

# Hochverfügbares Linux mit Bordmitteln

## *Ausfallsicherheit für Lau (naja, fast)*

Oliver Kügow

team(ix) GmbH

ok@teamix.de

## Überblick

● Inhaltsverzeichnis

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

# Überblick

Überblick

● Inhaltsverzeichnis

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Einführung in Hochverfügbarkeit (High Availability, HA)
- Wir bauen ein Cluster
- Die Komponenten:
  - ◆ Hardware
  - ◆ Ethernet Bonding
  - ◆ File-Service (NFS/CIFS)
  - ◆ DRBD
  - ◆ Heartbeat
  - ◆ Zeit!
- DRBD Hands-On Hacking-Session

Überblick

---

Einführung in HA

- Wichtigkeit von IT-Systemen
- Kosten von Ausfällen
- Messung von Verfügbarkeit
- Ausfälle / Downtime

Cluster-Vorbereitung

---

Cluster-Bau

---

# Einführung in HA

- Je nach Branche und Einsatzzweck müssen unsere Systeme mehr oder weniger "verfügbar" sein.
- Verfügbarkeit bedeutet, dass das System den Dienst kontinuierlich verrichtet.
- Dabei reicht die Skala von:
  - ◆ "Ein User kann momentan keine Briefe schreiben"  
(Aber eigentlich egal)
  - ◆ "Alle im Büro können nicht arbeiten"  
(Es geht der Firma Geld verloren)
  - ◆ "Unsere Kunden können alle nicht mehr arbeiten"  
(uns UND den Kunden geht Geld verloren)
  - ◆ "Menschen sind in Gefahr"  
(z.B. Verkehrschaos: Ausfall von Ampeln)
  - ◆ "Menschen werden sterben"  
(Koordinierung der Blutversorgung von Krankenhäusern)

Überblick

---

Einführung in HA

● Wichtigkeit von IT-Systemen

● **Kosten von Ausfällen**

● Messung von Verfügbarkeit

● Ausfälle / Downtime

Cluster-Vorbereitung

---

Cluster-Bau

---

Entsprechend der Wichtigkeit von IT-Systemen sind die Kosten für Ausfälle entsprechend hoch:

- Pay TV: 0.1 Mill. EUR / Stunde
- Kreditkarten: 2.6 Mill. EUR / Stunde
- Aktienhandel: 6.4 Mill. EUR / Stunde

Das bedeutet, die Verfügbarkeit sollte möglichst hoch sein!

- Meist wird Verfügbarkeit bei gemieteten Diensten in Form eines SLAs (Service Level Agreements) festgelegt.
- Das SLA definiert, wieviel Zeit pro Monat (oder Jahr) das System verfügbar sein muss.
- Sonst drohen Strafen....
- Häufig definierte SLA-Werte:
  - ◆ 99% im Jahresmittel: 3,65 Tage Ausfall erlaubt
  - ◆ 99,9% im Jahresmittel: 0,365 Tage (=8,76 Stunden) Ausfall erlaubt
  - ◆ 99,99% im Jahresmittel: 0,0365 Tage (=53 Minuten) Ausfall erlaubt
  - ◆ 99% im Monatsmittel: 0,3 Tage (=7,32 Stunden)
  - ◆ 99,9% im Monatsmittel: 0,03 Tage (=43 Minuten)
  - ◆ 99,99% im Monatsmittel: 0,003 Tage (=4,3 Minuten)

- Wichtigkeit von IT-Systemen
- Kosten von Ausfällen
- Messung von Verfügbarkeit
- Ausfälle / Downtime

- Ausfälle werden in 2 Kategorien eingeteilt:
  - ◆ Geplante Ausfälle (Wartungsarbeiten, Software-Updates)
  - ◆ Ungeplante Ausfälle (Hardware-Fehler, Software-Fehler)
- Gegenmassnahmen:
  - ◆ Geplante Ausfälle: Wartungsfenster einführen oder Redundantes System
  - ◆ Ungeplante Ausfälle: Vermeidung eines SPOF durch Redundantes System
- Es läuft im Prinzip auf ein ausfallsicherers System raus.
- Diese Systeme werden oft als Cluster bezeichnet.



