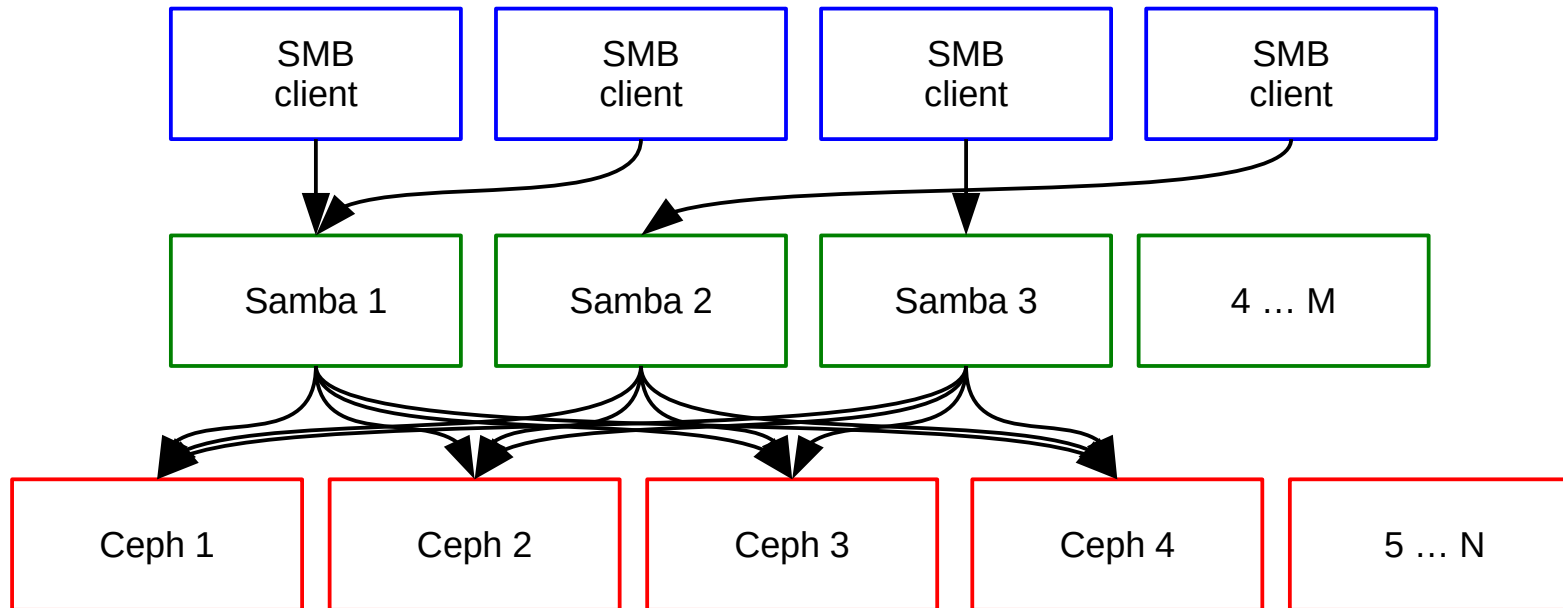
Samba

- gegründet 1992 von Andrew Tridgell
- Datei- und Druckerserver
 - SMB-Protokoll
- seit 2012 auch Active Directory Domain Controller
- Erweiterbar durch Virtual-FileSystem-Module
 - VFS_CEPH(8)

CTDB clustered trivial database

- TDB von Samba im Cluster
 - Datenkonsistenz
 - konsistentes Locking
- Automatisches Wiederverbinden und Reparatur der TDBs
- Hochverfügbarkeit durch
 - Überwachen der Knoten
 - IP-Adressübernahme
- Flexibel erweiterbar durch applikationsspezifische Managementskripte

Konzept



Ceph-Cluster Setup

- Benötigt wird
- ein laufender Ceph-Cluster

Einrichtung CephFS

- Mindestens zwei Metadata-Server
 - `ceph orch apply mds cephfs 'count:2'`

- Zwei Pools: Data & Metadata
 - Metadata-Pool auf schnellen OSDs (SSD, NVMe) platzieren
 - `ceph osd pool create cephfs_data`
 - `ceph osd pool create cephfs_metadata`

- Filesystem anlegen
 - `ceph fs new cephfs cephfs_metadata cephfs_data`

CephFS einbinden

- `/etc/ceph/ceph.conf`
 - Cluster-ID
 - IP-Adressen der Ceph-MONs
- Mount-Authentifizierung
 - `ceph fs authorize cephfs client.fsuser / rw`
 - `> /etc/ceph/ceph.client.fsuser.keyring`

 - damit auch möglich, nur Unterverzeichnisse zum Mount freizugeben

- `mount -t ceph -o name=fsuser :/ /mnt/cephfs`
- `ceph-fuse --id fsuser /mnt/cephfs`

