

Oracle VM 3.0 – alles neu?

Martin Bracher
Senior Consultant
21. Januar 2012



Dieses Dokument ist das Manuskript zum gleichnamigen DOAG2011 Vortrag.

1. Einleitung

Während wir mit Oracle VM (OVM) Server Version 2 eine schnelle, stabile Virtualisierungslösung hatten, war der (optionale) Administrationsteil in Form des OVM Managers, verglichen mit anderen Virtualisierungslösungen, doch sehr bescheiden. Für einige Kunden, die nicht auf der Kommandozeile arbeiten wollten, war dies ein No-Go. Viele Kunden schätzten aber gerade die Möglichkeit, die Virtualisierungsumgebung vollständig auf der Kommandozeile administrieren zu können.

Oracle liess sich viel Zeit, eine überarbeitete Version herauszugeben. Nachdem sie eigentlich für letztes Jahr angekündigt war, kam sie nun Ende August dieses Jahres endlich raus. Im Vortrag schauen wir uns an, ob sich das lange Warten gelohnt hat. Der Vortrag gibt einen Überblick und erste Erfahrungen mit den neuen Features von OVM 3.

Abgesehen von der nachfolgenden rein technischen Betrachtung sollten wir auch immer die folgenden Vorteile im Auge behalten: OVM ist die einzige zertifizierte Virtualisierungsumgebung in der x86 Prozessorfamilie für Oracle Produkte. Nur bei OVM anerkennt Oracle „Hard-Partitioning“ (d.h. man muss nur die physisch zugewiesenen CPU's für die Datenbank lizenzieren). Und OVM ist Lizenzkostenfrei: Man bezahlt nur für Support (optional).

2. Architektur

Oracle VM basiert auf Oracle/RedHat Enterprise Linux und verwendet für die Virtualisierung auch weiterhin XEN (RedHat setzt ab Version 6 auf KVM als Virtualisierungslösung), jedoch erfolgt jetzt der Wechsel von XEN3 auf XEN4. Ein grosser Vorteil von XEN ist die sogenannte Paravirtualisierung (PVM). Statt Hardware zu emulieren, kann eine Virtuelle Maschine (VM, DomU) mit entsprechend angepasstem Gast-Kernel die Ressourcen des Hosts (Dom0) direkt ansprechen und hat so kaum Performance-Verluste. XEN unterstützt aber auch Hardware-Virtualisierung (HVM), erforderlich etwa für Windows VM's. Wie bei anderen Virtualisierungslösungen auch, existieren für HVM paravirtualisierte Treiber, um zumindest die Performance-kritischste Hardware (Disk, Netz) effizienter nutzen zu können.

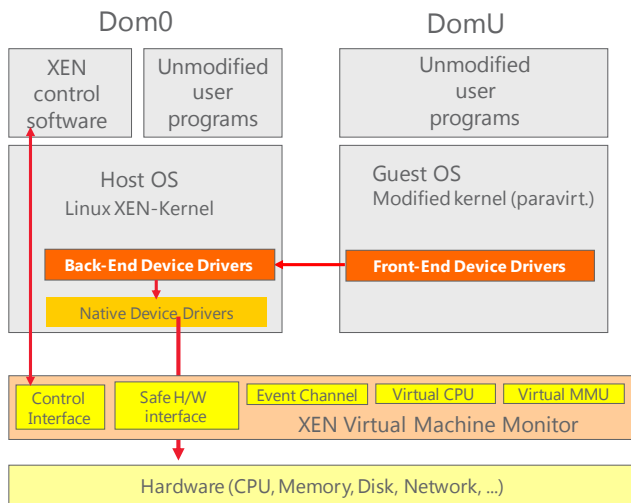


Abb. 1: XEN-Architektur (paravirtualisiert)

OVM sind eigentlich zwei Komponenten: Einerseits der OVM Server, andererseits der OVM Manager.