Überblick

---

Einführung in HA

---

Cluster-Vorbereitung

- Definition
- Hochverfügbarkeits-Cluster
- Wir bauen einen Cluster
- Zielaufbau
- Zutaten
- Vorgehensweise

Cluster-Bau

---

# Cluster-Vorbereitung

- Was versteht man unter einem Cluster?
  - ◆ Wörtliche Übersetzung: Anhäufung von etwas. In unserem Fall also eine "Anhäufung von Servern"
- Arten von Clustern:
  - ◆ Compute-Cluster: Kombination von Servern, um die Rechenleistung zu erhöhen
  - ◆ High Availability-Cluster: Kombination von Servern, um Redundanz zu erhöhen
  - ◆ Kombination aus beiden. Das wäre der Optimalfall, ist aber nur höherem Aufwand zu erreichen. (z.B. Google oder Oracle RAC)
- Eine Geschwindigkeitsverbesserung wäre schön, ist in unserem Fall aber nicht wichtig.
- Wir wollen die Verfügbarkeit verbessern.

Überblick

---

Einführung in HA

---

Cluster-Vorbereitung

---

● Definition

● **Hochverfügbarkeits-Cluster**

● Wir bauen einen Cluster

● Zielaufbau

● Zutaten

● Vorgehensweise

Cluster-Bau

---

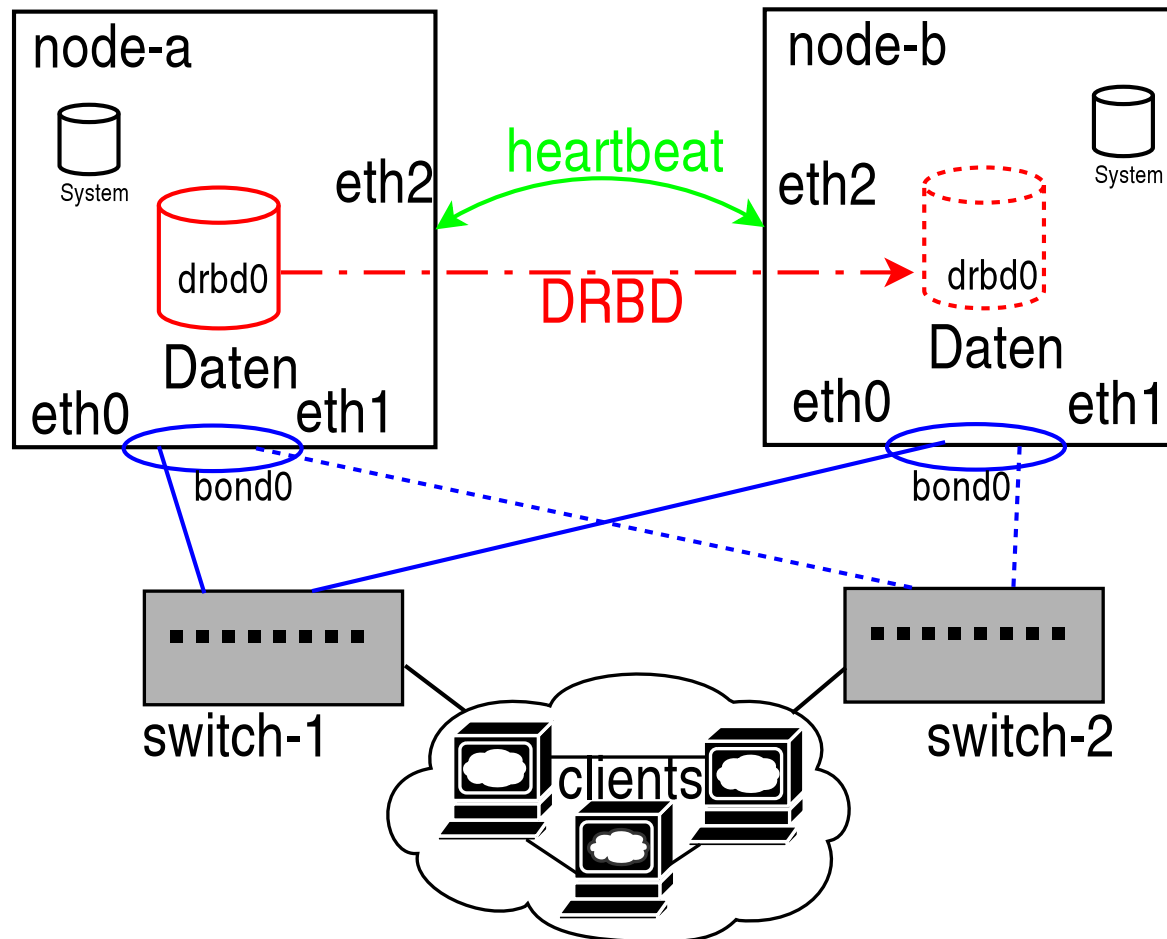
- (Ja, gleich ist Schluss mit der Theorie!!)
- Wir unterscheiden bei HA-Clustern noch zwischen:
  - ◆ Aktiv/Passiv Cluster: Ein Rechner (Knoten) arbeitet, während der andere auf den Ausfall des ersten Knoten wartet.
  - ◆ Aktiv/Aktiv Cluster: Beider Knoten arbeiten, fällt einer aus, so hat der andere die doppelte Arbeit.
- Der Einfachheit halber beschränken wir uns auf ein Aktiv/Passiv Cluster

