Getting Started with Dual VIO Servers White Paper

April 1, 2006

Dave Buffaline
UNIX Software Services
IBM Systems and Technology Group
buffalid@us.ibm.com

Bob Minns
UNIX Software Services
IBM Systems and Technology Group
bobminns@us.ibm.com

TABLE OF CONTENTS

TABLES

FIGURES

Introduction

The Advanced POWER™ Virtualization (APV) feature for System p™ servers provides a firmware activation key to enable the Micro-Partitioning™ technology and Virtual I/O (VIO) Server type partitions. The VIO Server partition type supports virtual SCSI **server** adapters while the AIX 5L™ (or Linux®) partition type only supports virtual SCSI **client** adapters. The other components of the APV feature are the software for the VIO Server and the Partition Load Manager (PLM).

The VIO Server provides two capabilities - Shared Ethernet Adapter (SEA) support and Virtual SCSI support. The SEA feature of the VIO Server provides a bridge between the internal virtual network(s) and an external physical network. The Virtual SCSI support on the VIO Server provides block storage to multiple clients partitions as virtual SCSI devices. The storage can be backed by locally attached SCSI drives or SAN-attached storage. Both features allow multiple client partitions to share the physical adapters – Ethernet network adapters, SCSI adapters connected to local disks or Fibre Channel adapters attached to SAN storage. Future VIO Server support will include iSCSI and Serial Attached SCSI (SAS) disks.

This white paper covers a number of topics related to the configuration of the AIX 5L V5.3 client partitions and dual VIO Servers. It is assumed that the reader is familiar with using the HMC (Hardware Management Console), creating and installing AIX 5L and VIO Server partitions, and AIX 5L administration. The configuration of the client partitions is AIX 5L specific but most of the principles apply to Linux.

The configuration of the VIO Servers is performed as the prime administrator (**padmin** userid) in a restricted shell and the actual commands are shown with the $ prompt. The configuration on the AIX 5L client partitions is performed by the **root** userid and the commands are identified with the # prompt.

General Configuration Tips

There is a lot of flexibility in the setup of these systems. Here are a few ideas to think about and adapt to your environment.

From a hardware point of view, there are some things to think about that might make life easier. There are two CD/DVD drives available for the p5 systems. The first option is a DVD-ROM drive while the second is a DVD-RAM drive. One option as part of a disaster recovery (DR) strategy might to backup an existing VIO Server to the DVD-RAM and recover the VIO Server without the dependency on an external network. It is faster to backup the VIO Server to an NFS Server than the DVD-RAM. Another suggestion for the hardware is to use hot-pluggable network adapters for the VIO Server instead of the built-in integrated network adapters. Multiple client partitions may be using the Shared Ethernet Adapter on the VIO Server so it is easier to replace the PCI-based adapters without taking a system outage to replace the planar for the integrated adapter. The integrated adapters can be used for one of the client partitions such that the unlikely event of a failure would be limited to that partition.

When the new System p servers are shipped, they have a full system partition defined that is used to deliver an AIX 5L install image. The name of this full system partition is the serial number for the system. Once the system is partitioned, there is no further use for the full system partition so it is recommended just to remove it. The only catch is that this full system partition is configured as the service partition so you have to change the setting for the Service partition on the managed system to “Unassigned” or another partition in order to delete this full-system partition. The only possible requirement for using the full system partition is running diagnostics from CDROM against the entire system. A full system partition can easily be recreated at a later date by creating a new logical partition and selecting the option for “Use all the resources in the system.”.

With multiple client partitions dependent on the VIO Server for resources, clients like to use dual VIO Servers for availability. The two VIO Servers may need to be created before the client partitions if the client partitions are using virtual resources for the boot disk and/or the network for installation. The partition IDs for the VIO Servers can be set to a higher number when they are created if the idea is to use the partition ID to reflect the numbering of the client partitions (1, 2, 3 …) on the system.

It is useful to spend some time upfront as part of the planning stage to create a naming scheme for the partition names. This helps identify individual partitions especially on the larger configurations. The partition name can be set to the hostname of

the partition but hostnames are usually a little cryptic and will not help identify which system the partition is associated with. Another point is that graphical user interface for the HMC, sorts the partitions alphanumerically. A suggestion is to name the partition based on the partition type (AIX 5L or Linux partition or VIO Server), the system and a number for the partition ID; for example, LPAR_A_1 or VIO_A_51. The client has used a letter ("A") to identify the managed system (with a single system in each rack or frame).

