

Fazit

In einer virtuellen Umgebung kommt man kaum darum herum, Netzwerk-NICs zu bündeln. In diesem Beispiel ist aufgezeigt, wie die Technologien „Link Aggregation“ und IPMP kombiniert werden können. In Zusammenarbeit mit den Kollegen vom Netzwerk-Team wurde mit der oben beschriebenen Konfiguration eine gute Lösung gefunden. Die Systeme sind redundant und mit optimaler Performance am Netzwerk angeschlossen. Auch im Fall von Wartungen im Netzwerk-Bereich können die Systeme ohne Unterbrechung weiterbetrieben werden. Weil „Load Balancing“ implementiert ist, zeigten

die Messungen eine optimale Verteilung der Netzwerk-Last über die Netzwerk-Ports. Auch der Durchsatz entsprach den Erwartungen.

Es lohnt sich, im Vorfeld genug Zeit zu investieren und ein gutes Netzwerk-Design auszuarbeiten. Wichtig sind auch ausführliche Tests der möglichen Varianten.

Literatur und Links

- https://blogs.oracle.com/droux/entry/link_aggregation_vs_ip_multipathing
- <http://docs.oracle.com/cd/E19253-01/816-4554/>
- http://docs.oracle.com/cd/E37707_01/html/E29665/preface.html
- http://www.ieee802.org/3/hssg/public/apr07/frazier_01_0407.pdf

1. <http://standards.ieee.org/findstds/standard/802.1Q-2011.html>

Roman Gächter
Roman.Gaechter@trivadis.com



Ein knappes Jahr ist seit der Veröffentlichung von Oracle VM 3.1 für x86 vergangen. Dies soll Anlass für ein Review und den Überblick über wesentliche Features der aktuellen Version sein.

OVM 3 (x86) – was sich getan hat

Dirk Läderach, Robotron Datenbank-Software GmbH

Während die ersten Versionen (3.0.1 bis 3.1) bei vielen Anwendern nach Tests oder Upgrades für Ablehnung beziehungsweise Schmunzeln bis Verärgerung sorgte, kann mittlerweile von vielen Seiten eine stete, schrittweise positive Resonanz im Umgang mit der Virtualisierung-Lösung beobachtet werden. Im letzten Jahr hat die DOAG eine Liste der aus Sicht ihrer Mitglieder aktuellen Probleme veröffentlicht und auch Oracle diesbezüglich um Stellung gebeten. Viele dieser Themen sorgten auch bei unseren Kunden für eine schleppende Akzeptanz der Lösung. In der Zwischenzeit ist die Version 3.2.2.520 verfügbar und viele der Mängel sind behoben oder es existiert zumindest ein zufriedenstellender Workaround.

Ein paar positive Beispiele

Upgrade-Probleme gehören seit der Version 3.1.1. nahezu der Vergangenheit an und mittlerweile gibt es auch ein funktionierendes Rollback. Bei fünfzehn vom Autor selbst durchgeführten Migrationen auf 3.2.1 und 3.2.2 hat eine einzi-

ge nicht funktioniert; diese wurde sauber wieder zurückgerollt und es erfolgte ein „redeploy“ der Anwendung.

Einbinden von ISO-Dateien funktioniert zwar weiterhin nur über den Umweg des OVM-Hosts, aber die Import-Limitierung (http, ftp) ist inzwischen einem funktionierenden Repository-Import gewichen, der über eine Aktualisierung des Repository gestartet wird. Diese Funktionalität ist übrigens auch in der Lage, per „scp“ in das Repository kopierte Objekte wie virtuelle Disks, Templates, Assemblies bis hin zu den Konfigurationsdateien der VMs zu importieren (wenn die Syntax stimmt) und über die Oberfläche bereitzustellen.

Das oft genutzte Feature der Hard-Partitionierung mittels CPU-Pinning erforderte anfangs noch manuelles Editieren der VM-Konfigurationsdatei („vm.cfg“). Ebenfalls genutzt werden konnten die OVM-Utills. Im aktuellen Release besteht nun mit den CPU-Pools eine sehr gelungene, automatisierte Variante, das CPU-Pinning für die jeweilige Lizenzierung und/oder die Performance-Ansprüche umzusetzen. Für die Oracle Database Appliance ist diese Methode Pflicht.

