



Linux Multipathing

Intended Audience:

Seasoned Linux administrators, familiar with working with SAN type storage in a Linux environment.

Notables:

This document was designed and written using Centos 5.2 as a Linux platform reference in combination with a QLogic HBA for FC connection.

Pre-requisites:

- **device-mapper-multipath** Package
- **dev_sg** Package
- **sg_tools/utils** Package
- **lsscsi** Package
- Up-to-Date Driver Package for your FC HBA

Version 1

Aug, 3rd 2010



Linux Multipathing

Instructions:

Step 1: Identify the Nexsan Storage.

- From the command line, type: `lsscsi -g |grep NEXSAN`

```
root@tornado ~]# lsscsi -g |grep NEXSAN
[5:0:0:0] disk NEXSAN SATABeast2 Nw66 /dev/sdb /dev/sg2
[5:0:0:1] disk NEXSAN SATABeast2 Nw66 /dev/sdc /dev/sg3
[5:0:0:2] disk NEXSAN SATABeast2 Nw66 /dev/sdd /dev/sg4
[6:0:0:0] disk NEXSAN SATABeast2 Nw66 /dev/sde /dev/sg5
[6:0:0:1] disk NEXSAN SATABeast2 Nw66 /dev/sdf /dev/sg6
[6:0:0:2] disk NEXSAN SATABeast2 Nw66 /dev/sdg /dev/sg7
[7:0:0:0] disk NEXSAN SATABeast2 Nw66 /dev/sdh /dev/sg8
[7:0:0:1] disk NEXSAN SATABeast2 Nw66 /dev/sdi /dev/sg9
[7:0:0:2] disk NEXSAN SATABeast2 Nw66 /dev/sdj /dev/sg10
[8:0:0:0] disk NEXSAN SATABeast2 Nw66 /dev/sdk /dev/sg11
[8:0:0:1] disk NEXSAN SATABeast2 Nw66 /dev/sdl /dev/sg12
[8:0:0:2] disk NEXSAN SATABeast2 Nw66 /dev/sdm /dev/sg13
root@tornado ~]#
```

This will display every NEXSAN device visible to the system. This will also display the **/dev/sd** mapping, the **/dev/sg** mapping, and the inquiry string of the Nexsan devices visible to the system.

You may wish to review the Nexsan configuration using the Web GUI of the attached Nexsan Storage to verify that the Nexsan devices visible to the system are expected. Keep in mind that you will see multiple entries for single devices until multipathing is configured. The LUN # and Host Port # are important bits of information to take note of. There is an example on the next page.

Linux Multipathing

	Fibre		iSCSI		
	Host 0	Host 1	Host 0	Host 1	
Volumes on Controller 0					
1: 'V1-A1-C0' Array: 'A1-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN0	LUN0	---	---	→
3: 'V1-A3-C0' Array: 'A3-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN1	LUN1	---	---	→
5: 'V1-A5-C0' Array: 'A5-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN2	LUN2	---	---	→
7: 'V1-A7-C0' Array: 'A7-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN3	LUN3	---	---	→
9: 'V1-A9-C0' Array: 'A9-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN4	LUN4	---	---	→
11: 'V1-A11-C0' Array: 'A11-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN5	LUN5	---	---	→
13: 'V1-A13-C0' Array: 'A13-C0', Controller 0 Capacity: 800.0 GB (745.0 GB)	LUN6	LUN6	---	---	→
Volumes on Controller 1					
2: 'V1-A2-C1' Array: 'A2-C1', Controller 1 Capacity: 800.0 GB (745.0 GB)	LUN0	LUN0	---	---	→
4: 'V1-A4-C1' Array: 'A4-C1', Controller 1 Capacity: 800.0 GB (745.0 GB)	LUN1	LUN1	---	---	→
6: 'V1-A6-C1' Array: 'A6-C1', Controller 1 Capacity: 800.0 GB (745.0 GB)	LUN2	LUN2	---	---	→
8: 'V1-A8-C1'					

From the Nexsan Storage WEB GUI, we can see that we actually have 14 true volumes, double mapped to each port on each controller. In this example, we are in “4 Port Active/Active” mode, so each controller only presents the volumes it owns. In APAL (All Ports, All LUNs) mode, each controller is able to present LUNs owned by either controller. For simplification, the host access control list (ACL) has been configured to present only LUNs 0-2 for each controller. For this example, the Linux system has access to a total of 6 physical volumes.


Step 2: Initial Multipath Configuration

Once you have a good idea of what you have mapped and what to expect, it's time to configure and initialize multipathing. This is a multi-step process that involves editing the multipath.conf file multiple times and starting and stopping the multipath service multiple times.

The first step is to configure the `devices` section from within the multipath configuration file.

