Using Oracle Linux 6.4 with NetApp ONTAP

ONTAP SAN Host

Sean Daffy
June 22, 2020

Table of Contents

Using Oracle Linux 6.4 with NetApp ONTAP

Installing the Linux Unified Host Utilities

The NetApp Linux Unified Host Utilities software package is available on the NetApp Support Site in a 32-bit and 64-bit .rpm file. If you do not know which file is right for your configuration, use the NetApp Interoperability Matrix Tool to verify which one you need.

Installing the Linux Unified Host Utilities is strongly recommended, but not mandatory. The utilities do not change any settings on your Linux host. The utilities improve management and assist NetApp customer support in gathering information about your configuration.

Before you begin

If you have a version of Linux Unified Host Utilities currently installed you should upgrade it or, you should remove it and use the following steps to install the latest version.

1. Download the 32-bit or 64-bit Linux Unified Host Utilities software package from the NetApp Support Site Site to your host.
2. Use the following command to install the software package:

   `rpm -ivh netapp_linux_unified_host_utilities-7-1.x86_64`

SAN Toolkit

The toolkit is installed automatically when you install the NetApp Host Utilities package. This kit provides the `sanlun` utility, which helps you manage LUNs and HBAs. The `sanlun` command returns information about the LUNs mapped to your host, multipathing, and information necessary to create initiator groups.

Example

In the following example, the `sanlun lun show command` returns LUN information.

```
# sanlun lun show all
controller(7mode/E-Series)/            device     host               lun
vserver(cDOT/FlashRay)   lun-pathname  filename   adapter  protocol  size    Product
------------------------------------------------------------------------------
data_vserver          /vol/vol1/lun1   /dev/sdb   host16   FCP       120.0g  cDOT
data_vserver          /vol/vol1/lun1   /dev/sdc   host15   FCP       120.0g  cDOT
data_vserver          /vol/vol2/lun2   /dev/sdd   host16   FCP       120.0g  cDOT
data_vserver          /vol/vol2/lun2   /dev/sde   host15   FCP       120.0g  cDOT
```

SAN Booting

Before you begin

If you decide to use SAN booting, it must be supported by your configuration. You can use the NetApp Interoperability Matrix Tool to verify that your OS, HBA, HBA firmware and the HBA boot BIOS, and ONTAP version are supported.

1. Map the SAN boot LUN to the host.
2. Verify multiple paths are available.

   Remember, multiple paths will only be available after the host OS is up and running on the paths.

3. Enable SAN booting in the server BIOS for the ports to which the SAN boot LUN is mapped.

   For information on how to enable the HBA BIOS, see your vendor-specific documentation.

4. Reboot the host to verify the boot is successful.

Multipathing

For Oracle Linux 6.4 the /etc/multipath.conf file must exist, but you do not need to make specific changes to the file. Oracle Linux 6.4 is compiled with all settings required to recognize and correctly manage ONTAP LUNs.
To Enable ALUA Handler, perform the following steps:

1. Create a backup of the initrd-image.
2. Append the following parameter value to the kernel for ALUA and non-ALUA to work:
   `rdloaddriver=scsi_dh_alua`

   Example

   ```
   kernel /vmlinuz-3.8.13-68.1.2.el6uek.x86_64 ro root=/dev/mapper/vg_ibmx3550m421096-
   lv_root rd_NO_LUKSrd_LVM_LV=vg_ibmx3550m421096/lv_root LANG=en_US.UTF-8
   rd_NO_MDSYSFONT=latarcyrheb-sun16 crashkernel=256M KEYBOARDTYPE=pc KEYTABLE=us
   rd_LVM_LV=vg_ibmx3550m421096/lv_swap rd_NO_DM rhgb quiet rdloaddriver=scsi_dh_alua
   ```

3. Use the `mkinitrd` command to recreate the initrd-image.
   Oracle 6x and later versions use either:
   The command: `mkinitrd -f /boot/ initrd-"uname -r".img uname -r`
   Or
   The command: `dracut -f`
4. Reboot the host.
5. Verify the output of the `cat /proc/cmdline` command to ensure that the setting is complete.
   You can use the `multipath -ll` command to verify the settings for your ONTAP LUNs.

There should be two groups of paths with different priorities. The paths with the higher priorities are Active/Optimized, meaning they are serviced by the controller where the aggregate is located. The paths with the lower priorities are active but are non-optimized because they are served from a different controller. The non-optimized paths are only used when no optimized paths are available.