- Definition
- Hochverfügbarkeits-Cluster
- **Wir bauen einen Cluster**
- Zielaufbau
- Zutaten
- Vorgehensweise

- (ja, das wars mit der Theorie!)
- Ziel: Aufbau eines redundanten File-Servers unter Linux
- Es soll wirklich alles redundant sein:
  - ◆ Netz-Anbindung
  - ◆ Storage
  - ◆ Applikationen
  - ◆ Rest der Hardware. . .
- Unabhängig von unserem "komplexen" Setup die einfachen Dinge nicht vergessen: Ausfallsicherer Strom!!

- Definition
- Hochverfügbarkeits-Cluster
- Wir bauen einen Cluster
- Zielaufbau
- Zutaten
- Vorgehensweise

So soll der Setup aussehen, wenn er fertig ist:



Überblick

---

Einführung in HA

---

Cluster-Vorbereitung

---

- Definition
- Hochverfügbarkeits-Cluster
- Wir bauen einen Cluster
- Zielaufbau
- **Zutaten**
- Vorgehensweise

Cluster-Bau

---

Man nehme:

**Hardware** 2xServer, 2xSwitch, ausreichend LAN-Karten

**”Linux”** für ein stabiles Betriebssystem

**”NFSd”** als NFS-Server

**”Ethernet-Bonding”** für redundante Netzanbindung

**”DRBD”** für Replikation der Nutz-Daten auf beide Rechner

**”Heartbeat”** für den eigentlich Cluster-Mechanismus

Überblick

---

Einführung in HA

---

Cluster-Vorbereitung

---

- Definition
- Hochverfügbarkeits-Cluster
- Wir bauen einen Cluster
- Zielaufbau
- Zutaten
- **Vorgehensweise**

Cluster-Bau

---

1. Linux-Installation mit richtiger Plattenpartitionierung
2. Installation des File-Service NFS
3. Replikation der Storage mit DRBD
4. Redundanz der Netzanbindung
5. Cluster-Modus mit Heartbeat
6. Den Cluster-Knoten klonen

Überblick

---

Einführung in HA

---

Cluster-Vorbereitung

---

**Cluster-Bau**

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

# Cluster-Bau



Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

● Linux-Installation

● File-Service NFS

● DRBD

● DRBD - Installation

● DRBD - Konfiguration 1

● DRBD - Konfiguration 2

● DRBD - Initialisierung

● DRBD - Erster Kontakt

● DRBD - Konfigurationsdatei 1

● DRBD - Konfigurationsdatei 2

● DRBD - Verwendung

● Ethernet-Bonding

● Bonding Konfiguration

● Heartbeat - Cluster

● Heartbeat - Installation

● Heartbeat - Konfiguration 1

● Heartbeat - Konfiguration 2

● Heartbeat - Konfiguration 3

● Heartbeat - Was geht?

● Fragen ?

