

Oracle VM: Tipps und Tricks aus dem Alltag

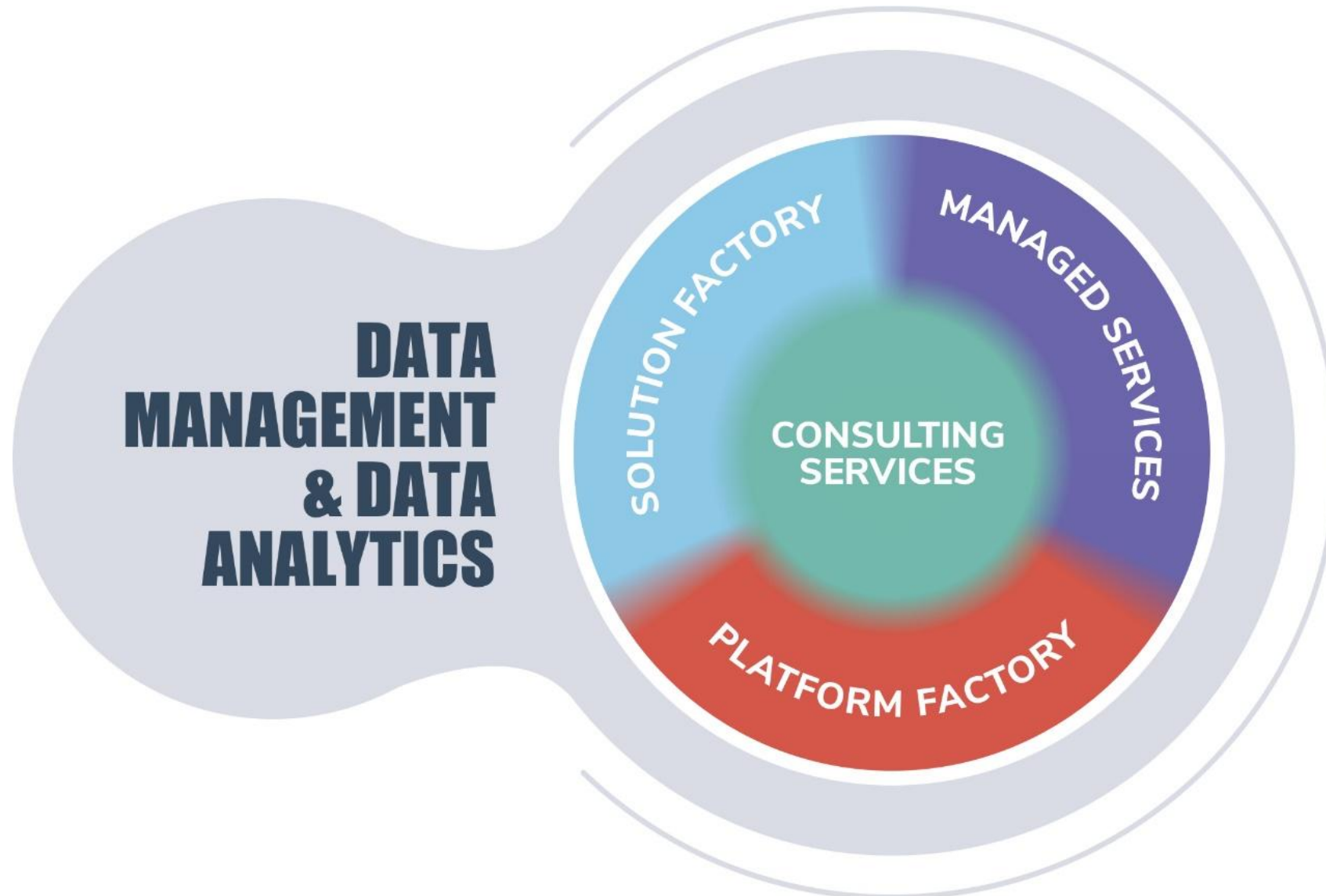
Martin Bracher



#



We help to generate added value from data



With over 650 specialists and IT experts in your region.



