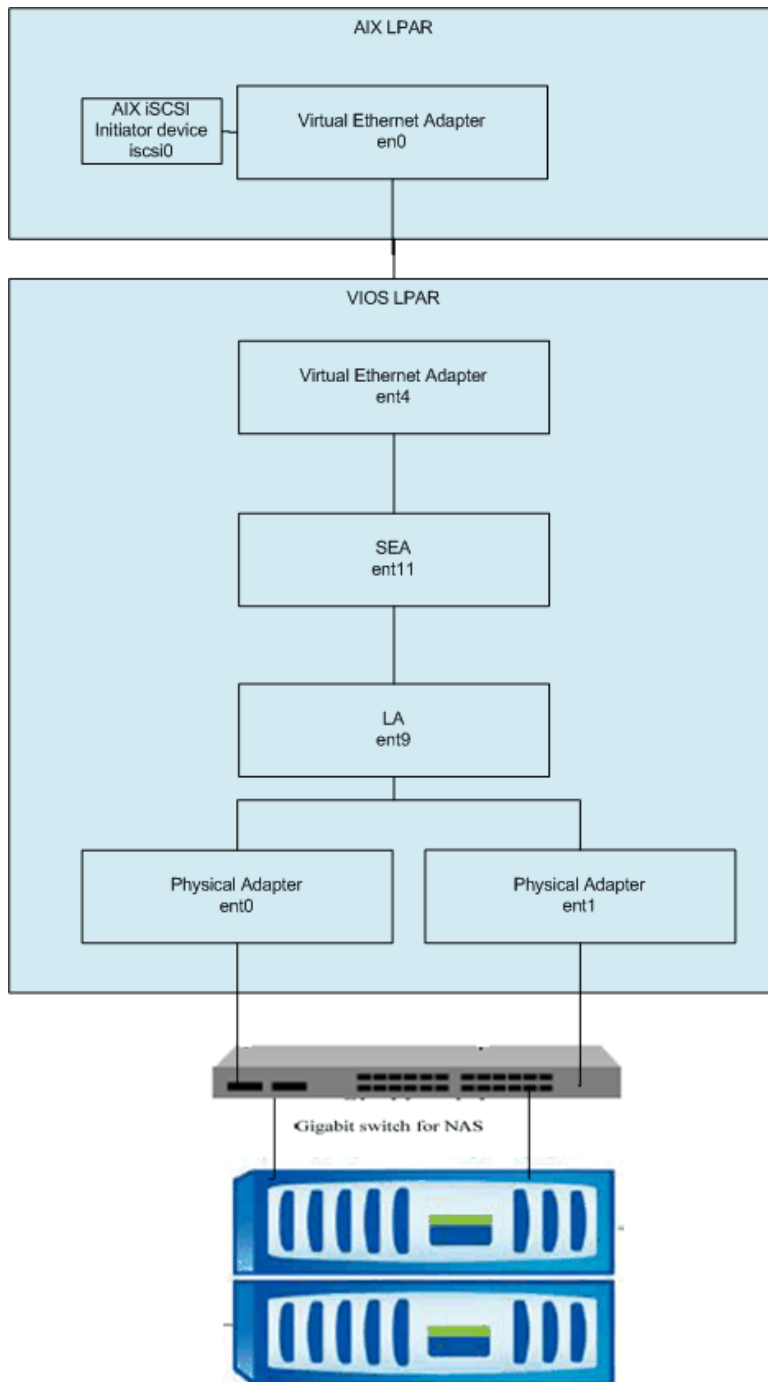


Configuring iSCSI on AIX using the iSCSI software initiator.

In this tip I'll share one method of configuring iSCSI on AIX. We will use the AIX iSCSI software initiator, a virtual Ethernet adapter, a Virtual I/O server and a NetApp storage device. The following figure illustrates the environment setup.