Einrichtung Samba-Gateway

- Pakete installieren:
 - ctdb
 - samba
 - ggfs winbind
- /etc/ceph/ceph.conf & /etc/ceph/ceph.client.fsuser.keyring
- /etc/fstab
 - `:/ /mnt/cephfs _netdev,name=fsuser 0 0`
- /etc/systemd/system/ctdb.service.d/override.conf
 - [Unit]
After=mnt-cephfs.mount
RequiresMountsFor=/mnt/cephfs
 - Damit CTDB erst nach dem Mount startet

Konfiguration CTDB

- /etc/ctdb/nodes
 - IPs der Knoten
- /etc/ctdb/public_addresses
 - Service-IPs, pro Knoten eine
- Service-IPs in DNS round-robin
 - file.s2day.helein-akademie.de IN A 10.128.17.21
IN A 10.128.17.22
IN A 10.128.17.23
 - SMB-Client nehmen zum Mount eine beliebige Service-IP
 - Lastverteilung
- `ctdb event script enable legacy 50.samba`

Konfiguration Samba

- /etc/samba/smb.conf
 - [global]
clustering = yes
include = registry

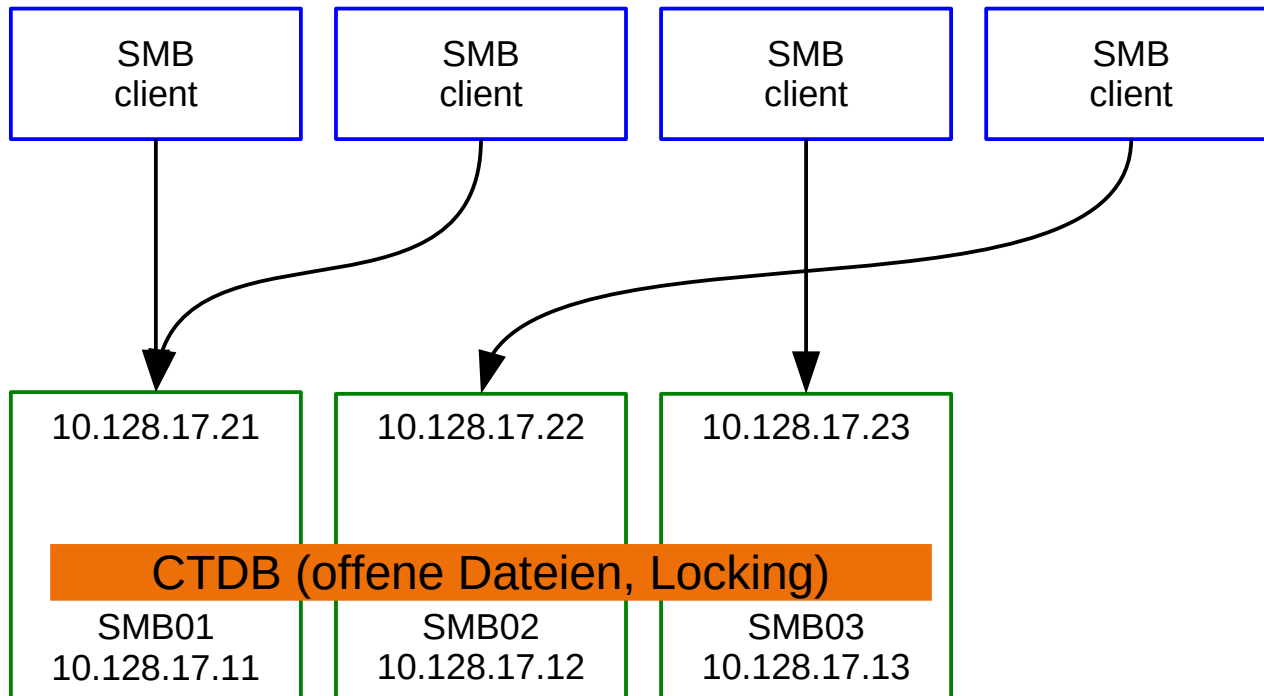
- net conf
 - net conf addshare
 - net conf setparm
 - net conf list
 - net conf import