Example

The following example displays the correct output for an ONTAP LUN with two Active/Optimized paths and two Active/non-Optimized paths:

```
# multipath -ll
3600a09803831347657244e527766394e dm-5 NETAPP,LUN C-Mode
size=80G features='4 queue_if_no_path pg_init_retries 50 retain_attached_hw_handle'
hwhandler='1 alua' wp=rw
|-+- policy='round-robin 0' prio=50 status=active
| |- 0:0:26:37 sdje 8:384   active ready running
| |- 0:0:25:37 sdik 135:64  active ready running
`-+- policy='round-robin 0' prio=10 status=enabled
  |- 0:0:18:37 sdda 70:128  active ready running
  |- 0:0:19:37 sddu 71:192  active ready running
```

Do not use an excessive number of paths to a single LUN. No more than 4 paths should be required. More than 8 paths might cause path issues during storage failures.

Recommended Settings

The Oracle Linux 6.4 OS is compiled to recognize ONTAP LUNs and automatically set all configuration parameters correctly.

The multipath.conf file must exist for the multipath daemon to start, but you can create an empty, zero-byte file using the command:
`touch /etc/multipath.conf`
The first time you create this file, you might need to enable and start the multipath services.

```
[root@jfs0 ~]# chkconfig multipathd on
[root@jfs0 ~]#/etc/init.d/multipathd start
```

There is no requirement to add anything directly to multipath.conf, unless you have devices that you do not want to be managed by multipath or you have existing settings that override defaults.
You can add the following syntax to the multipath.conf file to exclude the unwanted devices.

Replace the <DevId> with the WWID string of the device you want to exclude. Use the following command to determine the WWID:

```
blacklist {
        wwid <DevId>
        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
        devnode "^hd[a-z]"
        devnode "^cciss.*"
}
```

Example

In this example, sda is the local SCSI disk that we need to blacklist.

1. Run the following command to determine the WWID:

```
# /lib/udev/scsi_id -gud /dev/sda
360030057024d0730239134810c0cb833
```

2. Add this WWID to the blacklist stanza in the /etc/multipath.conf:

```
blacklist {
     wwid   360030057024d0730239134810c0cb833
     devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
     devnode "^hd[a-z]"
     devnode "^cciss.*"
}
```

You should always check your `/etc/multipath.conf` file for legacy settings, especially in the defaults section, that may be overriding default settings.
The table below shows the critical `multipathd` parameters for ONTAP LUNs and the required values. If a host is connected to LUNs from other vendors and any of these parameters are overridden, they will need to be corrected by later stanzas in `multipath.conf` that apply specifically to ONTAP LUNs. If this is not done, the ONTAP LUNs may not work as expected. These defaults should only be overridden in consultation with NetApp and/or OS vendor and only when the impact is fully understood.

Parameter	**Setting**
detect_prio	yes
dev_loss_tmo	"infinity"
failback	immediate
fast_io_fail_tmo	5
features	"3 queue_if_no_path pg_init_retries 50"
flush_on_last_del	"yes"

Parameter	Setting
hardware_handler	"0"
no_path_retry	queue
path_checker	"tur"
path_grouping_policy	"group_by_prio"
path_selector	"round-robin 0"
polling_interval	5
prio	"ontap"
product	LUN.*
retain_attached_hw_handler	yes
rr_weight	"uniform"
user_friendly_names	no
vendor	NETAPP

Example

The following example shows how to correct an overridden default. In this case, the `multipath.conf` file defines values for `path_checker` and `detect_prio` that are not compatible with ONTAP LUNs.
If they cannot be removed because of other SAN arrays still attached to the host, these parameters can be corrected specifically for ONTAP LUNs with a device stanza.

```
defaults {
 path_checker readsector0
 detect_prio no
 }
devices {
 device {
 vendor "NETAPP "
 product "LUN.*"
 path_checker tur
 detect_prio yes
 }
}
```

Known Problems and Limitations

NetApp Bug ID	Title	Description	Bugzilla ID
713555	QLogic adapter resets are seen on OL6.4 and OL5.9 with UEK2 on controller faults such as takeover/giveback, and reboot	QLogic adapter resets are seen on OL6.4 hosts with UEK2 (kernel-uek-2.6.39-400.17.1.el6uek) or OL5.9 hosts with UEK2 (kernel-uek-2.6.39 400.17.1.el5uek) when controller faults happen (such as takeover, giveback, and reboots). These resets are intermittent. When these adapter resets happen, a prolonged I/O outage (sometimes, more than 10 minutes) might occur until the adapter resets succeed and the paths' status are updated by dm-multipath. In /var/log/messages, messages similar to the following are seen when this bug is hit: kernel: qla2xxx [0000:11:00.0]-8018:0: ADAPTER RESET ISSUED nexus=0:2:13. This is observed with the kernel version: On OL6.4: kernel-uek-2.6.39-400.17.1.el6uek On OL5.9: kernel-uek-2.6.39-400.17.1.el5uek	13999

