

UGS-Failover unter UCS

Thema:	Einrichtung und Inbetriebnahme eines Cold-Standby-Failover-Groupwaresystems
Datum:	15. Dezember 2009
Seitenzahl:	6
Versionsnummer:	4532
Autoren:	Univention GmbH feedback@univention.de

Inhaltsverzeichnis

1	Erstellung der Failover-Kopie	3
2	Inbetriebnahme im Fehlerfall / Test der Failover-Kopie	5

In diesem Dokument wird die Bereitstellung eines Failover-Systems für einen UGS-Groupwareserver im Netz der Unfallkasse NRW beschrieben.

Dabei wird eine exakte Kopie des Servers als virtuelle Maschine erzeugt, mit der im Fall eines Hardwareschadens der Groupware-Betrieb weitergeführt werden kann. Die zur Inbetriebnahme der Kopie im Fehlerfall notwendigen Aktualisierungsschritte werden ebenfalls aufgeführt.

Die Failover-Kopie des Servers muss wie in Kapitel 1 beschrieben nach allen umfassenden System-Updates am Groupware-Server neu durchgeführt werden, da sonst ein veraltetes System für den Failover-Fall vorgehalten wird, das potentiell nicht mit dem aktuellen Datenbestand arbeiten kann.

Dasselbe gilt für wichtige administrative Änderungen am UGS-System; wenn diese sofort nach dem Ausfall des Groupware-Servers auch in der inbetriebgenommenen Kopie zur Verfügung stehen sollen, so müssen sie auch in dem kopierten System nachgetragen werden.

1 Erstellung der Failover-Kopie

Da die Mail-Daten auf einem Network Attached Storage Device liegen wird nur eine Kopie der System-Installation benötigt, da im Fehlerfall die aktuellen Groupware-Daten aus dem Mailspool bezogen werden und die LDAP-Daten über die Mittel der Listener-/Notifier-Replikation aktualisiert werden.

In folgendem Dokument wird davon ausgegangen, dass die beiden Datenverzeichnisse des Cyrus-Mailserver (`/var/spool/cyrus` und `/var/lib/cyrus`) als einzelne Partitionen von einem SAN-Device über iSCSI eingebunden werden und in die `/etc/fstab` eingetragen wurden.

Um eine verlässliche Kopie der Daten zu ermöglichen, sollte sowohl das zu clonende Groupware-DC-Slave-System (nachfolgend Quellsystem genannt), als auch die virtuelle Maschine, in die das Failover-System installiert werden soll (nachfolgend Zielsystem genannt) durch eine Linux-Live-CD gebootet werden. Beispielsweise Knoppix, welches unter der URL <http://www.knoppix.org> heruntergeladen werden kann.

Wenn beide Systeme von der Live-CD gebootet wurden, sollte zwischen ihnen eine Netzwerk-Verbindung hergestellt werden, indem bei beiden eine IP-Adresse und das Gateway konfiguriert wird. Beide Rechner sollten sich gegenseitig durch `ping` erreichen können.

Auf dem Quellsystem muss nun die bestehende Festplatten-Partitionierung ermittelt werden. Diese kann durch den Befehl

```
cfdisk /dev/sda
```

ermittelt werden (bzw. `/dev/sdb` für eine zweite System-Platte).

Auf dem Zielsystem müssen nun Partitionen gleicher Grösse angelegt werden. Dies geschieht ebenfalls durch das Kommando `fdisk`. Es reicht die Festplatteninformationen abschliessend mit **Write** abzuschliessen. Die neu angelegten Partitionen können auch grösser sein als die Partitionen auf dem Quellsystem, der überzählige Plattenplatz wird in diesem Fall nicht genutzt.

Es ist nicht notwendig mit `mkfs` Dateisysteme auf den angelegten Partitionen zu erzeugen, da im folgenden Schritt eine bitweise Kopie erzeugt wird.

Die Daten werden durch das `netcat`-Tool über das Netz kopiert:

Auf dem Quellsystem muss zur Bereitstellung der Daten folgender Befehl eingegeben werden:

```
dd if=/dev/sdaX | nc -l -p 10000
```

Auf dem Zielsystem werden die Daten durch den folgenden Befehl entgegen genommen und auf die Festplatte geschrieben:

```
nc IP-DES-QUELLSYSTEMS 10000 | dd of=/dev/sdaX
```

Das Kopieren der Daten kann je nach Geschwindigkeit der Netzverbindung und der Größe der kopierten Partitionen bis zu eine Stunde dauern.

Sobald alle Daten auf dem Quellsystem über das Netz kopiert wurden, erscheint dort eine Ausgabe, die etwa wie folgt aussieht: (die tatsächlichen Werte sind natürlich dementsprechend grösser)

```
41225+0 Datensätze ein
41225+0 Datensätze aus
21107200 Bytes (21 MB) kopiert, 0,609004 Sekunden, 34,7 MB/s
```

Dies bedeutet, dass auf dem Quellsystem alle Daten gesendet wurden. Mit Strg+C kann der Kopievorgang dann beendet werden.