- 16 Trivadis branches and more than 650 employees
- Experience from more than 1,900 projects per year at over 800 customers
- 250 Service Level Agreements
- Over 4,000 training participants
- Research and development budget: CHF 5.0 million
- Financially self-supporting and sustainably profitable

AGENDA

1. Using the latest hardware
2. Server with local SDCARD
3. BIOS
4. Upgrade of OVM server
5. Corrupt filesystem
6. Migration from VMWare to OVM

Using the latest hardware

Case: customer ordered the latest hardware

- Booting from virtual cdrom (iso image)
 - Choose the disk to install... but no disks were detected
 - Server had 2 disks with a hardware raid controller
 - We disabled the hardware raid, but the disks are still not recognized
 - The latest version of ovm server was older than the physical hardware
 - The disk- and raid-controller was not yet supported by the kernel
 - Verifying this thesis by booting the latest EL7 image which was a little bit newer. It recognized the disk
 - Remastering the iso image with a new kernel
 - Not so easy if you never have done so...

Using the latest hardware

■ Driver Update Disk

■ Good luck that we have seen a Driver Update Disk

- The Driver Update Disk provides a mechanism for delivering updated device drivers during system installation. This allows installing Oracle VM on certain systems that are not supported by the latest Oracle VM installation media. The installer loads the updated device driver kernel modules (kmods) from the DUD so it can access the devices, and it also installs the RPM packages containing the device drivers on the target system.
- Put it on an USB drive (usb stick) and attach it to the server, and reboot it together with the ISO
 - Unfortunately the server only supports 1 virtual drive
 - We had to physically go to the server room and plug it in 😞
 - But it was successful 😊



<input checked="" type="checkbox"/>	Oracle VM Server for x86 3.4.4 for x86 64 bit	
<input checked="" type="checkbox"/>	V952528-01.iso	Oracle VM Server 3.4.4 Source for x86_64 (64 bit), 1.3 GB
<input checked="" type="checkbox"/>	V952527-01.iso	Oracle VM Server 3.4.4 for x86_64 (64 bit), 844.0 MB
<input checked="" type="checkbox"/>	V974200-01.iso	Oracle VM Server for x86 3.4.4 Driver Update Disk, 1.8 MB
<input checked="" type="checkbox"/>	V974201-01.zip	Oracle VM Server for x86 3.4.4 Driver Update Disk Readme, 1.1 KB

Using the latest hardware

- Server setup completed successfully, but...
 - Not all network cards were detected
 - One card was correctly recognized (this one we used for the admin interface)
 - But the other one was not visible
 - The module for it was not on the driver update disk
 - Download and install the latest Kernel from public-yum.oracle.com
 - If the Kernel in the ovm3 repo is still too old, have a look in the OL6 repo (OVM3.4 is based on OL6)
 - OL and OVM should have the same kernel
 - After a reboot the card was recognized 😊
 - Now the configuration via OVM Manager could start

AGENDA

1. Using the latest hardware
- 2. Server with local SDCARD**
3. BIOS
4. Upgrade of OVM server
5. Corrupt filesystem
6. Migration from VMWare to OVM

Server with local sdcard (no harddisk)

Case: Cisco blade system, installing on 32GB sdcards (RAID)

- no local harddisk, no booting from SAN
- Installation begins as usual
 - Choosing the boot device (the only device is the sdcard)
 - Afterwards installation aborts with an error about formatting
 - Switch to textmode console: [Ctrl]+[Alt]+[F2]
 - Try to format it manually: `mkfs.ext4 /dev/sdd1`
 - gives `"/dev/sdd1 is apparently in use by the system"`
- Hmm, what now?

Server with local sdcard (no harddisk)

- Hmmm, what now?

- The device was in use before with another OS

-  Erasing the device with "dd" → did not help

-  Maybe the card is broken. Replace it with a new one → did not help

- Back to the message "in use by the system". Who is using it?

```
multipath -ll #look for device "sdd", multipath device is e.g. mpatha
ls -l /dev/mapper/mpatha #is a symlink to /dev/dm-0
```

- /dev/sdd was in use by the device mapper and ovm tried to format sdd instead of dm-0
- Re-start installation, before formatting the drive, switch to console [Ctrl]+[Alt]+[F2]

```
dmsetup remove /dev/dm-0
```

- Switch back to installer [Ctrl]+[Alt]+[F1] and continue

AGENDA

1. Using the latest hardware
2. Server with local SDCARD
- 3. BIOS**
4. Upgrade of OVM server
5. Corrupt filesystem
6. Migration from VMWare to OVM

BIOS

Case: re-install a server

- System installed, but after reboot the bootdevice is not found
 - Probably it was installed in "legacy" boot mode
 - The installer used the "uefi/gpt" boot mode
 - Solution
 - Change the server to "uefi" boot mode
- Background of the problem
 - The old OVM version was unable to boot with "uefi"
 - So we disabled uefi in the bios and activated legacy boot