The **multipath.conf** file is found in the **/etc** directory.

Linux Multipathing

There are a few important sections within the **Multipath.conf** file, that will need to be looked at to ensure proper operation with Nexsan's Storage Systems. Nexsan support will provide a sample configuration file to use as a template which includes all of the options needed to work with Nexsan's Storage Systems. 

The three main elements of the **multipath.conf** file are:

1. The **`devnode_blacklist`** section
 - a. This section tells the Linux Multipather what node NOT to use.
2. The **`multipaths`** section
 - a. This section is where you configure the individual volumes.
 - b. You will need to supply 2 key pieces of information for each volume:
 - 1) The **`wwid`** (Worldwide Identification)
 - i. This is a 33 character ID that is unique to each volume.
 - ii. This is found by issuing the **`multipath -ll`** command.
 - 2) The **`alias`**
 - i. This is the human-friendly name to use.
 - ii. It is manually created by you and must be unique for each volume.
3. The **`devices`** section
 - a. The section defines the Storage Device as a whole.
 - b. The key elements in this section are:
 - 1) The vendor
 - i. Must be **"NEXSAN "**
 - ii. This field must be 8 characters long, so add the 2 spaces after the name.
 - 2) The product
 - i. **"SATABeast.*"**
 - a) For **SATABeast** products
 - ii. **"SATABoy.*"**
 - a) For **SATABoy** products
4. Refer to the Nexsan **Multipath.conf** example file for exacting details!



Linux Multipathing

To configure the `devices` section:

- 1) Open the multipath.conf file for editing
- 2) Add the following entry into the `devices` section of the file

```
device {
    vendor                "NEXSAN  "
    product               "SATAB.*"
    path_grouping_policy  group_by_prio
    getuid_callout        "/sbin/scsi_id -g -u -s /block/%n"
    prio_callout          "/sbin/mpath_prio_alua %d"
    eatures               "1 queue_if_no_path"
    path_checker          tur
    path_selector         "round-robin 0"
    hardware_handler      "0"
    rr_weight             priorities
    no_path_retry         5
    failback              300
    rr_min_io             16
}
```

The devices section is what enables Linux multipathing to recognize a storage device as valid for multipathing. Linux multipathing will not work for Nexsan storage until you have completed this step.

Once the **multipath.conf** file has been edited and saved, you must load the multipath kernel module and restart the service for the changes to take effect. Issue the following commands:

- # **modprobe dm-multipath**
- # **service multipathd start**

```
[root@tornado ~]# modprobe dm-multipath
[root@tornado ~]# service multipathd start
Starting multipathd daemon:
[root@tornado ~]#
```

You must now verify that settings properly took effect and the Nexsan Storage is recognized as valid for multipathing.

Linux Multipathing

Issue these commands: (This will verify that the `devices` section is properly configured)

- # multipath -F
- # multipath -v2

You should get the following output:

```
[root@tornado ~]# multipath -F
[root@tornado ~]# multipath -v2
create: mpath22 (36000402001fc517565e2df5300000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  5:0:0:0 sdb 8:16 [undef][ready]
  round-robin 0 [prio=1][undef]
  8:0:0:0 sdk 8:160 [undef][ready]
create: mpath23 (36000402001fc517565e2dfe000000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  5:0:0:1 sdc 8:32 [undef][ready]
  round-robin 0 [prio=1][undef]
  8:0:0:1 sdj 8:176 [undef][ready]
create: mpath24 (36000402001fc517565e2de2100000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  5:0:0:2 sdd 8:48 [undef][ready]
  round-robin 0 [prio=1][undef]
  8:0:0:2 sdm 8:192 [undef][ready]
create: mpath29 (36000402001fc517565e2dfb800000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  6:0:0:0 sde 8:64 [undef][ready]
  round-robin 0 [prio=1][undef]
  7:0:0:0 sdh 8:112 [undef][ready]
create: mpath30 (36000402001fc517565e2dfc700000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  6:0:0:1 sdf 8:80 [undef][ready]
  round-robin 0 [prio=1][undef]
  7:0:0:1 sdi 8:128 [undef][ready]
create: mpath31 (36000402001fc517565e2de0f00000000) NEXSAN, SATABeast2
[size=745G][features=0][hwandler=0][n/a]
  round-robin 0 [prio=1][undef]
  6:0:0:2 sdg 8:96 [undef][ready]
  round-robin 0 [prio=1][undef]
  7:0:0:2 sdj 8:144 [undef][ready]
[root@tornado ~]#
```

These commands clear the multipath device listings and recreate them. If your output does not include Nexsan entries, re-edit the multipath.conf file and watch for typos and incomplete sections.