Auch die Integration in den Enterprise Manager Cloud Control, der Voraussetzung für eine rollenbasierte Nutzerverwaltung ist, und das OPS-Center

```
[root@ovm01 /]# dmidecode|grep UUID
UUID: 20BDCCAA-D378-4C3E-B968-74AB09200A4E

/etc/ovs-agent/agent.ini:
...
[server]
fakeuuid=20BDCCAA-D378-4C3E-B968-74AB09200A4E
...
```

Listing 1

funktioniert in der Version 3.2.2 nahezu reibungslos. Allerdings hätte man sich einen Hinweis im Upgrade-Guide oder in den Release-Notes darüber gewünscht, dass bei Verwendung des Cloud Control 12c R1 einiges zu tun ist, wenn das Plug-in für den OVM Manager nach dem Update auf 3.2.2.520 weiterhin funktionieren soll.

Was noch fehlt oder noch nicht zufriedenstellend umgesetzt ist

Man sollte niemals versuchen, einen Clustered-Server-Pool mit nur einem Host zu betreiben. Fällt dieser wider Erwarten einmal aus und wird komplett (oder nur das Mainboard) ausgetauscht, kann man sich schon mal eine geeignete Strategie zur erneuten Inbetriebnahme des Server-Pools überlegen. Was einfach klingt und beim Wettbewerber VMware relativ unproblematisch verläuft, kann sich auch in der aktuellsten Version zu einem „Trial & Error“-Szenario entwickeln. Folgende Strategie hat sich in der Praxis bewährt.

Bei Verwendung der vorherigen Server-Namen und IP-Adressen sowie in der Hoffnung, dass der FC-HBA wiederverwendet werden kann, ist es erfolgversprechend, den Host zu „discover“ und die Konfigurationsdatei des Agent („/etc/ovs-agent/agent.ini“) mit der über „dmidecode“ ermittelten „UUID“ zu bearbeiten (siehe Listing 1). Wenn das nicht funktioniert, muss die lokale RPM-DB neu erstellt werden (siehe Listing 2).

```
cd /var/lib
rm __db* && rpm -rebuilddb
reboot
```

Listing 2

Darüber hinaus konnten gute Erfahrungen mit dem Storage-Plug-in von Fujitsu gesammelt werden. Dieses funktionierte praktisch sehr gut und konnte die gestellten Aufgaben wie Anlegen von LUNs, Volumes, Host-Affinity-Groups sowie Snapshots mittels der OVM-Manager-Oberfläche fehlerfrei erledigen. Lediglich der Snapshot von virtuellen Disks mit Thin Provisioning funktionierte nachweislich nicht, hinterher war immer das „root“-Dateisystem des Klons defekt.

Wer bereits mit OVM 2.x gearbeitet hat, kennt die Rolle des Utility-Servers, der in Version 3.1.1 wieder eingeführt wurde. Dadurch ist es möglich, einzelnen OVM-Hosts explizit die Rollen „VM Server“ und „Utility Server“ zuzuweisen. Server mit der Rolle des „Utility Server“ werden für I/O-intensive Operationen (wie das Importieren von Templates) bevorzugt verwendet.

Einige Verbesserungen hat auch die Oberfläche erfahren. So ist es nun beispielsweise möglich, durch Mehrfachauswahl viele VMs auf einmal zu starten oder zu stoppen. Leider hat dieses Feature die Live-Migration noch nicht erreicht. Hier kann weiterhin nur sequenziell gearbeitet werden, beim Wettbewerber VMware ist die Migra-

tion von bis zu acht VMs gleichzeitig möglich.

Die OVM-Shell

Mit der Version 3.2.1 hielt die finale Version des OVM Command Line Interface (CLI) sowie der Shell Einzug. Damit ist nun endlich eine gut funktionierende Schnittstelle (API) für Skripte etc. vorhanden. Diese sollte aus Sicht des Autors stetig erweitert werden, um in Zukunft vielleicht den einen oder anderen Hersteller von Backup-Software zu animieren, diese Schnittstelle in der eigenen Lösung zu nutzen. Aktuell gibt es noch keine Funktionalität eines „Snapshot mit Redo-Log-Funktionalität“, wie es etwa VMware bietet. Dennoch kann mit der CLI die Backup/Restore-Funktionalität mit eigenen Skripten einigermaßen umgesetzt werden.

Snapshots eignen sich nur bei VMs, bei denen die Anwendungen ihre Daten nicht primär im Hauptspeicher vorhalten und bearbeiten. Besonders Application- und File-Server lassen sich damit gut und vor allem konsistent sichern. Bei Oracle-Datenbanken kann durch einen Snapshot allein kein konsistentes Backup erzielt werden, dies ist als alleiniges Sicherungskonzept daher völlig ungeeignet.