- Sharepfade im CephFS

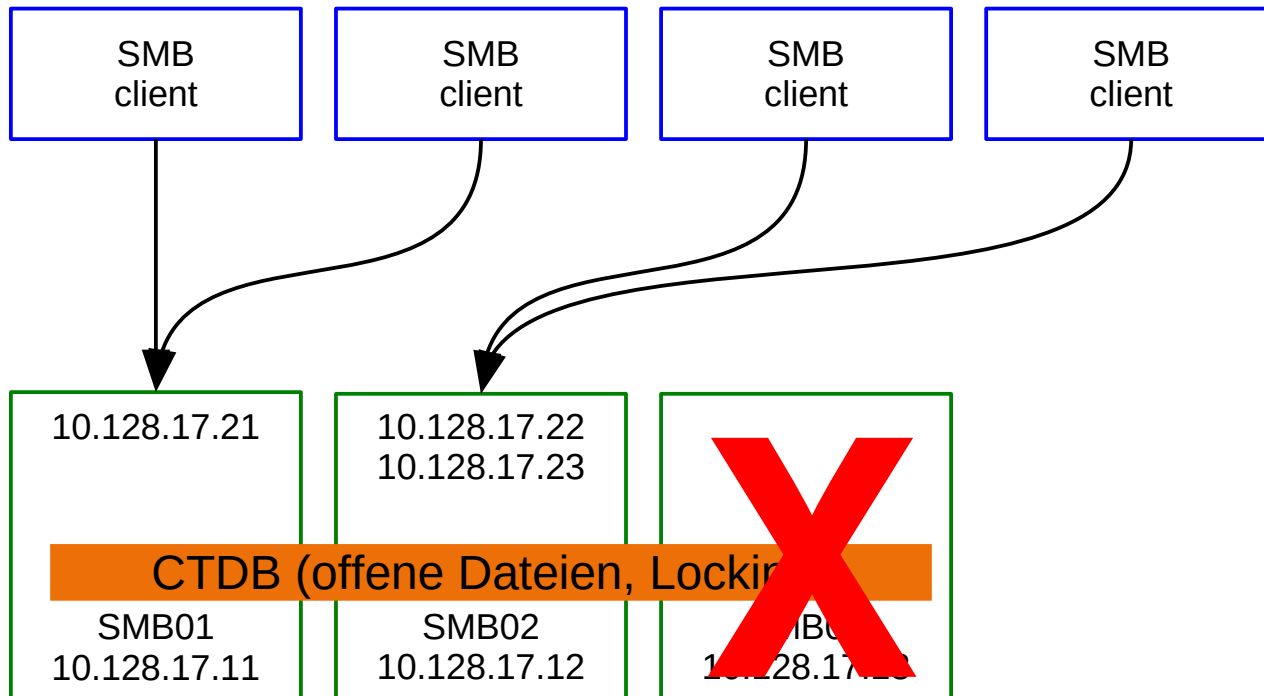
```
[share1]
    vfs objects = ceph
    path = /smb/share1
    comment = Share 1
    kernel share modes = no
    ceph:user_id = fsuser
    read only = no
```