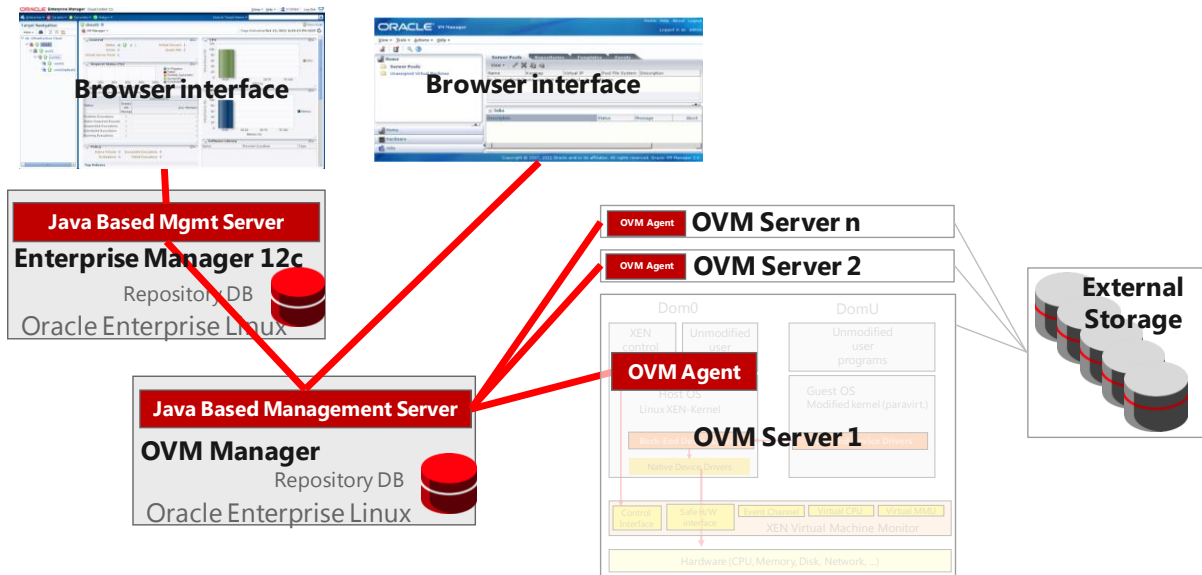


Abb. 2: Architektur OVM Manager

2.1 OVM Server 3.0

Der OVM Server ist eine minimale Linux-Distribution (ISO-Image ca. 200MB), basierend auf Oracle/RedHat Enterprise Linux (OEL), ergänzt mit XEN 4, um als Host für die VM's genutzt werden zu können. Dieser Server kann vollständig auf der Kommandozeile administriert werden. Wer dies schon mit OVM2 so gemacht hat, kann problemlos weiterhin so arbeiten. Die bestehenden VM's sind ohne Anpassungen weiterhin lauffähig. Wer jedoch einen hochverfügbaren Cluster aufbauen will, muss mit dem OVM Manager arbeiten, oder die Hochverfügbarkeits-Funktionalität selbst implementieren.



2.2 OVM Manager 3.0

Der OVM Manager dient zum Administrieren, Steuern und Überwachen mehrerer OVM Server. Er wird in ein bestehendes Oracle/RedHat Enterprise Linux System installiert. Wie viele andere Oracle Produkte auch, ist der Manager nun in Weblogic integriert und besitzt ein Repository, welches sich in einer Oracle Datenbank befindet. Bei der Verwaltung der Server hat sich gegenüber OVM2 konzeptionell kaum etwas geändert: Man verwaltet weiterhin VM Templates, CD-Images, shared Disks, lauffähige VM's usw., aber die technische Implementierung hat sich vollständig geändert, was die Migration problematisch macht.

Enterprise Manager 12c

Die OVM Server können auch über den neuen Enterprise Manager 12c verwaltet werden. Dabei wird eine bestehende OVM Manager Installation eingebunden, d.h. der OVM Manager ist also weiterhin notwendig. EM 12c bietet gute Möglichkeiten, den Ressourcenverbrauch über die Zeit hinweg zu betrachten. Ansonsten ist EM 12c eigentlich nur ein anderes Frontend zum OVM Manager.