- Das "richtige" Linux ist jedem selbst überlassen, für mich ist Debian richtig.
- Plattenpartitionierung ist wichtig
  - ◆ Teil für das Betriebssystem (muss nicht redundant sein)
  - ◆ Nutzerdaten (müssen repliziert werden)
- Deshalb wählen wir ca. 5 GB für das OS (inkl. Logfiles)  
(`/dev/hda1`)
- und den Rest der Platte für den File-Service  
(`/dev/hda2`)
- Auf "Nebensächlichkeiten" wie Swap, feinere Partitionierung usw verzichte ich hier bewusst!
- Bitte nur einen der beiden Cluster-Knoten installieren. Das System wird dann mit Hilfe von einer Knoppix-CD einfach geclont.

- Die eigentliche Applikation, die wir ausfallsicher machen wollen!
- `apt-get install nfs-user-server`
- Verzeichnisfreigabe anlegen: `mkdir /daten`
- Konfigdatei `/etc/exports` bearbeiten:  
`/daten/ * .firma.de(rw,all_squash, sync)`
- Sicherstellen, dass NFS NICHT automatisch gestartet wird:  
`rm /etc/rc{2,3}.d/*nfs-user-server*`
- Gestartet wird NFS nämlich von Heartbeat aus

- Das Distributed Replicating Block Device hilft uns die Daten die auf einen Rechner geschrieben werden automatisch in Echtzeit auf den passiven Clusterknoten zu kopieren (synchroner Spiegel)
- prinzipiell wäre sowas auch mit rsync möglich, dann wäre der Vorgang aber zeitversetzt (asynchroner Spiegel)
- Wie funktioniert DRBD?
  - ◆ Stellt ein neues Block-Device zur Verfügung (`/dev/drbd/0`)
  - ◆ Beim Schreiben auf die virtuelle Platte werden die Daten
  - ◆ - einmal auf die echte Platte geschrieben (`/dev/hda2`)
  - ◆ - zusätzlich auf die Platte des passiven Clusterpartners geschrieben. Das geschieht über das Netz
  - ◆ Ein Schreibzugriff wird erst als erfolgreich bestätigt, wenn BEIDE Clusterknoten die Daten auf Platte geschrieben haben.

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- **DRBD - Installation**
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Runterladen: von DRBD.org, momentan ist v0.7.4 aktuell und stabil
- Entpacken: `tar xvfz drbd-0.7.4.tar.gz`
- **ACHTUNG:** Die Kernel-Quellen müssen installiert sein, sonst kann das Modul nicht kompiliert werden!!1
- Kompilieren:
  - ◆ `make KDIR=/path/to/kernel/source`
  - ◆ `make install`
- Ergebnis:
  - ◆ Kernel-Modul `drbd.o`
  - ◆ User-Space-Tools `drbdsetup` und `drbdadm` (Wrapper-Skript für `drbdsetup`)
  - ◆ Beispiel-Configuration `/etc/drbd.conf`

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

Einmal den Vorgang von Hand durchspielen, damit wir wissen obs geht:

Modul laden: `insmod drbd.o`

```
node-a kernel: drbd0: mdev = 0xa0d3e000
```

```
node-a kernel: drbd1: mdev = 0xa0d3e4c8
```

```
node-a kernel: drbd: initialised. Version: 0.7.4 (api:7
```

```
node-a kernel: drbd: SVN Revision: 1537M build by root@
```

```
node-a kernel: drbd: registered as block device major 1
```

danach ein drbd-device erstellen:

```
drbdsetup drbddev disk lower_dev meta_data_dev +
                                meta_data_index
```

```
drbdsetup drbddev net localip remoteip protocol
```

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

## Beispiele:

```
node-a# drbdsetup /dev/drbd/0 disk /dev/hda2 +
                                internal -1
```

```
node-a# drbdsetup /dev/drbd/0 net 192.168.42.1 +
                                192.168.42.2 C
```

## Ergebnis:

```
node-a kernel: drbd0: Creating state block
node-a kernel: drbd0: resync bitmap: bits=98816 words
node-a kernel: drbd0: size = 386 MB (395264 KB)
node-a kernel: drbd0: Assuming that all blocks are
                        out of sync (aka FullSync)
node-a kernel: drbd0: 395264 KB now marked
                        out-of-sync by on disk bit-map.
node-a kernel: drbd0: drbdsetup [165]: cstate
                        Unconfigured --> StandAlone
node-a kernel: drbd0: drbdsetup [167]: cstate
                        StandAlone --> Unconnected
```

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Sobald beide Clusterknoten aktiv sind sollten sie sich finden, der Status ist dann `Connected`
- Beide werden anfangs automatisch Slave, der Master wird manuell bestimmt
- Das übernimmt später die Clustersoftware für uns
- Beim Initialisieren des DRBD wird vom Master alles auf den Slave kopiert
- Der erste Sync dauert lange, weil alle Blöcke kopiert werden müssen
- Danach sind beide im Status `Sync` und `Connected`
- Man kann dann wieder mit `cat /proc/drbd` den Status anschauen.