The virtual adapters can be defined in the initial profile for the VIO Server or added dynamically as the client partitions are created. If the virtual I/O resources are being predefined on the VIO Server, set the virtual SCSI Server adapter to allow "Any remote partition and slot can connect" until the client partition is created. Once the client partition is created, the Connection Information for the virtual SCSI Server adapter in the VIO Server profile can be updated to reflect the specific client partition and virtual slot number. If the virtual SCSI Server adapters are added dynamically, the profile for the VIO Server must also be updated with the new adapters for the next time the VIO Server is shutdown and activated again. These are two separate operations (adding the virtual adapter to the VIO Server using dynamic LPAR and adding the virtual adapter information to the VIO Server partition profile for the next activation) even though the information for the virtual adapter is the same.

The difference with virtual adapters is that they can only be added or removed using dynamic LPAR operations (whereas physical adapters can also be moved between partitions). When defining the partitions for the VIO Server, do not select the virtual adapters as "Required" (just "Desired") so that they can be removed and added while the VIO Server is up and running. If the virtual adapters are "Required", they can only be changed to "Desired" in the active profile and then removed. The physical adapters that are necessary for booting the VIO Server should be "Required". All the other physical adapters can be "Desired" so they can be removed or moved to other partitions dynamically.

When creating the virtual adapters, it is also possible to set the virtual slot number rather than defaulting to the next available slot number. A useful tip to help with identification of resources between the client partitions and the dual VIO Servers is using a range of virtual slot numbers for the virtual SCSI Server adapters on each VIO Server. For example, virtual slots 11-16 can be used on VIO Server #1 for client partitions 1-6 and the virtual slots 21-26 on VIO Server #2 for client partitions 1-6. To expand this approach, it is also possible to set the virtual slot number on the client partition to match the virtual slot number on the VIO Server. For example, slot number 11 on client partition #1 is the virtual SCSI client adapter connected to the virtual SCSI Server adapter in virtual slot number 11 on the VIO Server #1, and slot number 21 on client partition #1 is the virtual SCSI client adapter connected to the virtual SCSI Server adapter in slot number 21 on VIO Server #2. With this scheme, it is necessary to increase the maximum number of virtual slots (from the default of 10) in the VIO Server and client partition profiles.

Partition Name	**LPAR_A_1**	**LPAR_A_2**	…	**VIO_A_51**	**VIO_A_52**
Partition ID	**1**	**2**	…	**51**	**52**
Virtual Serial (VTERM)	0	0		0	0
Virtual Serial (VTERM)	1	1		1	1
Virtual Ethernet VLAN 1	2	2		2	
Virtual Ethernet VLAN 2	3	3			3
Virtual SCSI	11			11	
Virtual SCSI		12		12	
…				…	
Virtual SCSI	21				21
Virtual SCSI		22			22
…					…

Table 1: Virtual Slot Numbering Scheme for Dual VIO Servers

If setting up a new system, it is important to use the latest levels of AIX 5L, VIO Server, firmware and HMC. Once the environment is stable, upgrades should only be required for stability fixes and new functionality. The latest levels at the time of writing were AIX 5L V5.3 Technology Level 4 (TL4) Service Pack 1, VIO Server V1.2.1.0, System Firmware SF240_202, and HMC V5.2.0 + MH00586.

VIO Server Installation and Backup

This is a quick summary of the options for installing and backing up the VIO Server. The individual tasks are covered in more details in the Using the VIO Server document.

The VIO Server software is supplied on a DVD as part of the Advanced POWER Virtualization feature. The VIO Server partition can be installed from the DVD using the CD/DVD drive assigned to the partition or using NIM on Linux (NIMOL) and placing the VIO Server DVD in the CD/DVD drive on the HMC.

The VIO Server can be backed up to DVD, Tape, NFS Server or a NIM Server. The DVD or Tape Drive must be assigned to the VIO Server partition. The DVD drive must be the DVD-RAM drive with the correct media.

The NFS Server can be any system as long as the HMC is able to mount the exported file system. A *nim_resources.tar* file is created on the NFS Server from the restricted shell on the VIO Server using the backupios command. The backup can be restored from the NFS Server using the installios command using a local command shell on the HMC or remotely using the secure shell to the HMC.

With VIO Server Version 1.2, it is possible to backup the VIO Server to a NIM Server from the restricted shell. The VIO Server administrator would mount the exported directory from the NIM Server and use the backupios command with the –mksysb flag to create a regular *mksysb* file (not packed with the NIMOL resources in a nim_resources.tar file). The VIO Server can be reinstalled from the NIM Server with this backup by creating a SPOT resource from the mksysb resource and a bosinst.data resource for an unattended installation.

Dual VIO Server Example

For this dual VIO Server configuration example, we will use Network Backup Interface on the client partitions for network redundancy and the SAN-attached disks will be presented by both VIO Servers to provide redundant paths to the storage for the client partitions. In this simplistic example, we are using virtual adapters for both the network and storage so that multiple partitions can quickly be created and can share the bandwidth of the Ethernet and Fibre Channel adapters on the dual VIO Servers. Using virtual adapters for network and disks, the only additional resources required for each client partition are CPU and memory. The other design choices for LVM mirroring of boot disks and Shared Ethernet Adapter (SEA) Failover are mentioned briefly at the end when testing dual VIO Servers but are not covered in the configuration steps for clarity.