The VIOS is running 2.2.0.10-FP-24 SP-01 . The AIX LPAR (known as aix1), is running AIX 6.1 TL6 SP3 (shown in the following output). The following iSCSI filesets were installed by default.

```
# oslevel -s
6100-06-03-1048

# lslpp -l | grep -i iscsi
```

```

devices.common.IBM.iscsi.rte
6.1.6.0 COMMITTED Common iSCSI Files
devices.iscsi.disk.rte 6.1.6.0 COMMITTED iSCSI Disk Software
devices.iscsi.tape.rte 6.1.0.0 COMMITTED iSCSI Tape Software
devices.iscsi_sw.rte 6.1.6.0 COMMITTED iSCSI Software Device Driver
devices.pci.14102203.diag 6.1.0.0 COMMITTED IBM 1 Gigabit-TX iSCSI TOE
devices.pci.14102203.rte 6.1.0.0 COMMITTED IBM 1 Gigabit-TX iSCSI TOE
devices.pci.1410cf02.diag 6.1.0.0 COMMITTED 1000 Base-SX PCI-X iSCSI TOE
devices.pci.1410cf02.rte 6.1.0.0 COMMITTED 1000 Base-SX PCI-X iSCSI TOE
devices.pci.1410d002.com 6.1.6.0 COMMITTED Common PCI iSCSI TOE Adapter
devices.pci.1410d002.diag 6.1.6.0 COMMITTED 1000 Base-TX PCI-X iSCSI TOE
devices.pci.1410d002.rte 6.1.0.0 COMMITTED 1000 Base-TX PCI-X iSCSI TOE
devices.pci.1410e202.diag 6.1.0.0 COMMITTED IBM 1 Gigabit-SX iSCSI TOE
devices.pci.1410e202.rte 6.1.0.0 COMMITTED IBM 1 Gigabit-SX iSCSI TOE
devices.pci.14102e01.diag 6.1.0.0 COMMITTED 1000 Base-TX PCI-X iSCSI TOE
devices.pci.77102e01.rte 6.1.0.0 COMMITTED PCI-X 1000 Base-TX iSCSI TOE
devices.common.IBM.iscsi.rte
6.1.6.0 COMMITTED Common iSCSI Files
devices.iscsi_sw.rte 6.1.6.0 COMMITTED iSCSI Software Device Driver
devices.pci.1410d002.com 6.1.6.0 COMMITTED Common PCI iSCSI TOE Adapter
devices.pci.1410d002.rte 6.1.0.0 COMMITTED 1000 Base-TX PCI-X iSCSI TOE

```

The iSCSI software initiator enables AIX to access storage devices using TCP/IP on Ethernet network adapters. There are two **virtual** Ethernet adapters (VEAs) in this LPAR.

```

# lsdev -Cc adapter | grep ent
ent0 Available Virtual I/O Ethernet Adapter (1-lan)
ent1 Available Virtual I/O Ethernet Adapter (1-lan)

```

Two virtual SCSI (VSCSI) disks are used for rootvg. These disks map to logical volumes on internal SAS drives in the VIO servers (VIOS).

```

# lsdev -Cc disk
hdisk0 Available Virtual SCSI Disk Drive
hdisk1 Available Virtual SCSI Disk Drive

# lsvg
rootvg

# lspv
hdisk0 00f667587f4d5cc9 rootvg active
hdisk1 00f667588d6d7d58 rootvg active

```

Before I can discover my new iSCSI LUN, I must first configure my AIX iSCSI initiator (the `iscsi0` device) appropriately so that that I can connect to the storage device.

Essentially I need to supply an iSCSI qualified name (iqn). This provides my AIX system with a unique identity, of which the NetApp will use to verify that I am the correct host to assign storage. The *iqn* used in the following command was given to me by my storage administrator.

```

# chdev -l iscsi0 -a initiator_name=iqn.1986-03.com.ibm:aix1

# lsattr -El iscsi0
disc_filename /etc/iscsi/targets Configuration file False
disc_policy file Discovery Policy True
initiator_name iqn.1986-03.com.ibm:aix1 iSCSI Initiator Name True
isns_srvnames auto iSNS Servers IP Addresses True
isns_srvports iSNS Servers Port Numbers True
max_targets 16 Maximum Targets Allowed True
num_cmd_elems 200 Maximum number of commands to queue to driver True

```

The next step is to update the `/etc/iscsi/targets` file on my AIX system. This file must contain the hostname or IP address of the storage device providing the iSCSI LUN. The iSCSI port, listening on the storage server, is also entered. The default port is 3260. The last two entries identify the iqn of the storage system and a password. It is not always necessary to use a password but in this case, our storage administrator has set one, so we must specify it when we attempt to connect to the device.

```
# cd /etc/iscsi/
# tail -1 targets
10.2.6.1 3260 iqn.1992-08.com.netapp:filer1 "abc123"
```