3. OVM3 – alles neu?

3.1 Setup Server

Was ist denn nun neu? Beim Server sind die Änderungen auf den ersten Blick nicht riesig.

Bei der Installation erscheint nun ein Menüpunkt „Upgrade“, der aber leider nicht in der Lage ist, eine bestehende OVM2 Umgebung zu upgraden. Dies ist nur innerhalb von Version 3 möglich und funktioniert auch einwandfrei von 3.0.1 auf 3.0.2.

Nachdem die Installation abgeschlossen ist, fällt auf, dass auf der ersten Konsole der Status vom Server angezeigt wird. Zum Login muss man auf die zweite Konsole wechseln, auf der man sich aber wieder zuhause fühlt.

Die Installation erstellt nicht mehr wie früher ein ocfs2-Filesystem, welches unter /OVS gemounted wurde.

Im Netzwerkbereich setzt Oracle nun auf hohe Verfügbarkeit. Als Netzwerk kommen Bonding-Devices zum Einsatz: Dabei können mehrere Netzwerkkarten zu einer logischen Karte (bond0) zusammengefasst werden. Initial enthält die Konfiguration von bond0 jedoch nur eine Netzwerkkarte, was wir aber einfach erweitern können.

Falls wir schon OVM2 genutzt haben und den OVM Manager nicht verwenden wollen, können wir das Volume mit den bestehenden VM's wieder unter /OVS mounten und die VM's wie gewohnt starten. Wer jedoch mit einer VNC Konsole arbeitete stellt fest, dass der vncviewer leider fehlt. Das ist wohl einer der Gründe, wieso das Installations-ISO von 430 auf 197MB geschrumpft ist. Und gleich eine Warnung: ein OEL5 Repository einzubinden und vnc von dort zu installieren, quittierte der OVM Manager nach einem Reboot des Servers mit der Weigerung, weiterhin mit diesem Server zusammen zu arbeiten...

Einige weitere Änderungen sind nur für jene von Interesse, die auch entsprechende Hardware zur Verfügung haben. OVM3 kann nun bis zu 2 TB Memory und 160 CPU's verwalten.



3.2 Setup Manager

Nachdem vorallem der OVM2 Manager in der Kritik stand, hat Oracle diesen technisch nun völlig überarbeitet. Die Organisations-Architektur (ISO's, Shared Disks, Templates, VM's) ist weitgehend erhalten geblieben. Ebenso ist das Frontend eine Web-Applikation geblieben, d.h. es bleibt uns erspart, auf den Administrator-PC's einen Fat-Client installieren zu müssen und sich dabei vorschreiben lassen zu müssen, was wir als Betriebssystem dafür zu verwenden haben. Ansonsten ist nicht mehr viel gleich geblieben.

Das Installations-ISO ist von 585MB auf 2.4GB gewachsen und muss in ein bestehendes OEL/RHEL System mit mindestens Version 5.5 installiert werden. Für das Repository kann entweder eine eigene Datenbank, oder aber die mitgelieferte Oracle XE Version verwendet werden.

Nach der Installation (Login als User „admin“, eine weitergehende Benutzerverwaltung lässt sich nur über die Kommandozeile konfigurieren...) müssen nun mal die VM Server und der Storage konfiguriert werden. Diese Schritte sind nicht unbedingt selbst erklärend, ein Blick in die Dokumentation ist sehr hilfreich.

Zuerst müssen die OVM Server erkannt werden. Diese Funktionalität ist nicht auf der Einstiegsseite, sondern auf der Hardware-Seite versteckt.

Anschliessend ist der Storage zu registrieren. Als Storage werden nur ganze Disks/LUN's, also keine Partitions, akzeptiert. Alternativ kann auch NFS-Storage eingebunden werden. NFS-Shares werden nur erkannt, wenn noch mindestens 12GB Platz frei ist. Ebenso muss der OVM Server Schreibrechte darauf haben (standardmässig wird bei NFS von root ein Mapping auf nobody gemacht). Nach dem Zuweisen des Storage zu den Servern muss noch ein „Refresh File System“ gemacht werden, damit es funktioniert.