- /etc/ctdb/script.options
 - CTDB_SAMBA_SKIP_SHARE_CHECK=yes

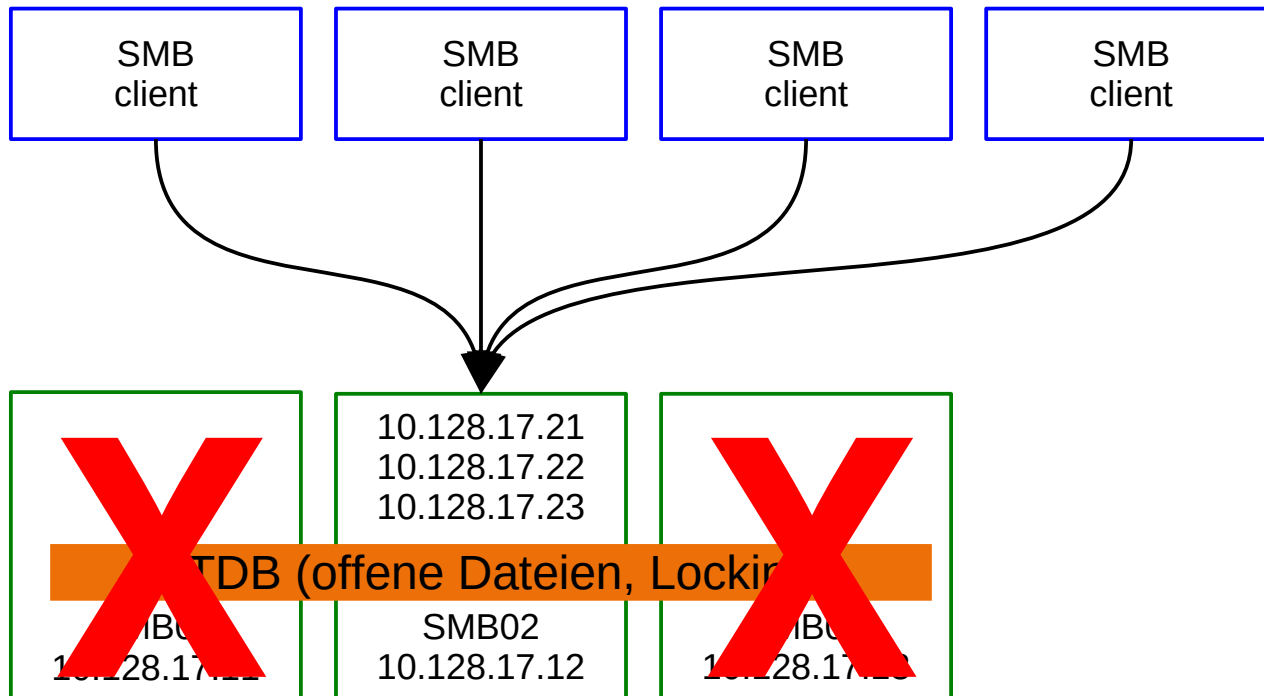
Normalzustand



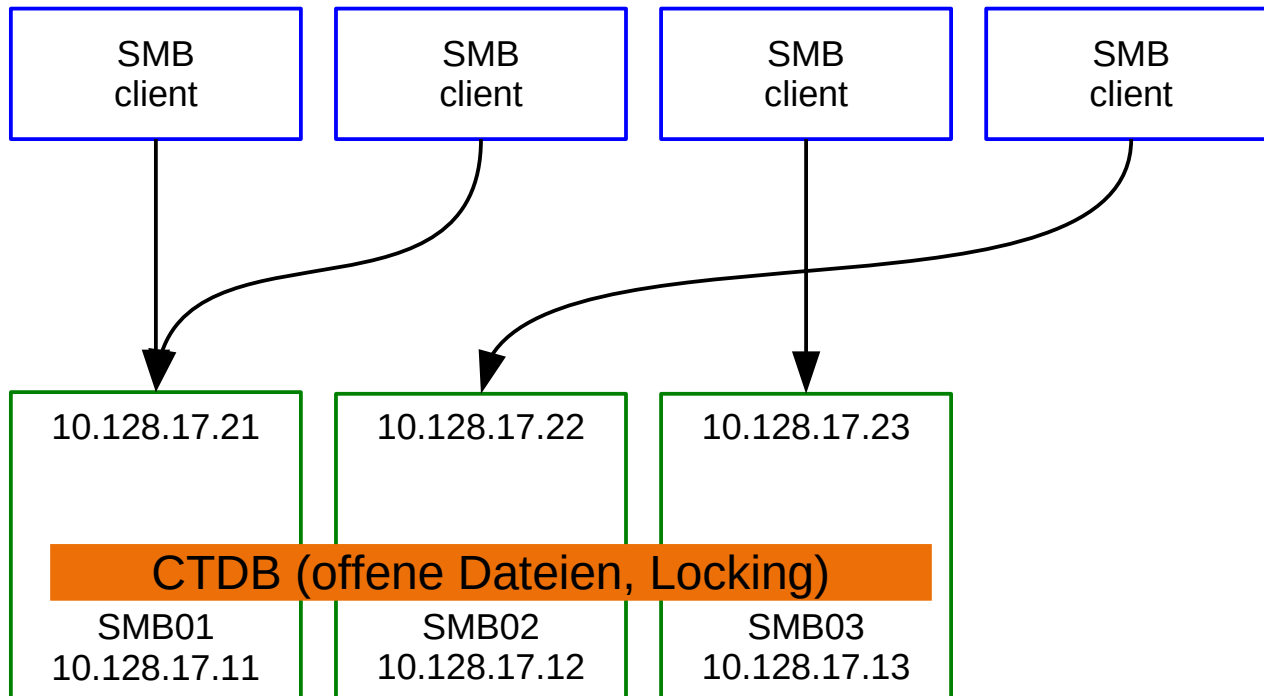
Failover



Failover



Recovery



Demo

Soweit, so gut.

**Gleich sind Sie am Zug:
Fragen und Diskussionen!**

Wir suchen:

Admins, Consultants, Trainer!

Wir bieten:

Spannende Projekte, Kundenlob, eigenständige Arbeit, ein tolles Team, Work-Life-Balance

...und natürlich: Linux, Linux, Linux...

<http://www.heinlein-support.de/jobs>

Heinlein Support hilft bei allen Fragen rund um Linux-Server

HEINLEIN AKADEMIE

Von Profis für Profis: Wir vermitteln die oberen 10% Wissen: geballtes Wissen und umfangreiche Praxiserfahrung.

HEINLEIN HOSTING

Individuelles Business-Hosting mit perfekter Maintenance durch unsere Profis. Sicherheit und Verfügbarkeit stehen an erster Stelle.

HEINLEIN CONSULTING

Das Backup für Ihre Linux-Administration: LPIC-2-Profis lösen im CompetenceCall Notfälle, auch in SLAs mit 24/7-Verfügbarkeit.