Looking at this output, you will notice each entry has a very long 360000xxxx number associated with it. You will need this number to complete the `multipaths` section of the multipath.conf file as part of the next step. Do not proceed until the Nexsan devices are recognized.

Linux Multipathing

Step 3: Multipaths configuration

****Do Not Skip This Step****

You must properly configure both the `device` and `multipaths` section in order to get Linux multipathing working properly!

Now that the devices section of the multipath.conf file is configured and the multipathing system recognizes Nexsan Storage, you can now edit the **`multipaths`** section of your **multipath.conf** file. The key item is the long number that starts with **36xxxxx**. This is your **wwid** that you will need to use. It is automatically created by the system, so you must use the number displayed. You must make a **multipath** entry in the **`multipaths`** for **EACH** Nexsan device

To Configure the **`multipaths`** section:

- 1) Open the multipath.conf file for editing
- 2) Add the following entry into the **`multipaths`** section of the file

```
multipath {  
    wwid                36000402001fc517565e2dfe000000000  
    alias               nex-c0-11  
    path_grouping_policy group_by_prio  
    prio_callout        "/sbin/mpath_prio_alua %d"  
    features             "1 queue_if_no_path"  
    path_selector        "round-robin 0"  
    path_checker         readsector0  
    failback            300  
    rr_weight            priorities  
    no_path_retry        5  
    rr_min_io           16  
}
```

Once the **multipath.conf** file has been edited and saved, the multipath service must be reloaded by re-issuing the following multipath commands:

- **# service multipathd reload**
- **# multipath -ll**

Linux Multipathing

If all is correct, the following output should be seen.

```
[root@tornado etc]# service multipathd reload
Reloading multipathd: [ OK ]
[root@tornado etc]# multipath -ll
nex-c1-12 (36000402001fc517565e2de0f00000000) dm-7 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][active]
    6:0:0:2 sdg 8:96 [active][ready]
    round-robin 0 [prio=50][enabled]
    7:0:0:2 sdj 8:144 [active][ready]
nex-c1-11 (36000402001fc517565e2dfc700000000) dm-6 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][active]
    6:0:0:1 sdf 8:80 [active][ready]
    round-robin 0 [prio=50][enabled]
    7:0:0:1 sdi 8:128 [active][ready]
nex-c1-10 (36000402001fc517565e2dfb800000000) dm-5 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][active]
    6:0:0:0 sde 8:64 [active][ready]
    round-robin 0 [prio=50][enabled]
    7:0:0:0 sdh 8:112 [active][ready]
nex-c0-12 (36000402001fc517565e2de2100000000) dm-4 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][active]
    5:0:0:2 sdd 8:48 [active][ready]
    round-robin 0 [prio=50][enabled]
    8:0:0:2 sdm 8:192 [active][ready]
nex-c0-11 (36000402001fc517565e2dfe000000000) dm-3 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][enabled]
    5:0:0:1 sdc 8:32 [active][ready]
    round-robin 0 [prio=50][enabled]
    8:0:0:1 sdj 8:176 [active][ready]
nex-c0-10 (36000402001fc517565e2df5300000000) dm-2 NEXSAN,SATABeast2
[size=745G][features=1 queue_if_no_path][hwandler=0][rw]
  round-robin 0 [prio=50][active]
    5:0:0:0 sdb 8:16 [active][ready]
    round-robin 0 [prio=50][enabled]
    8:0:0:0 sdk 8:160 [active][ready]
[root@tornado etc]#
```

There are several things that have changed for the first time we saw this output. Note the following:

- 1) The name mpath## is now your alias
- 2) The device now has a dm-# after the WWID
- 3) The paths now have an [active][ready] state instead of [undef][ready]
- 4) Each path now has a [prio=50] instead of [prio=1] and are also [enabled]
- 5) Compare this output to the original output for the `device` section step.

If your output does not look like this, carefully review the `multipaths` section in the multipath.conf file for typos and incomplete items.

Linux Multipathing

References:

- Device Mapper Multipath feature of Red Hat Enterprise Linux 5
 - http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5.5/html/DM_Multipath/index.html
- Online Storage Management Guide for Linux
 - http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/html/Online_Storage_Reconfiguration_Guide/index.html
- **Multipath.conf** Configuration File Information
 - <http://www.linuxcertif.com/man/5/multipath.conf/>
- Linux Multipath Implementation
 - <http://christophe.varoqui.free.fr/refbook.html>
- Multipath-Tools FAQs
 - <http://christophe.varoqui.free.fr/faq.html>
- The Linux SCSI Generic (sg) Driver
 - <http://sg.danny.cz/sg/>
- iSCSI Introduction for Linux
 - <http://sg.danny.cz/scsi/lsscsi.html>