Die obigen Schritte wurden exemplarisch als `/dev/sdaX` angegeben und müssen für jede Partition einzeln durchgeführt werden. Die Festplattenpartitionen werden dabei bitweise vom Quell- auf das System kopiert. Damit das Zielsystem auch von den kopierten Partitionen booten kann, muss noch der Bootloader GRUB installiert werden:

GRUB wird in den Master Boot Record und in das Verzeichnis `/boot` kopiert. Dazu muss das Verzeichnis einmalig manuell über `mount` eingebunden werden: (Hier am Beispiel eines `/boot` auf `/dev/sda1`.)

```
mkdir /tmp-boot
mount -t ext3 /dev/sda1 /tmp-boot/
```

Nun kann durch den Befehl `grub-install` ein Bootsektor in die Festplatte installiert werden. In diesem Beispiel wird GRUB in die erste Systemplatte installiert:

```
grub-install hd0 --root-directory=/tmp-boot/
```

Für die folgenden Tests sollte das Quellsystem heruntergefahren werden. Danach kann das Zielsystem als Kopie gebootet werden.

In UCS 2.0 werden Treiber durch den Gerätemanager udev verwaltet. Dieser bietet eine persistente Zuweisung von Netzwerk-Interfaces auf einzelne Hardware (identifizierbar bei Ethernet durch die MAC-Adresse). In dem hier skizzierten Anwendungsfall ist dieser Mechanismus jedoch eher hinderlich; da das Zielsystem ein neues Netzwerk-Interface verwendet, bekommt sie im kopierten System statt der vorher verwendeten **eth0** den Namen **eth1**.

Die Zuweisungen werden in der Datei `/etc/udev/rules.d/z25_persistent-net.rules` gespeichert. Der alte **eth0**-Eintrag muss dabei mit einem führenden `#` auskommentiert werden und der neu angelegte Eintrag auf **eth0** geändert werden, wie im nachfolgenden Beispiel zu sehen:

```
#SUBSYSTEM=="net", DRIVERS=="?*", ATTRS(address)=="00:0c:29:8d:71:3f", NAME="eth0"

# PCI device 0x1022x0x2000 (pcnet32)
SUBSYSTEM=="net", DRIVERS=="?*", ATTRS(address)=="00:0c:29:59:d3:36", NAME="eth0"
```

Nach einem Neustart des System steht dann wieder ein **eth0**-Netzwerk-Interface zur Verfügung.

Das Zielsystem kann anschliessend heruntergefahren werden und für den Fehlerfall vorgehalten werden.

2 Inbetriebnahme im Fehlerfall / Test der Failover-Kopie

Wenn auf dem Groupware-Server ein Hardware-Schaden auftritt, sind die folgenden Schritte notwendig, um wieder ein funktionsfähiges System zu erlangen. Derselbe Mechanismus kann auch zum Test der Failover-Kopie in einer lastarmen Zeit verwendet werden. Dabei sollte mit einem Vmware-Snapshot der Failover-Kopie gearbeitet werden, der einfach wieder zurückgesetzt werden kann. Für Tests sollte das Original-Groupware-System abgeschaltet werden.

- Der defekte Groupware-Server sollte abgeschaltet werden.
- Die virtuelle Maschine mit der System-Kopie wird aktiviert und ist unter dem identischen Hostnamen und IP-Adresse erreichbar. Da die IMAP-Daten noch nicht wieder verfügbar sind und die LDAP-Daten noch nicht aktuell sind, sollte der Cyrus-Server vorerst gestoppt werden:

```
/etc/init.d/cyrus2.2 stop
```

- Nun muss geprüft werden, ob die eigentlichen Daten des Mail-Spools vom SAN-Device eingebunden wurden. Ist die nicht der Fall (etwa weil das Netz-Device während des Bootvorgangs nicht verfügbar war) müssen sie von Hand durch `mount` eingebunden werden.

- Die Aktualisierung der LDAP-Daten startet automatisch. Es ist zu beachten, dass sämtliche LDAP-Transaktionen seit dem Anlegen der Groupware-Server-Kopie eingespielt werden. Der Status des Zurückspielens der Replikations-Daten kann entweder durch eine bestehende Nagios-Installation über das Plugin UNIVENTION_REPLICATION oder auch manuell geprüft werden:

Durch Abgleich der Transaktions-IDs geprüft werden. (Ermittelbar für die Quelldaten auf dem Domänencontroller Master durch den Befehl `tail -1 /var/lib/univention-ldap/notify/transaction` - die erste Zahl in der ausgegebenen Teile ist die Transaktions-ID - und für die Zieldaten auf dem in Betrieb genommenen Domänencontroller Slave durch den Befehl `cat /var/lib/univention-directory-listener/notifier_id`)

- Abschliessend kann der Cyrus-Dienst wieder gestartet werden:

```
/etc/init.d/cyrus2.2 start
```

- Die Benutzer sollten jetzt normal weiterarbeiten können.