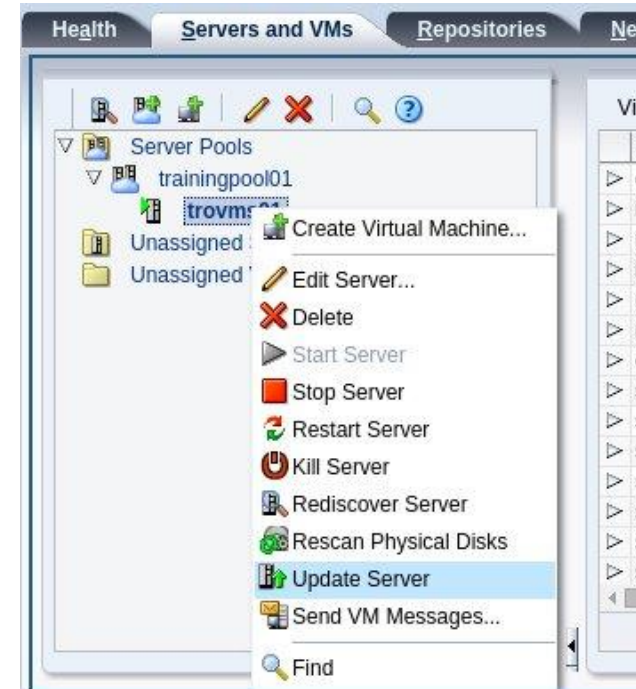
AGENDA

1. Using the latest hardware
2. Server with local SDCARD
3. BIOS
- 4. Upgrade of OVM server**
5. Corrupt filesystem
6. Migration from VMWare to OVM

Upgrade of OVM server

Case: Upgrade started from the manager

- Select the server and run "Update Server"
 - Puts it in maintenance mode
 - Runs "yum upgrade" on the server
 - Reboots the server
- Result
 - One server upgraded perfectly
 - The other server did not startup
 - After booting the kernel, it stops



Upgrade of OVM server

■ Analysis

- Reboot the server and choose the former Kernel
- Compare the /boot directory of the 2 servers

```
root      1024 Jun 22  2016 grub.tmp
root 39045032 Sep 11 14:16 initramfs-4.1.12-124.19.1.el6uek.x86_64.img
root 48571942 Jun 22  2016 initramfs-4.1.12-37.4.1.el6uek.x86_64.img
root 46623115 Sep 11 14:15 initramfs-4.1.12-94.3.8.el6uek.x86_64.img
```

```
root      1024 Jun 22  2016 grub.tmp
root 48526185 Jun 22  2016 initramfs-4.1.12-37.4.1.el6uek.x86_64.img
root 46253431 Jul 11  2017 initramfs-4.1.12-94.3.8.el6uek.x86_64.img
root      12288 Feb 16  2016 lost+found
```

- On one server the initramfs was missing! Try to re-create it

```
# dracut -f initramfs-4.1.12-124.19.1.el6uek.x86_64.img 4.1.12-124.19.1.el6uek.x86_64
E: Failed to install /etc/system-fips
```

- /etc/system-fips (an empty file) was missing

■ Solution

```
# touch /etc/system-fips
# yum reinstall kernel-uek-4.1.12-124.19.1.el6uek.x86_64
# init 6
```

AGENDA

1. Using the latest hardware
2. Server with local SDCARD
3. BIOS
4. Upgrade of OVM server
- 5. Corrupt filesystem**
6. Migration from VMWare to OVM

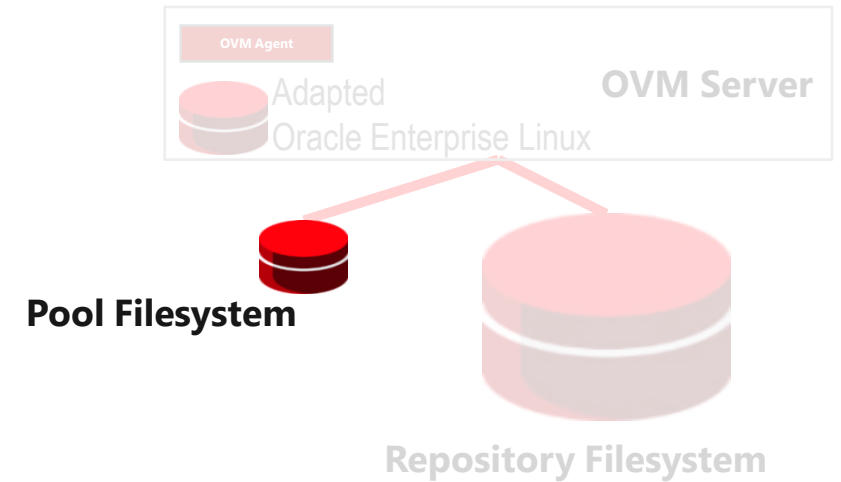
Corrupt filesystem

Corrupt or deleted OCFS2 filesystem (Storage-Repo, Pool-Filesystem)