The features of this sample configuration are:

- Only local SCSI storage is used for the VIO Server’s operating system
- Mirrored SCSI drives are used to increase availability of both VIO Servers due to a single disk failure
- The client partitions are configured to use Network Interface Backup configuration for network availability
- Sharing the SAN connections by multiple client partitions
- The SAN storage is provided by the SAN Volume Controller to the dual VIO Servers
- All SAN storage for the client partitions is zoned to both VIO Servers
- The client partitions will boot from the SAN via the VIO Servers

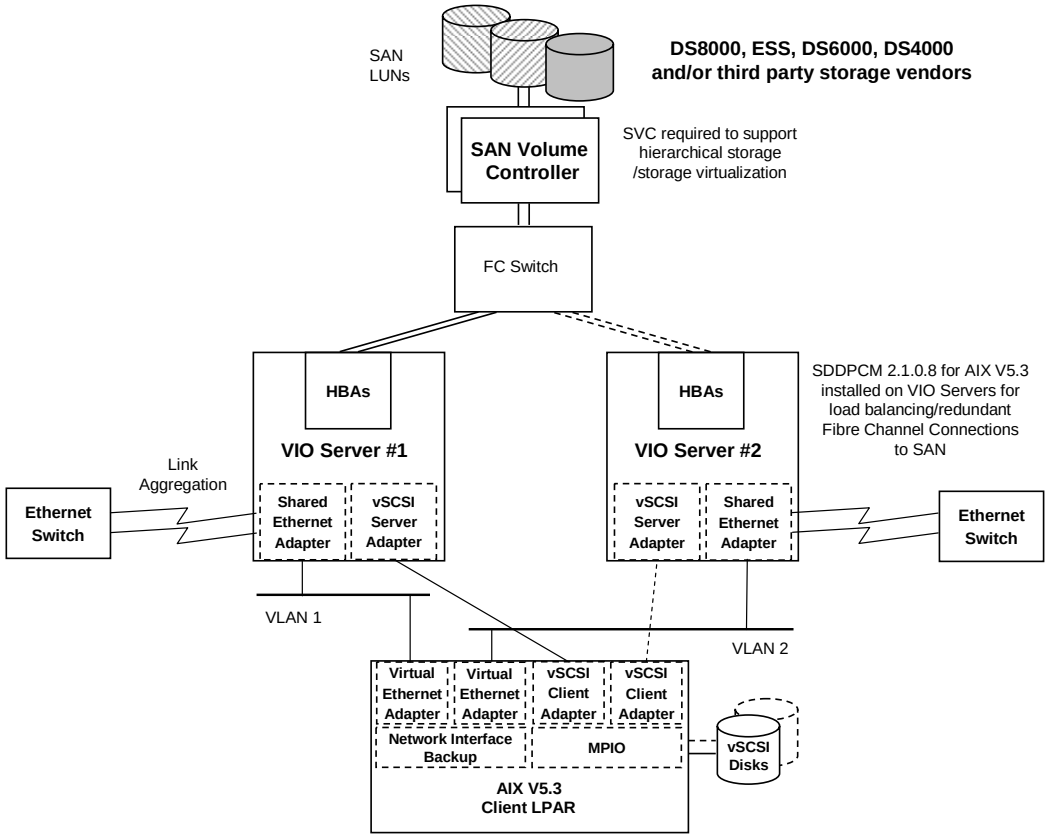

Figure 1: Dual VIO Server Example

The steps required to configure the dual VIO servers and a client partition for this configuration will be covered.

VIO Server Configuration

The following steps would be performed on the first VIO Server and then repeated on the second VIO Server. Both VIO Servers are installed and configured before installation of the first client partition as the VIO Servers are required for disk and network access.

Initial VIO Server Setup

Once the VIO Server partition reboots from the hard disk after installation from DVD media, login as the padmin user to set the new password.

Accept the license for the VIO Server software to continue using the command line using:

```
$ license –accept
```

To check the installed software level:

```
$ ioslevel
```

To check the date, time and time zone:

```
$ date
```

To set the date, time and time zone:

```
$ chdate 101802092006 -timezone CST6CDT
```

Note: If the date/time (for CUT – Central Universal Time) is set correctly on the service processor, the default date and time for any new partition should be correct, The only change required is the time zone. If the time zone is changed, logout and log back in again as the padmin userid.

To eliminate the syslog error message that appears during booting, log into the root shell on the VIO Server and touch the /tmp/syslog.out file.

Shared Ethernet Adapter

The first resource to configure on each VIO Server is the Shared Ethernet Adapter (SEA).

