# Configuring MPIO for the virtual AIX client

This document describes the procedure to set up Multi-Path I/O on the AIX clients of the virtual I/O server.

# Procedure:

This procedure assumes that the disks are already allocated to both the VIO servers involved in this configuration.

- **Creating Virtual Server and Client SCSI Adapters**

First of all, via HMC create SCSI server adapters on the two VIO servers and then two virtual client SCSI adapters on the newly created client partition, each mapping to one of the VIO servers´ server SCSI adapter.

***An example:***

Here is an example of configuring and exporting an ESS LUN from both the VIO servers to a client partition:

- **Selecting the disk to export**

You can check for the ESS LUN that you are going to use for MPIO by running the following command on the VIO servers.

*On the first VIO server:*

***$ lsdev -type disk***

```
name            status      description


..

hdisk3          Available   MPIO Other FC SCSI Disk Drive
hdisk4          Available   MPIO Other FC SCSI Disk Drive
hdisk5          Available   MPIO Other FC SCSI Disk Drive
..
```

***$lspv***

```
..
hdisk3          00c3e35c99c0a332                    None
hdisk4          00c3e35c99c0a51c                    None
hdisk5          00c3e35ca560f919                    None
..
```

In this case hdisk5 is the ESS disk that we are going to use for MPIO.

Then run the following command to list the attributes of the disk that you choose for MPIO:

***$lsdev -dev hdisk5 -attr***

..
algorithm fail_over Algorithm True
..
lun_id 0x5463000000000000 Logical Unit Number ID False
..
..
pvid 00c3e35ca560f9190000000000000000 Physical volume identifier False
..
reserve_policy single_path Reserve Policy True

Note down the lun_id, pvid and the reserve_policy of the hdisk4.

- **Command to change reservation policy on the disk**

You see that the reserve policy is set to single_path.

Change this to no_reserve by running the following command:

***$ chdev -dev hdisk5 -attr reserve_policy=no_reserve***
hdisk4 changed

*On the second VIO server:*

On the second VIO server too, find the hdisk# that has the same pvid, it could be a different one than the one on the first VIO server, but the pvid should the same.

$ lspv
..
hdisk7 00c3e35ca560f919 None
..

The pvid of the hdisk7 is the same as the hdisk5 on the first VIO server.

$ lsdev -type disk
name status description

..
hdisk7 Available MPIO Other FC SCSI Disk Drive
..

```
$lsdev -dev hdisk7 -attr
..
algorithm       fail_over                Algorithm                True
..
lun_id          0x5463000000000000       Logical Unit Number ID   False
..
pvid            00c3e35ca560f9190000000000000000 Physical volume identifier
False
..
reserve_policy  single_path              Reserve Policy           True
```

You will note that the lun_id, pvid of the hdisk7 on this server are the same as the hdisk4 on the first VIO server.

```
$ chdev -dev hdisk7 -attr reserve_policy=no_reserve
hdisk6 changed
```

- **Creating the Virtual Target Device**

Now on both the VIO servers run the mkvdev command using the appropriate hdisk#s respectively.

```
$ mkvdev -vdev hdisk# -vadapter vhost# -dev vhdisk#
```

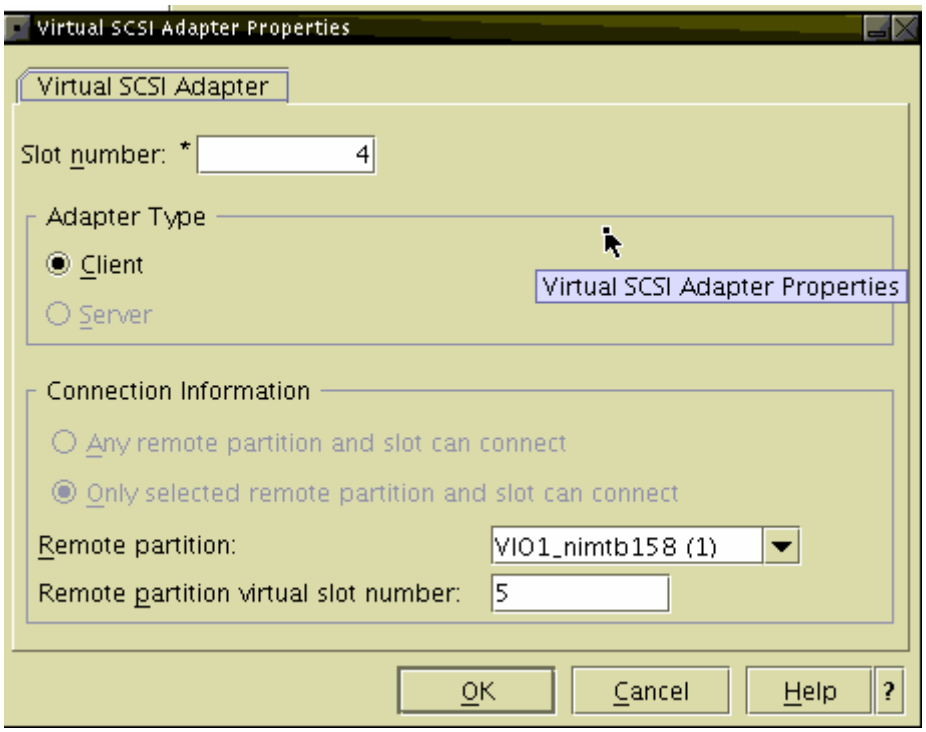
The above command might have failed when run on the second VIO server, if the reserve_policy was not set to no_reserve on the hdisk.

After the above command runs succesfully on both the servers, we have same LUN exported to the client with mkvdev command on both servers.