Beim Backup mit der OVM-Shell wird bei Verwendung des „OCFS2“-Dateisystems ein „Reflink“-Klon aller Disks der VM erzeugt. Die Verwendung der CLI hat zu den bisher gezeigten Varianten mittels „Reflink“ (siehe DOAG

```
OVM> clone Vm name=oe16_wls01 destType=VmTemplate destName=oe16_wls01_backup serverPool=prod_pool01
Command: clone Vm name=oe16_wls01 destType=VmTemplate destName=oe16_wls01_backup serverPool=prod_pool01
Status: Success
```

Listing 3

```
OVM> importVirtualDisk repository name=backup server=ovm01 url='http://nfs-server/ovmbackup/oe1_backup01.img'
Command: importVirtualDisk repository name=ovm_repo1tb server=ovm01 url='http://nfs-server/ovmbackup/oe1_backup01.img'
Status: Success
Time: 2013-04-15 15:51:12.512
OVM> create VM name=oe16_revover repository=Repo01 domainType=XEN_PVM memory=1024 on Server name=ovm01
Command: create VM name=oe16_revover repository=Repo01 domainType=XEN_PVM memory=1024 on Server name=ovm01
Status: Success
Time: 2013-05-15 16:04:06.071
OVM> create vmDiskMapping name=recoverMap1 slot=1 storageDevice=oe1_backup01.img on vm name=oe16_recover
Command: create vmDiskMapping name=recoverMap1 slot=1 storageDevice=oe1_backup01.img on vm name=oe16_recover
Status: Success
```

Listing 4

News, Ausgabe 03/2012) den entscheidenden Vorteil, dass lediglich der Serverpool und der VM-Name bekannt sein müssen. Die bisherigen Backup-Implementierungen mit „Reflink“-Technologie im Monitoring- und Backup-Tool „robotron*DBAcheck“ mussten noch mühsam feststellen, auf welchem Host die VM läuft und ob es sich um Shared Disks handelt. Da die CLI auf dem Manager läuft, kann natürlich auch auf alle Informationen des Manager zugegriffen werden (siehe Listing 3).

Wenn das Repository per NFS zugänglich gemacht wird, kann die Konfigurationsdatei der zu sichernden VM im Template-Verzeichnis lokalisiert werden. Nun gilt es, die zugehörigen Disks zu ermitteln. Das kann sehr einfach mit „grep“, „sed“, „cut“, „awk“ etc. erledigt werden. Dabei ist der originale Pfad durch den NFS-Mount-Punkt mittels „/OVS/Repositories/id/VirtualDisks/ -> /nfs/archiv/ovm-backup/VirtualDisks/“ zu ersetzen. Das Ergebnis stellt die Quelle für das Sichern der virtuellen Disks auf ein externes Medium dar. Darüber hinaus kann die Konfigurationsdatei sehr leicht im Template-

Verzeichnis gesichert werden. Das ist hauptsächlich für die MAC-Adresse interessant.

Beim Restore gibt es verschiedene Varianten. Die virtuellen Disks müssen auf jeden Fall wieder in das Repository und dort muss eine VM entstehen. Eine Möglichkeit zum Instant Recovery (Start der VM aus dem Backup-Verzeichnis) gibt es momentan noch nicht. Dieses Skript kann sehr einfach implementiert werden und erstellt aus der Sicherung eine neue VM (siehe Listing 4).

Die virtuelle Netzwerkkarte kann ebenfalls über die CLI zugewiesen werden. Falls die originale MAC-Adresse notwendig ist, kann diese ebenfalls mithilfe von „find . -type f -exec sed -ie ,s#alter Wert#neuer Wert#g' {} \;“ über das Restore-Skript ausgetauscht werden.

Fazit

Oracle VM ist schrittweise immer besser geworden und kann inzwischen auch als gut benutzbar betrachtet werden. Wer sich an ein paar Regeln hält, wie keinen Stand-Alone-Serverpool zu betreiben oder nicht zu exzessiv mit VLANs und PXE-Boot zu arbei-

ten, kann durchaus auch über einen Einsatz im Rechenzentrum nachdenken. Allerdings gibt es noch keine Lösung, um eine Windows-Installation zu Oracle VM zu portieren. Hier kann weiterhin nur der VMware-Konverter genutzt werden, der die Installation nach OVF portiert, das dann vom Manager wiederum in Form von Assemblies verstanden wird. Ironischerweise gibt Oracle das mittlerweile auch selbst als Workaround heraus.

Die OVM Shell hat ebenfalls noch Potenzial. Hier würde man sich beispielsweise ein alternatives Zielverzeichnis für VM-Kopien wünschen. Die VM-Messages bieten viele Möglichkeiten bei der Konfiguration und Kommunikation mit VMs ohne direkten Zugriff auf die VMs selbst.

Dirk Läderach
dirk.laederach
@robotron.de



www.dba-im-urlaub.de

MUNIQSOFT
Datenbanken mit iQ