- On a corrupt filesystem, first try to copy as much as possible to another device
- You can re-create the filesystem, if
 - the WWID of the LUN is unchanged
 - you have a backup of the contents

Re-create a corrupt Pool Filesystem

- Pre-requisite
 - Backup the contents of the filesystem regularly (with tar/cp/...)
- Re-create the filesystem on the same LUN/WWID



```
mkfs.ocfs2 -b 4096 -C 4096 -J size=4M -N 32 -L OVS_POOL_FILESYSTEM \  
-U 0004FB0000050000E6A66CB83CA0C328 --global-heartbeat \  
--cluster-stack=o2cb \  
--cluster-name=e778eb160b9eab47 \  
--force /dev/mapper/360002ac0000000000000000760000d140 ;
```

- The highlighted information can be found in the .ovspoolfs file (see pre-requisite 😊)

```
# cat /poolfsmnt/0004fb0000050000e6a66cb83ca0c328/.ovspoolfs  
OVS_POOLFS_UUID=0004fb0000050000e6a66cb83ca0c328  
OVS_POOLFS_MGR_UUID=0004fb000001000079a86253a43b91fb  
OVS_POOLFS_VERSION=3.0  
OVS_POOLFS_POOL_UUID=0004fb0000020000e778eb160b9eab47  
OVS_POOLFS_LUN_UUID=360002ac0000000000000000760000d140
```

- Afterwards, temporarily mount the device on /mnt and restore the backup, then reboot the server

Pool Filesystem - Dependencies

```
mkfs.ocfs2 -b 4096 -C 4096 -J size=4M -N 32 -L OVS_POOL_FILESYSTEM \  
-U 0004FB0000050000E6A66CB83CA0C328 --global-heartbeat \  
--cluster-stack=o2cb \  
--cluster-name=e778eb160b9eab47 \  
--force /dev/mapper/360002ac0000000000000000760000d140 ;
```

```
# multipath -ll  
360002ac0000000000000000760000d140 dm-2 3PARdata,VV  
size=13G features='1 queue_if_no_path' hwhandler='0' wp=rw ...
```

```
# strings /poolfsmnt/0004fb0000050000e6a66cb83ca0c328/db/server_pool | grep "U "  
U 0004fb0000020000e778eb160b9eab47q
```

```
# df -h |grep poolfsmnt  
/dev/mapper/360002ac0000000000000000760000d140          13G  263M  13G   2%  
/poolfsmnt/0004fb0000050000e6a66cb83ca0c328
```

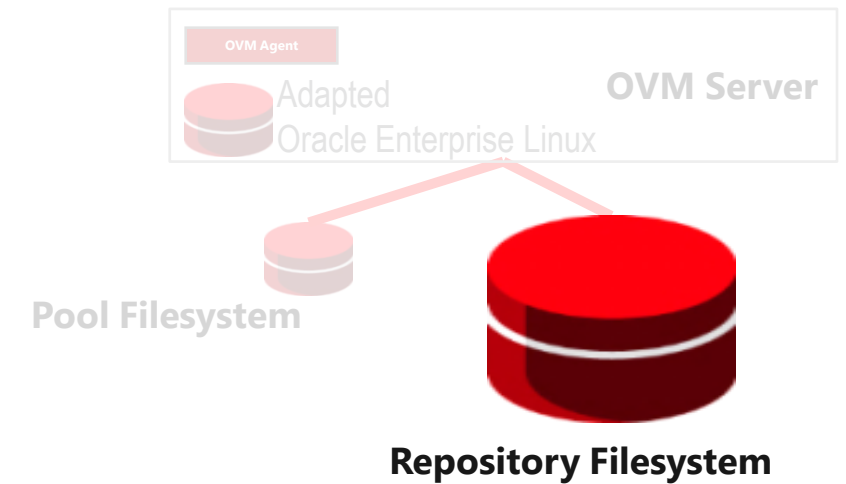
```
# cat /poolfsmnt/0004fb0000050000e6a66cb83ca0c328/.ovspoolfs  
OVS_POOLFS_UUID=0004fb0000050000e6a66cb83ca0c328  
OVS_POOLFS_MGR_UUID=0004fb000001000079a86253a43b91fb  
OVS_POOLFS_VERSION=3.0  
OVS_POOLFS_POOL_UUID=0004fb0000020000e778eb160b9eab47  
OVS_POOLFS_LUN_UUID=360002ac0000000000000000760000d140
```