The Shared Ethernet Adapter can use multiple Ethernet ports as an aggregated link to the physical network. The extra network adapters provide additional bandwidth over and above a single network connection and provide some redundancy if there is a port or cable issue with one of the connections. All ports must be connected to the same switch and the switch ports used for aggregation must be configured for EtherChannel®. If using link aggregation, each VIO server will be connected to a different switch for redundancy.

The link aggregation device must be created before creating the Shared Ethernet Adapter.

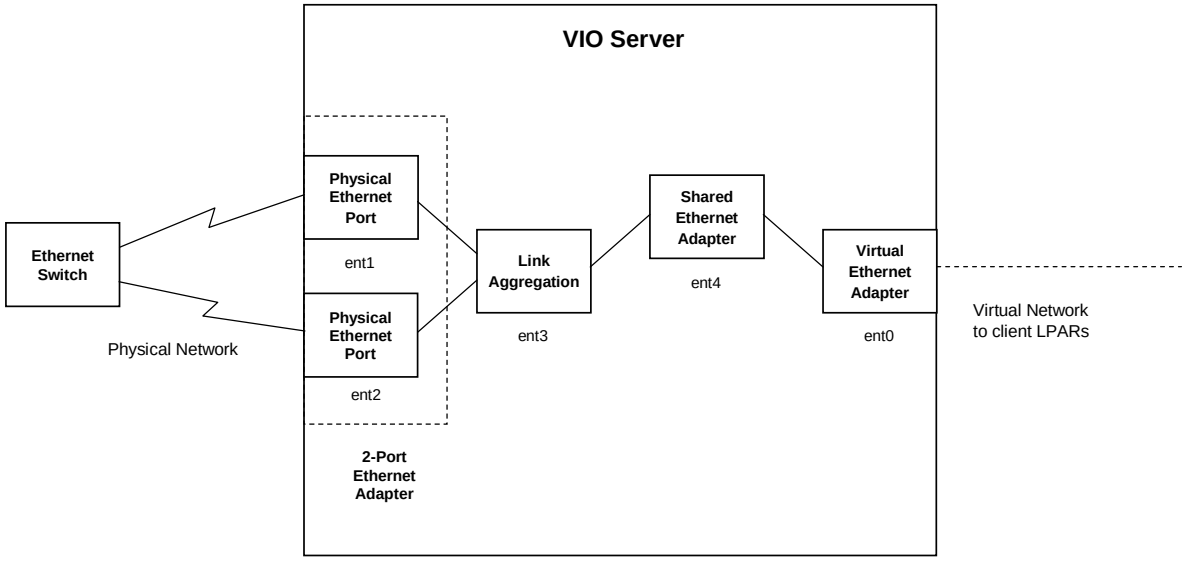

Figure 2: Link Aggregation for Shared Ethernet Adapter

To identify the physical and virtual adapters that will be used for the SEA adapter

```
$ lsdev -virtual
```

To create the link aggregation device from the 2 physical Ethernet ports:

```
$ mkvdev –lnagg ent1 ent2
```

Note: ent0 is the virtual adapter, ent1 and ent2 are the physical Ethernet ports (they could be on the same simgle 2-port adapter).

The next network adapter name, ent3 will be created for the (virtual) link aggregation device.

The new link aggregation device is then mapped to the virtual network adapter to create the Shared Ethernet Adapter (ent4):

```
$ mkvdev –sea ent3 –vadapter ent0 –default ent2 –defaultid 1
```

The defaultid option is 1 as the virtual Ethernet adapter in the partition profile for the first VIO Server is defined to use Virtual VLAN ID 1. The command option is changed to "`-defaultid 2`" on the second VIO Server for connecting to Virtual VLAN ID 2.

The hostname and IP address for the VIO Server is configured on the IP interface (en4) for the Shared Ethernet Adapter (SEA).

```
$ mktcpip –hostname <hostname> -inetaddr <ip_address> -interface en4 –netmask 255.255.255.0 \
–gateway <gateway_ip> -nsrvdomain <domain_name> -nsrvaddr <dns_server> -start
```

The en4 interface is only used for communicating with the VIO Server for administration tasks. The VIO Server is then rebooted so that the new hostname takes effect and the RMC daemons are recycled for DLPAR operations:

```
$ shutdown –restart
```

Once booted, the VIO Server should now be on the network and can be pinged from other systems. DLPAR operations should now be possible if this VIO Server has a network connection with the HMC.

To check the status of all the virtual and physical network adapters that form the SEA:

```
$ entstat –all ent4 | more
```

The second VIO Server can be installed and configured in the same way. The only difference is the Virtual VLAN ID and hostname/IP Address for the second VIO Server.

