

Aber mit einem kleinen Script kann man das Problem lösen:

```
# xm2 list
Shortname  Name                                     ID  Mem  VCPUs  State  Time (s)
slot031    0004fb0000060000059ec32f98918b07      20 2048    2 -b---- 51600.7
ovminfra   0004fb00000600004c953ac4246637eb      19  512    1 -b----  756.4
master249  0004fb00000600008219e1e797e01b0d      29 2048    2 -b----  1003.8
slot001    0004fb000006000088cde8cce067a63b      24 2048    2 r----- 60829.3
           Domain-0                                0 1134    16 r----- 474750.4
```

Der „xm“ Befehl kann auch zum Starten und Stoppen von VM's verwendet werden. Wenn wir also das oben erwähnte Script etwas erweitern, dann können wir bequem mit „xm2 stop slot033“ die so benannte VM stoppen. Aber dürfen wir das? Im Vortrag schauen wir an, wie sich das auf das Framework vom OVM Manager auswirkt.

Die Konsole einer VM erreicht man über ein entsprechendes Icon im OVM Manager. Die Implementierung erfolgt über das VNC-Protokoll. Im XEN-Umfeld ist dies aber nur eine Variante für die Konsolendarstellung. XEN selbst bietet auch eine Textmodus-Konsole an. Im Vortrag werden wir anschauen, wie man diese Konsole nutzen kann, sei es temporär, oder dauerhaft. Keine Lust auf Kommandozeile? Spätestens wenn man seine paravirtualisierte VM so verkonfiguriert hat, dass sie gar keinen Kernel mehr zum Booten findet, wird man die Textmodus-Konsole benötigen. Denn die VNC-Konsole funktioniert erst ab dem Zeitpunkt, wo der Kernel gefunden wurde. Der Boot-Manager ist nämlich nur in der Textmodus-Konsole sichtbar.

Sie wollen 100 neue virtuelle Disks erstellen, aber sich nicht 100x durch das GUI klicken? Auch das lässt sich auf der Kommandozeile elegant erledigen.

Snapshots

Ein Snapshot speichert den Zustand der Disks einer VM. Diese Disks sind in der Regel durch Dateien auf dem Host abgebildet. Produkte wie VirtualBox oder VMWare erstellen Snapshots applikatorisch. Nach dem Erstellen eines Snapshots wird die Original Disk-Datei nicht mehr verändert, sondern Änderungen werden in einer Differenzdatei gespeichert. Es ist auch möglich, von einem Snapshot wieder einen Snapshot zu erstellen, wobei sich die Virtualisierungslösung die Blöcke aus der eigenen und der vorangegangenen Differenzdatei, sowie aus der Originaldatei zusammensuchen muss. Vorteil dieser Lösung ist, dass sie unabhängig vom Betriebssystem oder Dateisystem funktioniert. Der Nachteil ist jedoch, dass solche Snapshots durch den Verwaltungs-Overhead eine schlechtere Performance haben. Und wenn man dann einen Snapshot löscht, müssen die geänderten Blöcke aufwendig wieder in die Originaldatei zurückgeschrieben werden

Mit dem aktuellen OVM3 Manager sind wir nicht in der Lage, den Zustand einer VM zwischenspeichern. Stattdessen gibt es die Möglichkeit, die gesamte VM zu klonen. Dabei werden die Disk-Files kopiert und ein neues Konfigurationsfile (die Definition der virtuellen Maschine) erzeugt. Wir halten also nicht einen Zustand unserer VM fest, sondern wir erzeugen eine neue VM (oder ein sogenanntes Template), die aktuell identischen Inhalts ist wie das Original, und beim Starten neu konfiguriert werden muss (Anpassen Netzwerk-Konfiguration (neue MAC-Adresse), Hostname usw.).

Falls wir unsere VM's auf lokalem LUN-Storage (SAN, iSCSI) abgelegt haben, dann verwendet OVM ein ocfs2 Clusterfilessystem. Dieses bietet seit Kurzem eine sehr praktische Funktion: Es können von Files schreibbare Snapshots erstellt werden, in der ocfs2-Terminologie sogenannte ReFlinks.

Wenn eine Datei so kopiert wird, dann wird nicht der Inhalt, sondern vorerst nur die Inodes (die Zuordnungstabelle von den Blöcken zur Datei) kopiert. Solange man keine der beiden Dateien verändert, braucht die Kopie keinen zusätzlichen Platz.

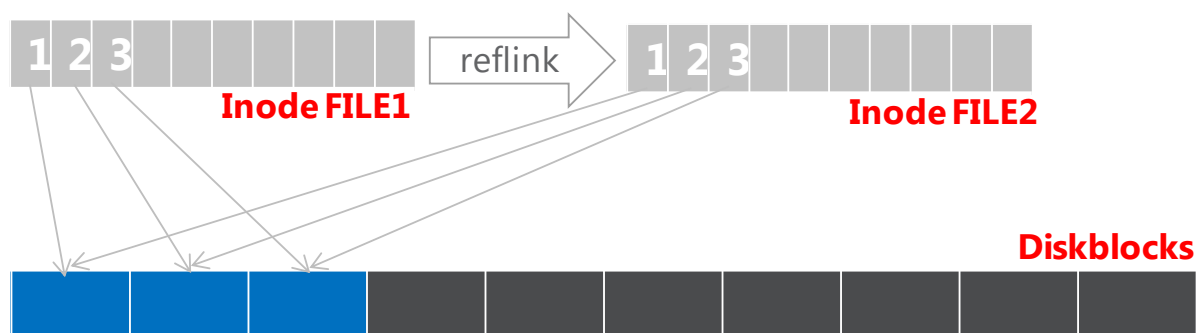


Abb. 1: Ergebnis einer Reflink-Kopie

Erst beim Ändern eines Blocks wird der Block kopiert (Copy-On-Write), beziehungsweise der Originalblock bleibt erhalten, und für den geänderten Block wird ein neuer alloziert. Änderungen sind sowohl an File1 als auch an File2 möglich.