OVM Server: Repository Filesystem

- Stores diskfiles and definition files of VM's
- Initially backup the whole filesystem
 - .ovsrepo file and directory structure

```
[root@elektra ~]# ls -l /OVS/Repositories/0004fb00000300009be1901acc5c697d
total 0
drwx----- 2 root root 3896 Mar  1 17:16 Assemblies
drwx----- 2 root root 3896 Mar  1 17:16 ISOs
drwx----- 3 root root 3896 Mar  1 17:28 Templates
drwx----- 2 root root 3896 Mar  2 23:22 VirtualDisks
drwx----- 4 root root 3896 Mar  2 23:21 VirtualMachines
drwxr-xr-x  2 root root 3896 Mar  1 17:16 lost+found
```

- Backup the contents regularly
 - Caution: do not backup open files of virtual disks

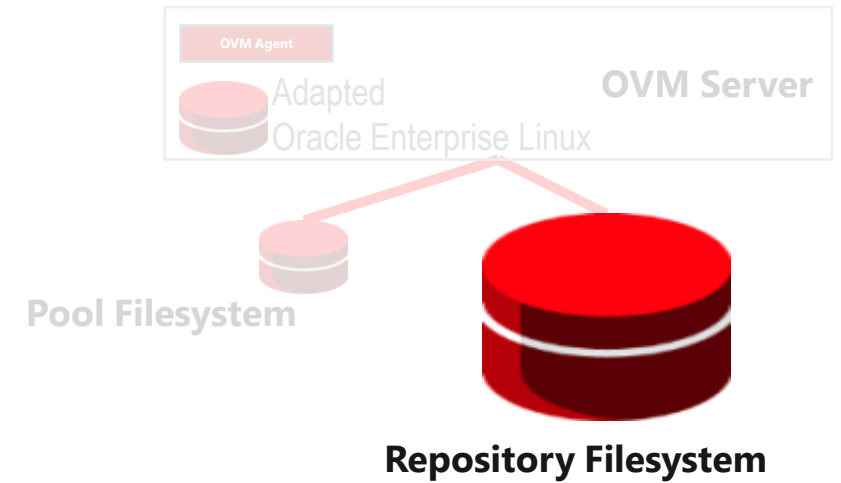


OVM Server: Repository Filesystem

- Re-create the filesystem

```
mkfs.ocfs2 -J block64 -b 4096 \  
-L OVS48c3355bb2934 \  
-U 0004FB0000050000B4C48C3355BB2934 -T vmstore -N 32 \  
/dev/mapper/360002ac000000000000000740000d140;
```

- -U is the ID of the physical disk on the "storage" tab in OVM manager
- Temporarily mount the filesystem and restore the backup
 - Probably you have to restart o2cb before: `service o2cb restart`



Repository Filesystem - Dependencies

```
mkfs.ocfs2 -J block64 -b 4096 \  
-L OVS48c3355bb2934 \  
-U 0004FB0000050000B4C48C3355BB2934 -T vmstore -N 32 /dev/mapper/360002ac000000000000000740000d140;
```

```
# multipath -ll  
360002ac000000000000000740000d140 dm-1 3PARdata,VV  
size=5.0T features='1 queue_if_no_path' hwhandler='0' wp=rw ...
```

```
/etc/ovs-agent/db # strings repository  
mount_pointq  
U2/OVS/Repositories/0004fb0000030000c43cdd00584279eaq  
...  
fs_locationq  
U-/dev/mapper/360002ac000000000000000740000d140q  
0004fb0000030000c43cdd00584279ea
```

```
# df -h |grep Repo  
/dev/mapper/360002ac000000000000000740000d140 5.0T 3.4T 1.7T 68%  
/OVS/Repositories/0004fb0000030000c43cdd00584279ea
```

```
# cat .ovsrepo  
OVS_REPO_UUID=0004fb0000030000c43cdd00584279ea  
OVS_REPO_VERSION=3.0  
OVS_REPO_MGR_UUID=0004fb000001000079a86253a43b91fb  
OVS_REPO_ALIAS=tr_repo01
```

AGENDA

1. Using the latest hardware
2. Server with local SDCARD
3. BIOS
4. Upgrade of OVM server
5. Corrupt filesystem
- 6. Migration from VMWare to OVM**