Mirror the rootvg (for the VIO Server)

On each VIO Server, the root volume group (hdisk0) is extended to include the second SCSI disk (hdisk1), mirrored and the VIO Server is rebooted as a result of running the mirrorios command:

```
$ extendvg –f rootvg hdisk1
$ mirrorios hdisk1
```

All the remaining storage on the VIO server is SAN storage which will be used by the client partitions.

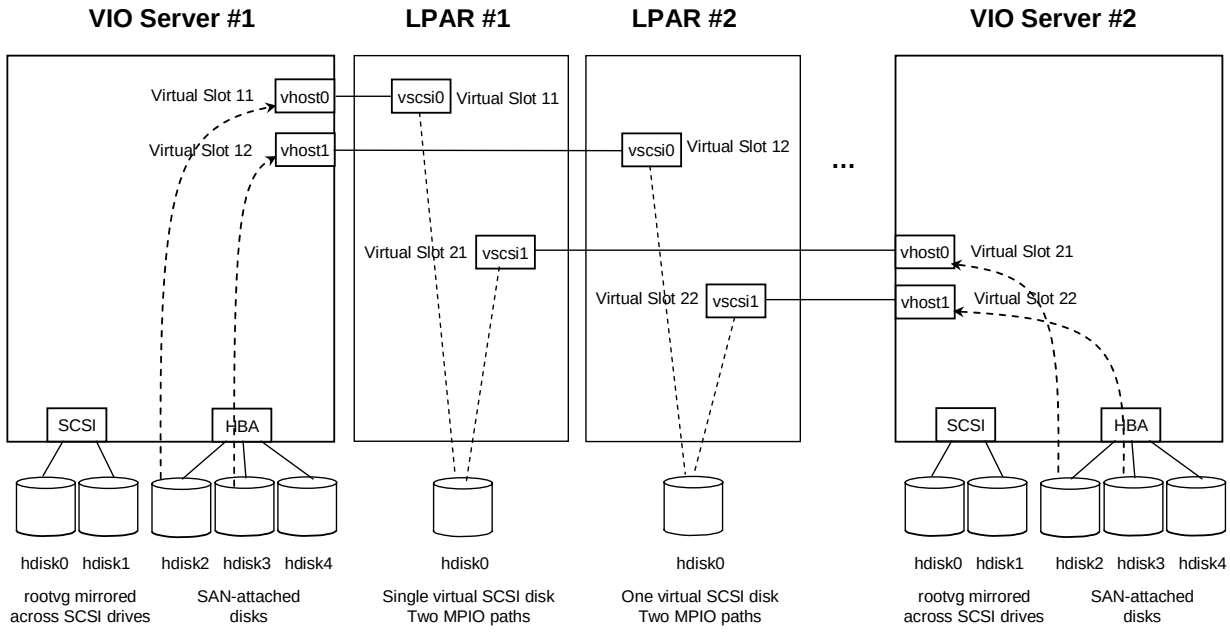

Figure 3: Dual VIO Server Storage Layout

The following steps are required on both VIO Servers to configure the connection to the SAN and the individual drives.

FC (HBA) Device Configuration

(Optional) The init_link attribute for the Fibre Channel adapters can be changed to pt2pt on all HBAs from the default setting of Fibre Channel Arbitration Loop (FCAL) to connect to the Fibre Channel switch ports:

```
$ chdev –dev fcs0 –attr init_link=pt2pt –perm
$ chdev –dev fcs1 –attr init_link=pt2pt –perm
```

This is an optional step as the setting should change to pt2pt when connected to a switch and can be verified using:

```
$ lsdev -dev fcs0 -attr
```

Some SAN administrators want to know that it has been set correctly prior to connection to the FC switch.

FC SCSI Device Configuration

The following attribute changes for all the fscsix devices must also be changed on both VIO Servers:

```
$ chdev –dev fscsi0 -attr fc_err_recov=fast_fail dyntrk=yes –perm
$ chdev –dev fscsi1 -attr fc_err_recov=fast_fail dyntrk=yes –perm
```

Changing the fc_err_recov attribute to “fast_fail” will fail any I/Os immediately if the adapter detects a link event such as a lost link between a storage device and a switch. The “fast_fail” setting is only recommended for dual VIO Server configurations. Setting the dyntrk attribute to “yes” allows the VIO Server to tolerate cabling changes in the SAN. The VIO Servers needs to be rebooted for these changed attributes to take effect.

```
$ shutdown -restart
```

(Optional) The FC SCSI devices can be unconfigured and reconfigured for the settings to take effect, instead of rebooting the VIO Servers.