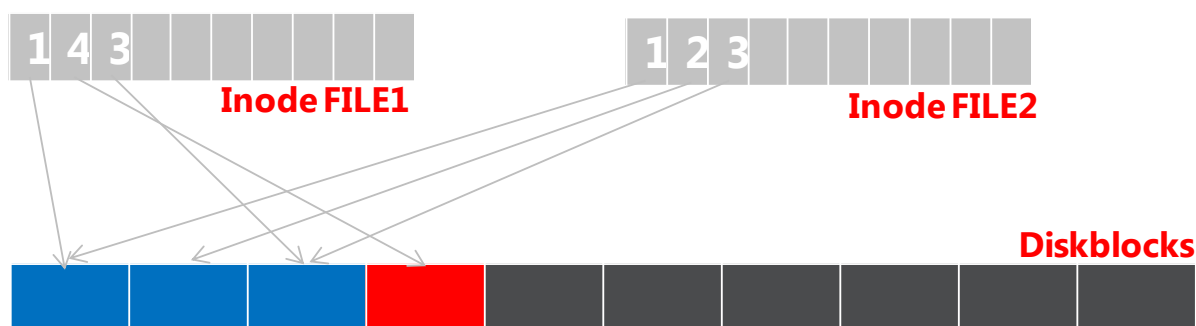


Abb. 2: Ergebnis nach Änderung eines Blocks in der Reflink-Datei FILE1

OVM3 bietet keine direkte Möglichkeit, einen Snapshot zu erstellen, und später wieder zu diesem Snapshot zurückzukehren. Wer sich hingegen auf die Kommandozeile getraut, kann sich diese Funktionalität leicht selbst implementieren. Im Vortrag werden wir anschauen, wie sich das implementieren lässt.

Neben den Snapshots von Disks ist es möglich, auch einen Snapshot einer laufenden VM (also vom Memory-Inhalt) zu erzeugen und später zu restoren.

Hochverfügbare Datenbankverbindung

Die Datenbankverbindung vom OVM Manager zur Datenbank erfolgt über eine jdbc-Verbindung mit dem Thin-driver. Der Zugriff ist nur sehr rudimentär über Host:Port:SID spezifiziert. Mit dieser Konfiguration ist es nicht möglich, das Repository in einer RAC-Umgebung zu haben, da wir hier zwingend über Service-Namen statt SID verbinden müssen. Und eine Data Guard Umgebung ist so auch nicht möglich, da wir zwei Hostnamen spezifizieren müssen. Glücklicherweise nutzt der OVM Manager für Datenbankverbindungen die Möglichkeiten von Weblogic, so dass wir die Verbindungskonfiguration dort korrigieren können. Der verwendete Jdbc Thin-Driver bietet zwar

keine Möglichkeit, abgebrochene Verbindungen wieder transparent aufzubauen, aber wir können die Überwachung und den Neu-Aufbau dem Weblogic-Server delegieren.

Hard-Partitioning

Unter „Hard-Partitioning“ wird die explizite Zuweisung von einer physischen CPU (Core) zu einer Virtuellen Maschine verstanden.

Normalerweise sind bei Oracle Datenbankservern, z.B. bei der Enterprise Edition, sämtliche vorhandenen CPUs zu lizenzieren, im Fall der x86 Architektur sind für 2 CPU-Cores 1 CPU-Lizenz notwendig (Faktor 0.5), auch wenn nur ein Bruchteil der Leistung für Oracle verwendet wird. Dies gilt auch bei der Lizenzierung auf virtualisierten Servern. Massgebend sind die auf dem Host vorhandenen CPU's. Bei einem VMWare Cluster sind sogar sämtliche CPUs im Cluster zu lizenzieren, auf denen eine VM theoretisch laufen könnte, denn Oracle akzeptiert im Fall von VMWare die Zuweisung von CPU's zu den Maschinen nicht als Hard-Partitioning.

Bei Oracle VM wird jedoch die explizite Zuweisung als Hard-Partitioning anerkannt. Man kann also auf einem Server mit 16 Cores z.B. zwei davon explizit einer VM zuordnen und muss nur diese zwei lizenzieren. Diese Möglichkeit ist einer der grossen Vorteile von Oracle VM. Dass man das Hard-Partitioning nicht über die Web-Oberfläche konfigurieren kann, sondern manuell auf dem Server das Konfigurationsfile editieren muss, ist mir jedoch unverständlich.

Im Vortrag werden wir anschauen, wie die Zuweisung eines Cores zu einer VM funktioniert. Das ist zwar grundsätzlich auch im Handbuch beschrieben. Das ist aber nur die Variante, wie die VM random auf diese zugewiesenen Cores zugreift. Es kann aber sinnvoll sein, ein explizites Mapping von physischem Core zu virtueller CPU zu machen.

Kontaktadresse:

Martin Bracher
Trivadis AG
Europa-Strasse 5
CH-8152 Glattbrugg

Telefon: +41 (0) 44-808 70 20
Fax: +41 (0) 44-808 70 21
E-Mail: martin.bracher@trivadis.com
Internet: www.trivadis.com