Migration from VMware to OVM

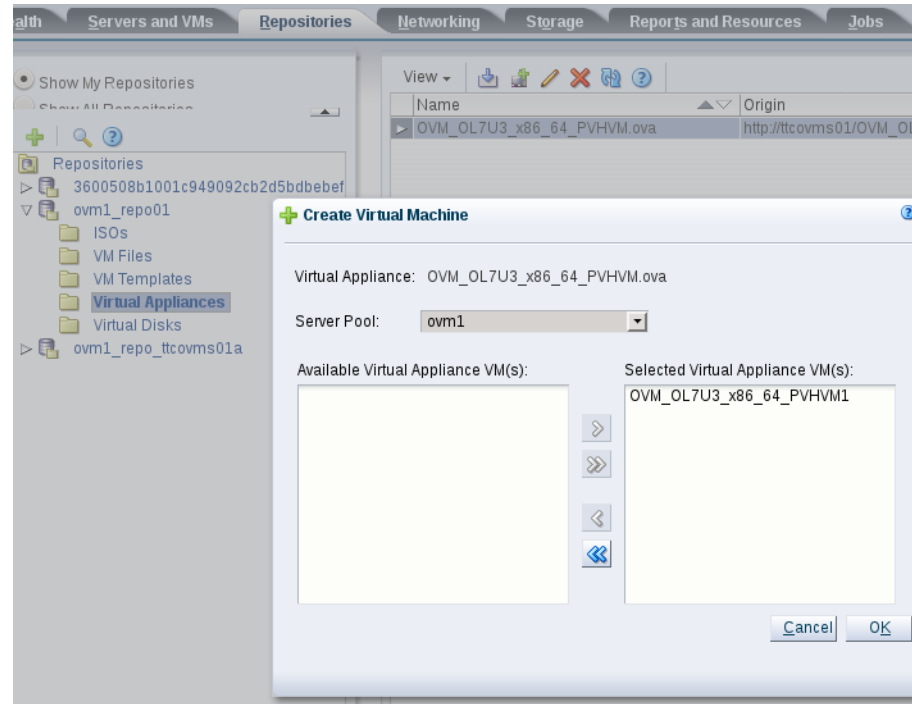
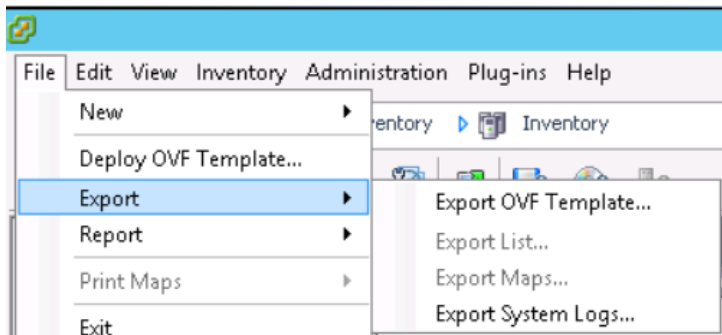
Preparation

- To boot an OS, you need drivers for your (virtual) hardware
 - For VM's there are optimized paravirtualized drivers
- Install/update UEK Kernel if possible
 - Contains XEN paravirtualized drivers and non-paravirtualized drivers
- Re-create the initial ramdisk and add the XEN modules
`dracut --add-drivers "xen-blkfront xen-netfront" --force`
- Adapt `/etc/fstab` and replace `sdX` with `xvdX` (e.g. `/dev/sdb` with `/dev/xvdb`)
 - Better: use `LABEL=` or `UUID=` instead of the device name
 - `# tune2fs -L u01 /dev/sdb #afterwards, set LABEL=u01 instead of /dev/sdb`
it will not change between vmware and ovm

Migration from VMware: OVF format

Different possibilities

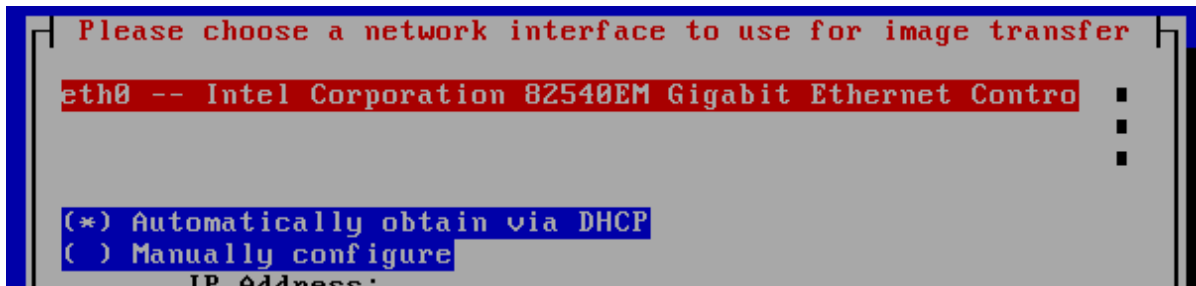
- See also <http://www.oracle.com/technetwork/server-storage/vm/ovm-migration-3708320.pdf>
- Export it in OVF format, import it from OVF format
 - Contains the disk images and configuration of VM (memory, cpus, diskmapping)
 - Clone process will convert the disk format