SDDPCM Installation

For dual VIO Servers attached to IBM SAN storage (SVC, DS8000, DS6000 and ESS), SDDPCM multipathing software for AIX 5L V5.3 is installed on the VIO Servers to provide the load-balancing and redundancy across multiple Fibre Channel connections between the VIO Server(s) and the IBM SAN Storage. The SDDPCM software can be downloaded from the IBM Storage website (see the "Useful Links" section at the end of the document). It has a prerequisite fileset (devices.fcp.disk.ibm.mpio.rte) referred to as the Host Attachment Script for SDDPCM on AIX 5L which is also available from the web. At the time of writing, SDDPCM V2.1.0.8 was the latest version and the required level for dual VIO Servers with the SAN Volume Controller (SVC).

Access the root shell on the VIO Server to install the SDDPCM multi-pathing software:

```
$ oem_setup_env
```

The SAN Volume Controller (SVC) storage is used as an example but the setup is similar for the SDDPCM supported IBM SAN storage products.

If the SVC LUNs have already been zoned to both VIO Servers and the SDDPCM software was not already installed, remove the hdiskx devices for the SVC LUNs (excluding hdisk0 and hdisk1).

Install the following SDDPCM filesets:

```
devices.fcp.disk.ibm.mpio.rte   1.0.0.2  COMMITTED  IBM MPIO FCP Disk Device
devices.sddpcm.53.rte      2.1.0.8  COMMITTED  IBM SDD PCM for AIX 5L V53
```

Once the SDDPCM software is successfully installed, exit from the root user shell.

To configure the SVC MPIO LUNs:

```
$ cfgdev
```

To check the SVC LUNs are identified correctly:

```
$ lsdev –type disk
```

The SVC LUNs will appear as the following devices:

```
hdisk2 Available 08-08-02     SAN Volume Controller MPIO Device
```

To check that the SVC LUNs are visible via both HBAs, look for both adapters (fscsi0 and fscsi1) in the output of the lspath command:

```
$ lspath
```

To check a specific SVC LUN:

```
$ lspath –dev hdisk2

Enabled hdisk2 fscsi0
Enabled hdisk2 fscsi1
```

Note: The SVC appliances are installed in pairs. If there is more than one pair, you will see additional paths on each adapter for each SVC pair.

hdiskx Device Configuration

The other requirement for presentation of a physical drive to a client partition by dual VIO Servers is the reserve_policy attribute is set to no_reserve.

For SVC devices with SDDPCM, the attribute for algorithm is load_balance instead of the typical round_robin for the base AIX 5L MPIO support.

Note: To enable load balancing across multiple HBAs within the VIO Servers when using the base AIX 5L MPIO support, the algorithm must be set to round_robin for each physical drive.

For SVC MPIO devices, algorithm and reserve policy are set correctly by **default**:

```
$ lsdev -dev hdisk2 –attr | more
PCM             PCM/friend/sddpcm               PCM                                 True
…
algorithm       load_balance                    Algorithm                           True
…
reserve_policy  no_reserve                      Reserve Policy                      True
```

Other SAN subsystems may require these attributes to be changed using:

```
$ chdev –dev hdisk2 –attr reserve_policy=no_reserve
$ chdev –dev hdisk2 –attr algorithm=round_robin
```

Virtual SCSI Server Support

For multipathing support of virtual SCSI devices in the AIX 5L client partition, the SAN LUN must be presented as a physical drive (hdiskx) from the VIO Server to the client partition. It is **not** possible to provide a large SAN LUN and then further sub-divide it into logical volumes at the VIO server level when using dual VIO servers. The storage management for this configuration is performed in the SAN so there is a one to one mapping of SAN LUNs on the VIO Servers to virtual SCSI drives on the client partition.

To check which virtual SCSI Server adapter corresponds to which virtual slot number, use:

```
$ lsdev –vpd | grep vhost
```

The slot number will show something like U9117.570.10XXXXX –V51-C11 where Un is the machine type, model number and serial number for the system, V51 is the partition ID and C11 is the virtual slot number.

The SAN LUN (hdiskx) can now be mapped to the respective virtual SCSI Server adapters for use by individual client partitions. It is not required to have a Virtual SCSI Server Adapter for each LUN.

Note: If each VIO Server has a different number of drives or the drives were zoned at different times, the device names (hdiskx) may be different between VIO Servers so check that the LUN IDs match when presenting a drive to the same client partition via dual VIO Servers. It is useful from an administration point of view to have the same device names on both VIO Servers.

There should not be a volume group created on the SAN LUNs on the VIO Server. To map a physical disk to the virtual SCSI Server adapter (vhost0):

```
$ mkvdev –vdev hdisk2 –vadapter vhost0 –dev vtscsi0
```

The default device name for the Virtual Target Device (VTD) is vtscsix. If the -dev option is not specified, the next device number (vtscsix) will be created. The device name can be used to identify different mapping options. Instead of using or defaulting to vtscsix, vtdatax could be used as a naming scheme for mapping the data disks.