In this example, the `en0` interface is connected to our “storage” network. The interface was configured according to the IBM recommendations on iSCSI performance with AIX. Jumbo frames (MTU set to 9000) and `largesend` are enabled on the interface, along with larger values for `tcp_sendspace` and `tcp_recvspace`. We also disabled the Nagle algorithm and enabled `tcp_nodelay`.

```
# ifconfig en0
en0:
flags=1e080863,4c0<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OF
FLOAD(ACTIVE),LARGESEND,CHAIN>
inet 10.2.6.11 netmask 0xffffffff80 broadcast 10.2.6.127
    tcp_sendspace 262144 tcp_recvspace 262144 tcp_nodelay 1 rfc1323 1

# lsattr -El en0 -a mtu
mtu 9000 Maximum IP Packet Size for This Device True

# no -a |grep nagle_limit
    tcp_nagle_limit = 0
```

Please refer to the following site for more iSCSI performance on AIX considerations:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.commadmn/doc/commadmn dita/iscsi_performance.htm

On the VIO server, we enabled `jumbo_frames`, `largesend` and `large_receive`. The SEA (Shared Ethernet Adapter) device is `ent11` and the backing device is `ent9` (which is in fact an LACP aggregated link). The aggregated link device, `ent9`, consists of two physical 1GB Ethernet ports, `ent0` and `ent1`.

```
$ lsmmap -all -net
SVEA Physloc
-----
ent5 U8233.E8B.066758P-V1-C21-T1

SEA ent11
Backing device ent9
Status Available
Physloc

$ chdev -dev ent11 -attr largesend=1

$ lsdev -dev ent11 -attr
attribute value description user_settable

accounting enabled Enable per-client accounting of network statistics True
ctl_chan ent7 Control Channel adapter for SEA failover True
gvrp no Enable GARP VLAN Registration Protocol (GVRP) True
ha_mode auto High Availability Mode True
jumbo_frames yes Enable Gigabit Ethernet Jumbo Frames True
large_receive yes Enable receive TCP segment aggregation True
largesend 1 Enable Hardware Transmit TCP Resegmentation True
netaddr 0 Address to ping True
pvid 1 PVID to use for the SEA device True
pvid_adapter ent5 Default virtual adapter to use for non-VLAN-tagged packets True
qos_mode disabled N/A True
real_adapter ent9 Physical adapter associated with the SEA True
thread 1 Thread mode enabled (1) or disabled (0) True
virt_adapters ent5 List of virtual adapters associated with the SEA (comma separated) True

$ lsdev -dev ent9 -attr
attribute value description user_settable
adapter_names ent0,ent1 EtherChannel Adapters True
alt_addr 0x000000000000 Alternate EtherChannel Address True
auto_recovery yes Enable automatic recovery after failover True
backup_adapter NONE Adapter used when whole channel fails True
hash_mode src_dst_port Determines how outgoing adapter is chosen True
interval long Determines interval value for IEEE 802.3ad mode True
```

```

mode          8023ad      EtherChannel mode of operation      True
netaddr       0          Address to ping                      True
noloss_failover yes      Enable lossless failover after ping failure True
num_retries   3          Times to retry ping before failing   True
retry_time    1          Wait time (in seconds) between pings True
use_alt_addr  no          Enable Alternate EtherChannel Address True
use_jumbo_frame yes      Enable Gigabit Ethernet Jumbo Frames True

```

```

$ lsdev -dev ent0 -attr | grep -i large
large_send    yes          Enable hardware TX TCP resegmentation      True

```

```

$ lsdev -dev ent1 -attr | grep -i large
large_send    yes          Enable hardware TX TCP resegmentation      True

```

```
$ entstat -all ent11
```

```
....
```

```
-----
Statistics for adapters in the Shared Ethernet Adapter ent11
-----
```

```
Number of adapters: 2
```

```
SEA Flags: 00000023
```

```
< THREAD >
```

```
< LARGESEND >
```

```
< ACCOUNTING >
```

```
.....
```