Migration from VMware: P2V

P2V – Physical to VM → also works with VMware to OVM

- Physical server / VMWare will be converted to a template
 - No changes on physical server / VMWare
 - On the fly conversion, presented via webserver to import in ovm manager
- Conversion program is on the OVM server installation CD
 - Boot the physical server / VMWare with this CD, and enter on boot-prompt
boot: **p2v**
 - Choose network card and configure network settings (dhcp/static IP)



```
Please choose a network interface to use for image transfer
eth0 -- Intel Corporation 82540EM Gigabit Ethernet Contro
(*) Automatically obtain via DHCP
( ) Manually configure
IP Address:
```


Migration from VMware: P2V

- Afterwards a webserver is started to download the files as a template

```
Starting web server
HTTPS web server is running on 10.0.2.15 port 443...
Interrupt with control-C
```

- import it as a template in OVM Manager: specify following Template URLs
 - <https://<ip>/System-sda.img>
 - <https://<ip>/vm.cfg>

```
acpi = 1
apic = 1
keymap = 'en-us'
on_crash = 'restart'
on_reboot = 'restart'
device_model = '/usr/lib/xen/bin/qemu-dm'
kernel = '/usr/lib/xen/boot/hvmloader'
builder = 'hvm'
pae = 1
serial = 'pty'
vif = ['bridge=xenbr0,type=ioemu']
vnc = 1
vncconsole = 1
vnclisten = '0.0.0.0'
vncunused = 1
vncpasswd = 'manager'
memory = '2048'
name = 'myConvertedServer'
vcpus = 2
disk = ['file:System-sda.img,hda,w',
]
```

Migration from VMware: copying the disks

Migration by copying the disks

- Transfer the vmdk file(s) to the OVM repository
 - NFS export, ftp, scp, ...

```
scp vm3_sda.vmdk ovms01:/OVS/Repositories/<UUID>/VirtualDisks/
```

- Convert the disk to the native format (RAW)

```
qemu-img convert -f vmdk vm3_sda.vmdk -O raw vm3_xvda.img
```

- Can be combined with the copy step (via pipes: `cat file | ssh ovms01 "cd repo; cat - | qemu-img ..."`)
- The image name must be unique over **all** repositories
e.g. `<vmname>_<diskname>.img`





- Re-scan the repository

```
OVM> refresh repository name=ovm1_repo01  
OVM> show virtualdisk name=vm3_xvda.img
```

Migration from VMware: copying the disks

- Create a new VM and map the copied disk
 - Use Hardware-virtualization, not paravirtualisation

- Add the virtual disk you copied

Slot	Disk Type	Name	Size (GiB)	Actions
0	Virtual Disk	vm3_xvda.img	200.0	   

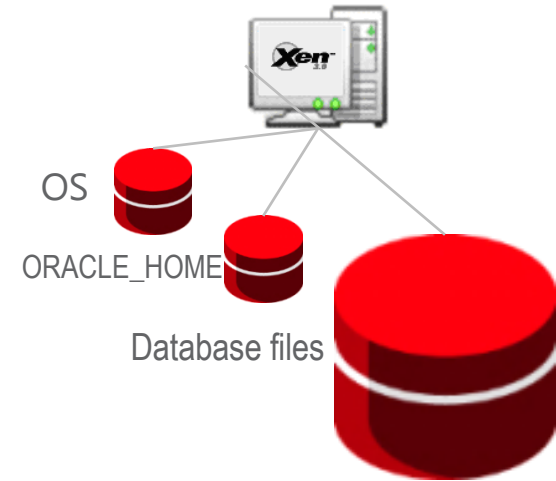
* Server Pool: ovm1
Server: ttcovms02
* Repository: 3600508b1001c949092cb2c
* Name: doag2018
 Enable High Availability
 Enable Huge Pages
Description:

Operating System: Oracle Linux 7
Mouse Device Type: OS Default
Keymap: en-us (English, United State)
Domain Type: Xen HVM PV Drivers
Start Policy: Current Server
Max. Memory (MB): 2048
Memory (MB): 2048
Max. Processors: 4
Processors: 4
Priority: 50
Processor Cap %: 100
Restart Action On Crash: Restart

- If the disk is a physical LUN: simply present it to the OVM server

Migration from VMware: Reducing downtime

- OS, Application and Database are usually on different disks
 - OS and Application binary (ORACLE_HOME) are mostly static files
- Create a clone of the VMWare VM
 - Put database in backup mode before
 - DO NOT start the database in the clone
- Transfer and convert the cloned VM to OVM (the original is still online)
 - Optionally: it is a good moment to upgrade the OS:
Install a new EL7 and integrate the software/database disks
 - – or –
 - Install a new EL7 and a new ORACLE_HOME and integrate the database disks, or restore the database from a backup



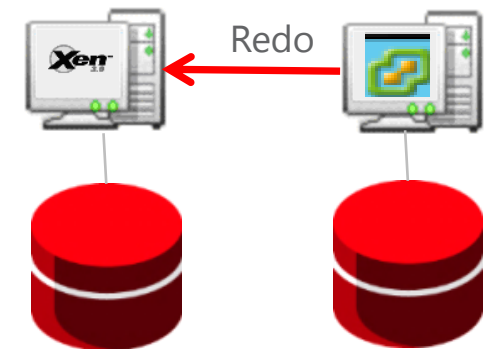
Migration from VMware: Reducing downtime

- Downtime: the final migration steps
 - Stop the application on the original VM
 - "rsync" the filesystem differences to the new VM (if required)
 - Transfer the database (see next slide)
 - Stop the original VM, configure and start the new VM with the original name

Migration from VMware: Reducing downtime

Migration of the database

- Transfer/convert the disk-images from the original VM
 - Time-consuming for big databases
- Update the clone
 - The clone already contains outdated datafiles
 - Create and apply an incremental backup beginning with scn of the clone, afterwards apply the latest redo
- Use Data Guard
 - After the initial clone, replace the controlfile by a standby controlfile
 - Configure Data Guard between the original and the migrated VM
 - Downtime is the time of the switchover



Convert to OVM: Possible problems

■ Possible problems

- If it boots in a dracut shell and complains about filesystems (EL7):

- https://docs.oracle.com/cd/E64076_01/E85146/html/vmrns-bugs-3.4.1-virtualbox-export-ol7-does-not-start.html
- Edit the grub config (e), replace `'rhgb quiet'` by `'xen_emul_unplug=never'`
- After booting: `dracut --add-drivers "xen-blkfront xen-netfront" --force`
- Better: do last step in the old environment before stopping and converting

- If it boots to an emergency prompt (EL6)

- `/dev/sd?` used in `/etc/fstab` and grub
 - replace it by `/dev/xvd?`
 - Add the `xen-*` modules as described above
- <https://support.oracle.com/epmos/faces/DocumentDisplay?id=2033007.1>

Convert to OVM: Possible problems

- Possible problems
 - No X11 in console
 - Wrong driver in /etc/X11/xorg.conf
 - Replace Driver "**VMware**" by Driver "**vesa**"
 - <https://support.oracle.com/epmos/faces/DocumentDisplay?id=2066952.1>

Enhance Disk I/O

- PV drivers process disk-I/O through a ring-buffer between domU and dom0
 - Default: 32 entries of 128KB
 - Requests larger 128KB are split in multiple requests
- New with OVM 3.4, can be adjusted
 - `xen-blkfront.max` (<4.6), `xen-blkfront.max_indirect_segments`: number of 4KB units, up to 256 (=1MB)
 - `xen-blkfront.max_ring_page_order=N`: 2^N buffer pages
- Requires Kernel UEK3 3.8.13-90 or later
 - in grub: `kernel ... xen-blkfront.max_indirect_segments=256 xen-blkfront.max_ring_page_order=4`
- Example on a customer system
 - `xen-blkfront.max=64 xen-blkfront.max_ring_page_order=4`
 - Virtual disk on a ocfs2 storage repository
 - 2-3 times faster with UEK Kernel and this setting compared to RH Kernel (with 'dd' utility)

Trivadis @ DOAG 2018

#opencompany

- Booth: 3rd Floor – next to the escalator
- We share our Know how!
Just come across, Live-Presentations
and documents archive
- T-Shirts, Contest and much more
- We look forward to your visit