To check the virtual SCSI mapping, use:

```
$ lsmap –all
```

To check the mapping for an individual adapter, use:

```
$ lsmap –vadapter vhost0
```

VIO Client Configuration

Once both VIO Servers have been installed, the SEA configured and virtual SCSI resources mapped, it is possible to install the client partitions. The client partition can be installed from media (DVD) or using one of the network connections to a NIM Server via a VIO Server.

Network Interface Backup

Once the client partition reboots from the installation process, the first resource to configure is the Network Interface Backup device and configure TCP/IP.

The Network Interface Backup configuration will use one of the virtual network adapters as the interface to the primary VIO Server and the second virtual network as the backup network connection though the second VIO Server.

To share the network bandwidth that is available on both VIO Servers, half the client partitions should be configured to use ent0 as the primary adapter with ent1 as the backup adapter. The other group of client partitions will use ent1 as the primary adapter with ent0 as the backup adapter.

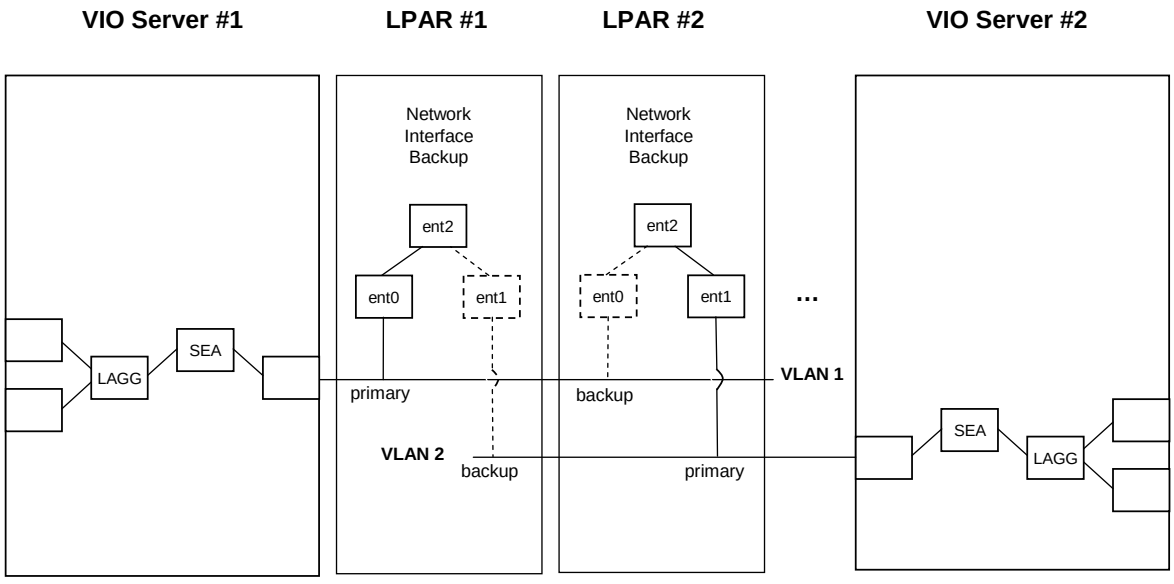

Figure 4: Network Interface Backup for Client Partitions

To identify the slot numbers for the virtual network adapters:

```
# lsslot –c slot
```

To check which virtual network interface is configured:

```
# netstat -rn
```

Tip: If one of the network interfaces (en0 or en1) on the client partition has already been configured as part of a NIM installation, it is necessary to take the interface down as well as remove it so that it is not configured (before the Network Interface Backup interface is configured) on reboot:

```
# ifconfig en0 detach
# rmdev –dl en0
```

Then use SMIT, to configure the Network Interface Backup device using the two virtual network interfaces:

```
# smitty etherchannel
```

To configure the Network Interface Backup, select the correct adapter (ent0) from Available Network Interfaces, ent1 for the Backup Adapter and enter the gateway IP address for the physical network for the Internet Address to Ping.

Configure alternate client partitions to use the second VIO Server for the primary connection to the physical network, by selecting ent1 from the Available Network Interfaces and using ent0 as the Backup Adapter.

To check which adapter is the primary and which is the backup for the Network Interface Backup configuration (ent2):

```
# lsattr -El ent2
```

To check which adapter is currently active for the Network Interface Backup:

```
# entstat –d ent2 | grep “Active channel”
```

To force a failover to the alternate network adapter:

```
# /usr/lib/methods/ethchan_config -f  ent2
```

A quick test of the Network Backup Interface configuration is to force a failover and ping the gateway, and then failback to the primary adapter when finished.

MPIO Path Management

The MPIO support over virtual SCSI between the client partition and the dual VIO Servers **only** supports failover mode. The client partition will use a primary path to one VIO Server and failover to the secondary path to use the other VIO Server. Only one path is used at a given time even though both paths may be enabled.