Nun ist das Netzwerk zu konfigurieren. Hierbei stehen Konfigurationen zur Verfügung mit Bonding und VLAN. Dies konnte man auch schon früher konfigurieren, aber nicht im Manager. Anschliessend können wir virtuelle Netzwerkkarten definieren. Wir geben einen Range von MAC-Adressen vor, und lassen danach eine Menge von virtuellen Interfaces definieren, welche wir danach den virtuellen Maschinen zuweisen können. Im Gegensatz zur alten Version haben wir jetzt hier gute Kontrolle über die vergebenen Interfaces. Wir können in unserem dhcp-Server den MAC's eine konstante IP und im DNS einen passenden Namen zuweisen.

Nun ist es an der Zeit, den Serverpool zu definieren. Im Gegensatz zur alten Version geschieht dies am Schluss, nachdem wir Storage und Server definiert haben.

Dem Pool muss ein sogenanntes Pool File System zugefügt werden. Bei SAN und ISCSI wird direkt auf die LUN ein OCFS2-Filesystem gelegt. Bei Verwendung von NFS Storage wird ein File von etwa 10GB darauf abgelegt. Das File ist ein Disk-Image mit einem OCFS2-Filesystem, das anschliessend per Loopback auf den OVM-Servern gemountet wird.

```
watt:/OVS/ovs3 on /nfsmnt/6126e965-5644-44c0-b710-92c907d3013c type nfs (rw,addr=192.168.2.12)
-rw----- 1 65534 65534 10G Sep 12 2011 ovspoolfs.img
/dev/mapper/ovspoolfs on /poolfsmnt/0004fb000005000008ab46d2e9bbd60b6 type ocfs2
(rw,_netdev,heartbeat=global)
```

Erst jetzt sollen die zuvor erkannten Server dem Pool zugewiesen werden.

Um nun virtuelle Maschinen speichern zu können, müssen sogenannte Storage Repositories zugefügt werden. Als Repository-Storage wird eine zuvor erkannte Disk/LUN oder einer der NFS-Shares verwendet. Wer jetzt denkt, das war's, der hat sich getäuscht. Zuerst muss man diese Repositories noch den Servern präsentieren, sonst sind sie da nicht verfügbar. Dass dies nicht automatisch geschieht, oder zumindest automatisch danach gefragt wird, ist verwirrend und ärgerlich.



Und wie kommen wir nun zu einer virtuellen Maschine? Entweder wir erstellen eine aus einem Template (Ein Template ist eine etwas spezielle VM, welche nicht als VM, sondern nur zum Klonen einer neuen VM verwendet wird). Oder wir installieren eine neue Maschine ab den Installationsmedien (DVD-Image). Vor Verwendung müssen wir die Templates zuerst via OVM Manager importieren. Und hier kann man, diplomatisch ausgedrückt, die Implementierung als „wenig gelungen“ bezeichnen. Als Quelle für den Import lässt sich nur http(s) oder ftp angeben. Falls sich das Template bereits auf dem Server auf Disk befindet, lässt sich dies nicht direkt von Disk lesen. Möglicher Workaround ist, einen ftp-Server zu installieren (Vorsicht: hinkopieren, nicht als RPM installieren; siehe Bemerkung wegen zusätzlicher Software oben) und dann über diesen die Software zu lesen (benötigt dann temporär den doppelten Platz). Noch weniger gelungen ist es, wenn man bestehende VM's von einer OVM2 Installation übernehmen will. Diese wären an sich ohne Änderungen auch auf einem OVM3 Server lauffähig, nur ist der OVM Manager nicht in der Lage, diese in sein Repository aufzunehmen. Einzige Möglichkeit, die ich aktuell sehe: Die VM in ein tar.gz File packen, dieses wie oben beschrieben über den ftp-Server herunterladen und als Template importieren und danach aus diesem Template wieder eine VM zu klonen. Benutzerfreundlich ist anders...

Wer jetzt denkt, er könne passende Einträge direkt im Repository vornehmen (wie man das beim alten OVM noch leicht machen konnte), den muss ich enttäuschen. Praktisch alle Tabellen bestehen aus 2 Feldern. Einer ID, und einem BLOB. Wieso Oracle die Informationen nicht in normal lesbaren Datentypen speichert, bleibt mir ein Rätsel.