We also installed the NetApp MPIO Host Utilities kit on the AIX LPAR (not on the VIOS).

```

# lsllpp -L NetApp*
Fileset              Level  State  Type  Description (Uninstaller)
-----
NetApp.MPIO_Host_Utilities_Kit.config
                    5.1.0.0  C    F    NetApp MPIO PCM Host Utilities
                    Kit Path Config Tool
NetApp.MPIO_Host_Utilities_Kit.fcp
                    5.1.0.0  C    F    NetApp MPIO PCM Host Utilities
                    Kit FCP Disk ODM Stanzas
NetApp.MPIO_Host_Utilities_Kit.iscsi
                    5.1.0.0  C    F    NetApp MPIO PCM Host Utilities
                    Kit iSCSI Disk ODM Stanzas
NetApp.MPIO_Host_Utilities_Kit.pcmmodm
                    5.1.0.0  C    F    NetApp MPIO PCM Host Utilities
                    Kit PCM Friend ODM Stanzas
NetApp.SAN_toolkit.sanlun  5.1.0.0  C    F    NetApp SAN Toolkit sanlun
NetApp.iSCSI_Host_Utilities_Kit.LUN.msg.en_US
                    5.1.0.0  C    F    NetApp iSCSI Host Utilities
                    Kit US English Message Catalog
NetApp.iSCSI_Host_Utilities_Kit.ODM_stanzas
                    5.1.0.0  C    F    NetApp iSCSI Host Utilities
                    Kit ODM Stanzas

```

Now we run the **cfgmgr** command on the AIX system to configure our new iSCSI disks.

```

# cfgmgr -vl iscsi0
-----
attempting to configure device 'iscsi0'
Time: 0 LEDES: 0x25b0
invoking /usr/lib/methods/cfgiscsi -l iscsi0
Number of running methods: 1
-----
Completed method for: iscsi0, Elapsed time = 0
return code = 0
***** no stdout *****
***** no stderr *****
-----
Time: 0 LEDES: 0x539
Number of running methods: 0
-----
calling savebase
return code = 0
***** no stdout *****
***** no stderr *****
Configuration time: 0 seconds

```

We now have two new iSCSI disks.

```
# lsdev -Cc disk | grep -i iscsi
hdisk2 Available MPIO NetApp iSCSI Disk
hdisk3 Available MPIO NetApp iSCSI Disk
```

```
# lspv
hdisk0          00f667587f4d5cc9      rootvg          active
hdisk1          00f667588d6d7d58      rootvg          active
hdisk2         none                 None
hdisk3         none                 None
```

```
# lspath
Enabled hdisk0  vscsi0
Enabled hdisk1  vscsi1
Enabled hdisk2  iscsi0
Enabled hdisk3  iscsi0
```

If there were any problems with the iSCSI configuration, either at the storage end or at the AIX end, I would see an error in the AIX error report after running **cfgmgr**; similar to the one shown below.

```
D3EF661B 0429100711 T H iscsi0 COMMUNICATIONS SUBSYSTEM FAILURE
```

This error could be the result of a misconfigured `/etc/iscsi/targets` file e.g. incorrect format, wrong password, etc.

The default queue depth for the disks was 8. You may consider changing this value for better performance. Although in our environment, we found that changing to a larger value did not help with performance, in fact it had a negative impact.

```
# lsattr -El hdisk2
PCM                PCM/friend/iscsiother      Path Control Module      False
algorithm          fail_over                  Algorithm                 True
clr_q              no                          Device CLEARS its Queue on error True
dist_err_pcnt     0                           Distributed Error Percentage True
dist_tw_width     50                          Distributed Error Sample Time True
hcheck_cmd        test_unit_rdy              Health Check Command      True
hcheck_interval   60                          Health Check Interval     True
hcheck_mode       nonactive                  Health Check Mode         True
host_addr         10.2.6.11                  Hostname or IP Address    False
location          Location Label              Location Label             True
lun_id            0x0                         Logical Unit Number ID    False
lun_reset_spt     yes                          LUN Reset Supported       True
max_retry_delay   60                           Maximum Quiesce Time      True
max_transfer      0x40000                     Maximum TRANSFER Size     True
port_num          0xcbc                       PORT Number               False
pvid              00f667589edfb68c0000000000000000 Physical volume identifier False
q_err             yes                          Use QERR bit              True
q_type            simple                       Queuing TYPE              True
queue_depth      8                          Queue DEPTH             True
reassign_to       120                          REASSIGN time out value  True
reserve_policy    single_path                  Reserve Policy             True
rw_timeout        30                           READ/WRITE time out value True
start_timeout     60                           START unit time out value True
target_name       iqn.1992-08.com.netapp:clip-fashdb-05.et1 Target NAME                False
unique_id         260CHndLSJcUeLBC07FAS608006NETAPPiscsi Unique device identifier  False
```