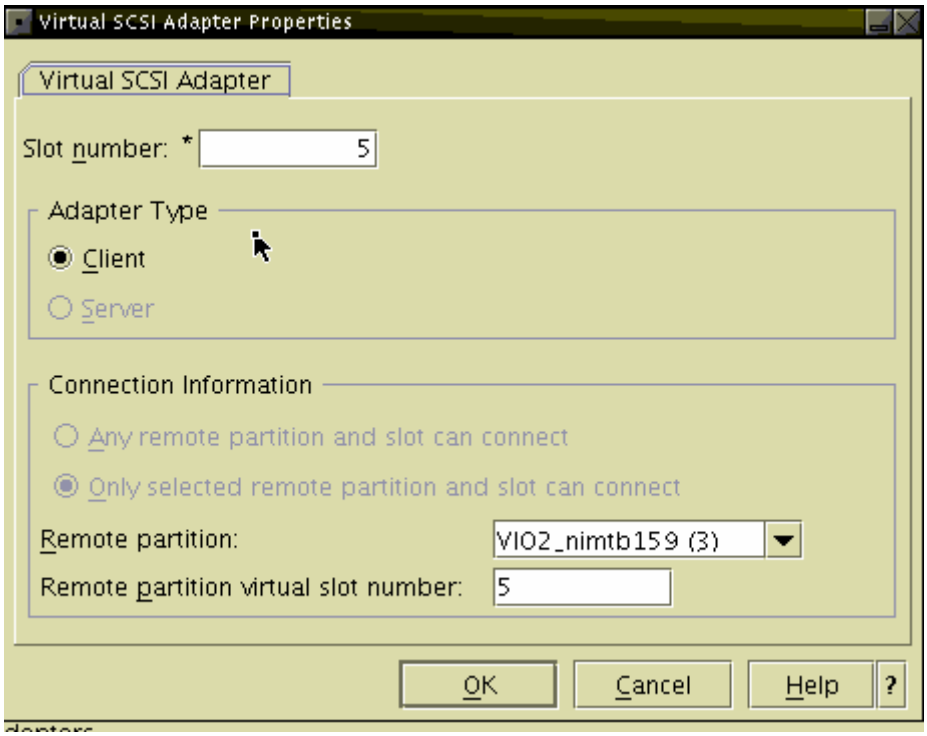
- **Check for correct mapping between the server and the client**

Double check the client via the HMC that the correct slot numbers match the respective slot numbers on the servers.

In the example, the slot number 4 for the client virtual scsi adapter maps to slot number 5 of the VIO server VIO1_nimtb158.




And the slot number 5 for the client virtual SCSI adapter maps to the slot number 5 of the VIO server VIO1_nimtb159.



- **On the client partition**

Now you are ready to install the client. You can install the client using any of the following methods described in the red book on virtualization at http://www.redbooks.ibm.com/redpieces/abstracts/sg247940.html:

1. NIM installation
2. Alternate disk installation
3. using the CD media

Once you install the client, run the following commands to check for MPIO:

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

***# lspv***
hdisk0  00c3e35ca560f919  rootvg  active

***# lspath***
Enabled  hdisk0 vscsi0
Enabled hdisk0 vscsi1

- **Dual Path**

When one of the VIO servers goes down, the path coming from that server shows as failed with the lspath command.
***# lspath***
Failed  hdisk0 vscsi0
Enabled hdisk0 vscsi1

- **Path Failure Detection**

The path shows up in the "failed" mode, even  after the VIO server is up again. We need to either change the status with the “chpath” command to “enabled” state or set the the attributes “hcheck_interval” and “hcheck_mode”  to “60” and “nonactive” respectively for a path failure to be detected automatically.

- **Setting the related attributes**

Here is the command to be run for setting the above attributes on the client partition:

$ chdev -l hdisk# -a hcheck_interval=60 –a hcheck_mode=nonactive -P

The VIO AIX client needs to be rebooted for hcheck_interval attribute to take effect.

- **EMC for Storage**

In case of using EMC device as the storage device attached to VIO server, then make sure of the following:

1. Powerpath version 4.4. is installed on the VIO servers.
2. Create hdiskpower devices which are shared between both the VIO servers.

- **Additional Information**

Another thing to take note of is that you cannot have the same name for Virtual SCSI Server Adapter and Virtual Target Device. The mkvdev command will error out if the same name for both is used.

***$ mkvdev -vdev hdiskpower0 -vadapter vhost0 -dev hdiskpower0***

Method error (/usr/lib/methods/define -g -d):
0514-013 Logical name is required.

The reserve attribute is named differently for an EMC device than the attribute for ESS or FasTt storage device. It is “reserve_lock”.

Run the following command as padmin for checking the value of the attribute.

***$ lsdev -dev hdiskpower# -attr reserve_lock***

Run the following command as padmin for changing the value of the attribute.

***$ chdev -dev hdiskpower# -attr reserve_lock=no***

- **Commands to change the Fibre Channel Adapter attributes**

And also change the following attributes of the fscsi#, fc_err_recov to “fast_fail” and dyntrk to “yes”

***$ chdev -dev fscsi# -attr fc_err_recov=fast_fail dyntrk=yes –perm***

The reason for changing the fc_err_recov to “fast_fail” is that if the Fibre Channel adapter driver detects a link event such as a lost link between a storage device and a switch, then any new I/O or future retries of the failed I/Os will be failed immediately by the adapter until the adapter driver detects that the device has rejoined the fabric. The default setting for this attribute is 'delayed_fail’.

Setting the dyntrk attribute to “yes” makes AIX tolerate cabling changes in the SAN.

The VIOS needs to be rebooted for fscsi# attributes to take effect.