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

Prüfung jeweils mit `cat /proc/drbd`

Partner noch nicht da:

```
0: cs:WFConnection st:Secondary/Unknown ld:Consistent
```

Partner da:

```
0: cs:Connected st:Secondary/Secondary ld:InConsistent
```

Jetzt Master werden:

```
drbdadmin primary all
```

Erster Sync

```
0: cs:SyncSource st:Primary/Secondary ld:InConsistent
[==>.....] sync'ed: 2.1% (389668/395264)
finish: 0:03:28 speed: 1,864 (1,864) K/sec
```

Danach:

```
0: cs:Connected st:Primary/Secondary ld:Consistent
```



- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- **DRBD - Konfigurationsdatei 1**
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

Jetzt wo alles geht, die Konfiguration persistent machen:

Datei ist: `/etc/drbd.conf`

```
resource "r1" {
    protocol C;
    incon-degr-cmd "sleep 60 ; halt -f";
    startup {
        wfc-timeout          0;    ## Infinite!
        degr-wfc-timeout    120;   ## 2 minutes.
    }
    disk {
        on-io-error detach;
    }
    net {
        # timeout          60;
        # connect-int     10;
        # ping-int        10;
        # max-buffers      2048;
        # max-epoch-size   2048;
    }
}
```

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- **DRBD - Konfigurationsdatei 2**
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

```

syncer {
    rate      4M;
    group    1;
}
on node-a {
    device    /dev/drbd/0;
    disk      /dev/hda2;
    address   192.168.42.1:7789;
    meta-disk internal;
}
on node-b {
    device    /dev/drbd/0;
    disk      /dev/hda2;
    address   192.168.42.2:7789;
    meta-disk internal;
}
}
    
```

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- ERST VERWENDBAR, WENN MAN MASTER IST!!
- Wie eine normale Festplatte/Partition
- `mkfs.ext3 /dev/drbd/0`
- `mount /dev/drbd/0 /daten`
- am besten einfach in die `/etc/fstab` eintragen:  
`/dev/drbd/0 /daten ext3 noauto 0 0`
- **WICHTIG:** nicht automatisch beim booten mounten. Das kann nur der Cluster-Master
- Wer Master ist, bestimmt heartbeat.

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- **Ethernet-Bonding**
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Mehrere physikalische Netzwerkkarten werden zu einer "virtuellen Karte" zusammengefasst
- Zweck kann Lastverteilung oder Ausfallsicherheit sein. Wir wollen letzteres.
- Karte eth0 wird an Switch 1 angeschlossen, Karte eth1 wird an Switch 2 angeschlossen
- Es soll immer nur eine Karte aktiv sein, die andere wartet auf einen Ausfall. . .
- . . . Ausfall des Switches oder der anderen Karte.

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- **Bonding Konfiguration**
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Tools fürs Bonding installieren:  
`apt-get install ifenslave`
- Dann Kernelmodul laden:  
`insmod bond miimon=100 mode=active-backup`
- `ifenslave bond0 eth0 eth1`  
(eth0 in dem Fall das aktive Interface)
- Danach wie eine normale Ethernetkarte verwenden:
- `ifconfig bond0 192.168.42.1 up`
- Es kann dabei zu Timing Problemen mit div. Switches kommen, Doku lesen!
- Alle Konfigurationen müssen noch persistent gemacht werden!!  
`/etc/network/interfaces`

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- **Heartbeat - Cluster**
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Heartbeat ist die eigentliche Cluster-Software
- Heartbeat tut sehr wenig, die komplexen Probleme (Storage Replikation, Applikationsintegrität) sind anders zu lösen.
- Es werden keine Dienste von HB überwacht, sondern nur ob der andere Knoten läuft
- Ist er ausgefallen, dann werden die Dienste auf dem Backup-Server gestartet und die Service-IP übernommen
- Um zu vermeiden, dass der beide Knoten glauben Master zu sein, kann HB einem Knoten den Saft abdrehen (STONITH)

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- **Heartbeat - Installation**
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- `apt-get install heartbeat`
- und bitte dafür sorgen, dass ALLE Dienste, die im Cluster laufen sollen NICHT automatisch gestartet werden.