At this point, I can now create a volume group using standard AIX LVM commands.

```
# mkvg -S -y iscsivg hdisk2 hdisk3
# mklv -tjfs2 -e x -y iscsilv iscsivg 1499
# crfs -vjfs2 -d iscsilv -m /iscsifs -a logname=INLINE -A yes
# mount /iscsifs
# df -g | grep iscsi
/dev/iscsilv 749.50 748.65 1% 4 1% /iscsifs

# lspv
hdisk0          00f667587f4d5cc9      rootvg          active
hdisk1          00f667588d6d7d58      rootvg          active
hdisk2         00f667589edfb68c      iscsivg        active
hdisk3         00f667589ee00fe7      iscsivg        active
```

It was interesting to see that there was a single TCP session open between the AIX LPAR and the NetApp filer.

```
# netstat -na | grep 3260
tcp4          0  32864  10.2.6.11.54730  10.2.6.1.3260      ESTABLISHED
```

We confirmed that *largesend* was in fact being used on the AIX LPAR by checking the output from the **netstat** command.

```
# netstat -p tcp | grep -i large
      178509 large sends
      1291861075 bytes sent using largesend
      2751348 bytes is the biggest largesend
```

Based on the recommendations on the IBM website, we disabled auto-varyon on the volume group.

```
# chvg -an iscsivg
# lsvg iscsivg
VOLUME GROUP:      iscsivg          VG IDENTIFIER:    00f6675800004c000000012f9ee01030
VG STATE:          active          PP SIZE:          512 megabyte(s)
VG PERMISSION:    read/write      TOTAL PPs:        1499 (767488 megabytes)
MAX LVs:          256            FREE PPs:         0 (0 megabytes)
LVs:              1              USED PPs:         1499 (767488 megabytes)
OPEN LVs:         1              QUORUM:           2 (Enabled)
TOTAL PVs:        1              VG DESCRIPTORS:   2
STALE PVs:        0              STALE PPs:        0
ACTIVE PVs:       1              AUTO ON:        no
MAX PPs per VG:   32768          MAX PVs:          1024
LTG size (Dynamic): 256 kilobyte(s)  AUTO SYNC:        no
HOT SPARE:        no              BB POLICY:        relocatable
PV RESTRICTION:   none
```

The IBM website states:

- *Configure volume groups that are created using iSCSI devices to be in an inactive state after reboot. After the iSCSI devices are configured, manually activate the iSCSI-backed volume groups. Then, mount any associated file systems.*

Volume groups are activated during a different boot phase than the iSCSI software driver. For this reason, it is not possible to activate iSCSI volume groups during the boot process.

Please refer to the following site for more information:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=/com.ibm.aix.commadmn/doc/commadmndita/iscsi_considerations.htm

I'm not recommending the use of iSCSI with AIX, but I hope this tip helps if you ever need to configure iSCSI disk on an AIX system.

You may wish to refer to the following links for further information on this topic.

Configuring iSCSI software initiator

http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.commadmn/doc/commadmndita/iscsi_config.htm

iSCSI performance considerations

http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.commadmn/doc/commadmndita/iscsi_performance.htm

iSCSI software initiator considerations

http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=/com.ibm.aix.commadmn/doc/commadmndita/iscsi_considerations.htm

HOWTO creating Qnap iSCSI disk on AIX 5.3 and AIX 6.1

http://wiki.qnap.com/wiki/HOWTO_creating_Qnap_iSCSI_disk_on_AIX_5.3_and_AIX_6.1

PowerVM Live Partition Mobility on IBM Power Systems using DB2 and NetApp Storage System over iSCSI

<https://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/whitepaper/power/lpm/use>