NetApp Bug ID	Title	Description	Bugzilla ID
715217	Delay in path recovery on OL6.4 or OL5.9 hosts with UEK2 may result in delayed I/O resumption on controller or fabric faults	When a controller fault (storage failover or giveback, reboots and so on) or a fabric fault (FC port disable or enable) occurs with I/O on Oracle Linux 6.4 or Oracle Linux 5.9 hosts with UEK2 Kernel, the path recovery by DM-Multipath takes a long time (4mins. to 10 mins). Sometimes, during the paths recovering to active state, the following lpfc driver errors are also seen: kernel: sd 0:0:8:3: [sdlt] Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK Due to this delay in path recovery during fault events, the I/O resumption also delays. OL 6.4 Versions: device-mapper-1.02.77-9.el6 device-mapper-multipath-0.4.9-64.0.1.el6 kernel-uek-2.6.39-400.17.1.el6uek OL 5.9 Versions: device-mapper-1.02.77-9.el5 device-mapper-multipath-0.4.9-64.0.1.el5 kernel-uek-2.6.39-400.17.1.el5uek	14001

NetApp Bug ID	Title	Description	Bugzilla ID
709911	DM Multipath on OL6.4 & OL5.9 iSCSI with UEK2 kernel takes long time to update LUN path status after storage faults	On systems running Oracle Linux 6 Update4 and Oracle Linux 5 Update9 iSCSI with Unbreakable Enterprise Kernel Release 2 (UEK2), a problem has been seen during storage fault events where DM Multipath (DMMP) takes around 15 minutes to update the path status of Device Mapper (DM) devices (LUNs). If you run the "multipath -ll" command during this interval, the path status is shown as "failed ready running" for that DM device (LUN). The path status is eventually updated as "active ready running." This issue is seen with following version: Oracle Linux 6 Update 4: UEK2 Kernel: 2.6.39-400.17.1.el6uek.x86_64 Multipath: device-mapper-multipath-0.4.9-64.0.1.el6.x86_64 iSCSI: iscsi-initiator-utils-6.2.0.873-2.0.1.el6.x86_64 Oracle Linux 5 Update 9: UEK2 Kernel: 2.6.39-400.17.1.el5uek Multipath: device-mapper-multipath-0.4.9-64.0.1.el5.x86_64 iSCSI: iscsi-initiator-utils-6.2.0.872-16.0.1.el5.x86_64	13984

NetApp Bug ID	Title	Description	Bugzilla ID
739909	The SG_IO ioctl system call fails on dm-multipath devices after an FC fault on OL6.x and OL5.x hosts with UEK2	A problem is seen on Oracle Linux 6.x hosts with UEK2 kernel and Oracle Linux 5.x hosts with UEK2 kernel. The sg_* commands on a multipath device fail with EAGAIN error code (errno) after a fabric fault that makes all the paths in the active path group go down. This problem is seen only when there is no I/O occurring to the multipath devices. The following is an example: # sg_inq -v /dev/mapper/3600a098041764937303f436c75324370 inquiry cdb: 12 00 00 00 24 00 ioctl(SG_IO v3) failed with os_err (errno) = 11 inquiry: pass through os error: Resource temporarily unavailable HDIO_GET_IDENTITY ioctl failed: Resource temporarily unavailable [11] Both SCSI INQUIRY and fetching ATA information failed on /dev/mapper/3600a098041764937303f436c75324370 # This problem occurs because the path group	14082

For Oracle Linux (Red Hat compatible kernel) known issues, see the Known Issues section in the corresponding Red Hat Enterprise Linux release documentation.

Release Notes

ASM Mirroring

ASM mirroring might require changes to the Linux multipath settings to allow ASM to recognize a problem and switch over to an alternate fail group. Most ASM configurations on ONTAP use external redundancy, which means that data protection is provided by the external array and ASM does not mirror data. Some sites use ASM with normal redundancy to provide two-way mirroring, normally across different sites. See Oracle Databases on ONTAP for further information.