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- **Heartbeat - Konfiguration 1**
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- 3 Konfigurationsdateien:
- `/etc/ha.d/ha.cf` - Globale Konfiguration, welche Knoten existieren, wie wird überwacht.
- `/etc/ha.d/haresources` - Welche Dienste werden auf den Knoten gestartet
- `/etc/ha.d/authkeys` - Enthält Schlüssel zur Authentifizierung der Cluster-Partner
- und in `/etc/ha.d/rc.d` werden die normalen Startskripte der Applikationen verlinkt. So kann Heartbeat selbst die Dienste starten



Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

`/etc/ha.d/ha.cf` - Globale Konfiguration, welche Knoten existieren, wie wird überwacht.

```
logfacility daemon # Log to syslog as facility "daemon"
node node-a node-b # List our cluster members
keepalive 1 # Send one heartbeat each second
deadtime 10 # Declare nodes dead after 10 seconds
bcast eth2 # Broadcast heartbeats on eth2
ping 192.168.42.254 # Ping our router to monitor ethernet
auto_failback no # Don't fail back to node-a automatically
respawn hacluster /usr/lib/heartbeat/ipfail # Failover
```

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- `/etc/ha.d/haresources` - Welche Dienste werden auf den Knoten gestartet

```
node1 datadisk::drbd0 Filesystem::/dev/drbd/0::/data+
192.168.42.42 nfs-user-server

node2
```

- `/etc/ha.d/authkeys` - Enthält Schlüssel zur Authentifizierung der Cluster-Partner

```
auth 1
1 sha1 SehrGeheimesPasswort
```

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Beide Clusterknoten booten
- *ABER*: Sie starten nicht die üblichen Dienste wie NFS o.Ä. (schliesslich haben wir ja die rc-Skripte gelöscht!)
- Dafür wird heartbeat gestartet.
- Der prüft, ob er Master oder Slave Knoten ist.
- Im Falle des Masters führt wird `/etc/ha.d/haresources` interpretiert und ...
- ... alle Start-Skripte der gewünschten Dienste werden ausgeführt.
- In unserem Fall sind das:
  - DRBD-Master werden, und das Block-Device Mounten
  - Dann sich die Service-IP zusätzlich aufs Netzwerk-Interface binden
  - Und zuletzt den NFS-Server starten.
  - Der Cluster sollte nun laufen

Überblick

Einführung in HA

Cluster-Vorbereitung

Cluster-Bau

- Linux-Installation
- File-Service NFS
- DRBD
- DRBD - Installation
- DRBD - Konfiguration 1
- DRBD - Konfiguration 2
- DRBD - Initialisierung
- DRBD - Erster Kontakt
- DRBD - Konfigurationsdatei 1
- DRBD - Konfigurationsdatei 2
- DRBD - Verwendung
- Ethernet-Bonding
- Bonding Konfiguration
- Heartbeat - Cluster
- Heartbeat - Installation
- Heartbeat - Konfiguration 1
- Heartbeat - Konfiguration 2
- Heartbeat - Konfiguration 3
- Heartbeat - Was geht?
- Fragen ?

- Die erste Frage lautet:
- Mit welchem Tool wurde diese Präsentation erstellt?
- Das Tool heisst HA-Prosper. Es ist eine Erweiterung des  $\text{\LaTeX}$  Paketes namens Prosper, das zur Folien-Erstellung dient. Die wichtigste Erweiterung von HA-Prosper sind die Menüs auf der linken Seite.
- Die zweite Frage lautet:
- Warum erzählst du uns nur die Theorie und zeigst uns nicht einfach wie es geht?
- w00t !! Ihr seid viel zu ungeduldig!!11
- Dafür ist die Hacking-Session vorgesehen, die jetzt im Anschluss stattfindet!
- So, und jetzt eure Fragen
- und: Danke fürs Zuhören