3.3 Was der Manager auf dem Server einrichtet

Was ist nun das Ergebnis auf dem OVM Server? Gegenüber dem alten OVM hat die Verzeichnisstruktur geändert. Die Unterverzeichnisse mit VM's, Templates usw. liegen nicht mehr direkt unter /OVS, sondern unter /OVS werden zuerst die Repositories gemounted: /OVS/Repositories/0004fb000003000000d15c102379bfaa

Auch die darunter liegende Directory-Struktur hat geändert. Statt running_pool, seed_pool, iso_pool usw, hat man jetzt in jedem Repository VirtualMachines, VirtualDisk, ISOs usw. Es handelt sich nicht bloss um eine Umbenennung der Verzeichnisse, auch das Konzept ist anders. Während z.B. in running_pool und seed_pool Unterverzeichnisse für die VM's existierten, welche die Disk-Images sowie das Konfigurationsfile enthielten, werden neu die Konfigurationsfiles in VirtualMachines und die Diskfiles in VirtualDisks gespeichert. Es gibt dabei kein Unterverzeichnis mehr mit dem Namen der VM, sondern alle Files liegen auf gleicher Ebene und haben einen UUID Namen, der nicht mehr auf die Zugehörigkeit oder Verwendung schliessen lässt.

3.4 Green IT

Konfiguriert über den Manager, lässt sich ein geclusterter Serverpool ökologischer betreiben. Ein Ziel der Virtualisierung ist, die Hardware besser zu nutzen, also so viele VM's auf einer Maschine laufen zu lassen, dass die Ressourcen ausgelastet sind. In der Praxis schwankt der Leistungsbedarf. OVM3 soll nun in der Lage sein, wenn ein Server am Limit ist, und ein anderer freie Ressourcen hat, eine VM im laufenden Betrieb auf einen anderen Server zu verschieben. Falls die Summe der nicht benötigten CPU-Ressourcen grösser ist, als die Ressourcen eines Servers, kann OVM die VM's von einem Server auf die anderen verteilen und diesen stoppen um Strom zu sparen. Wenn der Leistungsbedarf steigt, kann dieser Server wieder gestartet werden.



4. Fazit

Während der OVM Server überzeugt, macht der OVM Manager einen sehr zwiespältigen Eindruck. Obwohl das Produkt mit massiver Verspätung veröffentlicht wurde, hinterlässt es an vielen Stellen einen unfertigen Eindruck, wie beispielsweise die fehlende Benutzerverwaltung, oder die fehlende Möglichkeit, bestehende VM's ab Disk importieren zu können. Auf der anderen Seite steht eine ansprechende Web-Oberfläche und gute neue Konfigurationsmöglichkeiten wie etwa im Netzwerkbereich mit Bonding und VLAN.

Man merkt dem Produkt an, dass es vor allem für grosse Umgebungen konzipiert wurde. Beispielsweise wenn man eine Umgebung mit 2 Servern hat, fragt man sich, wieso beim Hinzufügen eines Servers die Pool-Filesysteme nicht automatisch hinzugefügt werden. Bei 20 Servern möchte man es aber vielleicht nicht mehr überall automatisch präsentieren. Aber eine automatisch erscheinende Abfrage, ob man auch noch Filesysteme präsentieren möchte, wäre doch ganz angenehm.

Es ist davon auszugehen, dass Oracle in einer nächsten Version viele der kritisierten Punkte noch verbessern wird. Wer jetzt schon OVM2 einsetzt, kann mit der Migration noch warten, wer neu beginnen will, sollte jetzt schon mit der neuen Version starten.

Viel Erfolg beim Einsatz von Trivadis-Know-how wünscht Ihnen

Martin Bracher

Trivadis AG

Europa-Strasse 5

CH-8152 Glattbrugg

Internet: www.trivadis.com

Tel: +41-31 928 09 60

Fax: +41-31 928 09 64

Mail: info@trivadis.com