To balance the load of multiple client partitions across dual VIO Servers, the priority can be set to select the primary path and therefore a specific VIO Server for each client partition. The priority is set on a per virtual SCSI disk drive basis so all the disks or alternate disks on a client partition can use one of the VIO Servers. The recommended method is to divide the client partitions between the two VIO Servers.

To display the MPIO paths to all the virtual SCSI drives:

```
# lspath
```

To display additional MPIO information:

```
# lspath –F”status:name:path_id:parent:connection”
```

To favor vscsi1 as the primary path by lowering the priority of the vscsi0 path, use:

```
# chpath –l hdisk0 –a priority=2 –p vscsi0
```

To check the priority for either path, use:

```
# lspath –E –l hdisk0 –p vscsi0
path 2 Priority True
# lspath –E –l hdisk0 –p vscsi1
path 1 Priority True
```

The health check interval attribute for each virtual SCSI disk that is being used with MPIO to the dual VIO Servers must be changed to enable automatic path failure detection. The default value for the health_interval attribute is 0 (disabled). The health_interval attribute can be changed to 20 seconds for each virtual SCSI disk drive using:

```
# chdev –l hdisk0 –a hcheck_interval=20 –P
```

and the client partitions must be rebooted for the attribute to take effect.

The attributes for the disk should appear as:

```
# lsattr –E –l hdisk0

hcheck_interval 20                      Health Check Interval      True
…
hcheck_mode     nonactive               Health Check Mode          True
```

If the primary VIO Server is shutdown for any reason, the client partition will failover to the secondary VIO Server. The path to the primary VIO Server will be displayed as “Failed”. When the primary VIO Server is booted again, the client partition will detect the path and the path will show as “Enabled”. The client partition will continue to use the secondary VIO Server and if the secondary VIO Server is taken down, it can fail back to the original primary VIO Server if the path is “Enabled”.

If automatic failback to the primary VIO Server is required, install the software support for APAR IY79741 on all the client partitions. With this feature, the client partitions will start using their primary VIO Server when it becomes available after an outage.

To check the status of the paths for a specific virtual SCSI disk drive (hdisk0), use:

```
# lspath –l hdisk0
```

Tip: If the virtual SCSI drives have not been configured with a volume group, the size of the drive can be checked using:

```
# bootinfo –s hdisk1
```

Testing Dual VIO Servers

With any high availability environment, it is important to test the environment and have a clear procedure for recovering from individual failures and minimizing any potential outages. With the two VIO Servers and the first client partition installed and configured, this is the ideal time to test the environment this client partition can continue to operate with either VIO Server.

For testing the network interface backup configuration, shutdown the backup VIO Server and ping a network address on the physical network from the client partition. Activate the backup VIO Server and then shutdown the primary VIO Server. Check that network access to the physical network is still available via the backup VIO Server.

For SAN attachment, shutdown one of the VIO Servers and check that the paths to the virtual SCSI disk drives via one VIO Server are shown as “Failed” and then “Enabled” after the VIO Server is rebooted.

The SAN booting of the client partitions via the dual VIO Servers has eliminated the requirement for LVM mirroring of the rootvg in this configuration. However with LVM mirroring of the rootvg, it is necessary to resync the stale copy of the rootvg if there is an outage of one of the VIO Servers.

The following procedure is required to resync rootvg for the client partitions.

1. Check the status of the rootvg mirror copies using:
   `# lsvg –p rootvg`
2. If either of the hdisks are in “missing” state, use:
   `# varyonvg rootvg`
3. Once the command completes and both drives show as “active”, then resynchronize the mirrors using:
   `# syncvg –v rootvg`

Updating Live VIO Servers

With a number of client partitions using the VIO Server for network and Virtual SCSI support, it is important to maintain these servers without any outage to the client partitions. With dual VIO Servers on the same system, it is possible to take one VIO Server offline and upgrade it.

Here is a procedure to update the VIO Servers one at a time:

1. If using LVM mirroring on the client partitions, check that all client partitions have two good copies of the rootvg.
2. Check that the network connection via the second VIO Server is functioning. Force a failover to the alternate network adapter on one of the client partitions using the first VIO Server and check that the client partition can still ping the network gateway.
3. Create a second profile for the first VIO Server without any virtual adapters (the easiest way is to copy the current profile and just remove the virtual adapters)
4. Shutdown the first VIO Server and the activate it with the new profile
5. Update the first VIO Server with the new fixpack
6. Shutdown the first VIO Server
7. Activate it with the original profile
8. If using LVM mirroring, you will need to resync the rootvg on all clients when the first VIO Server is available.

Then repeat the process for the second VIO Server on the same system.

When both VIO Servers are available, half the client partitions will require a forced failover of the Network Backup Interface to use the second VIO server for network connectivity.

Mapping Client Partition Drives to VIO Server Resources

The HMC V5 interface simplifies the process of correlating the virtual SCSI disks on the client partition to the physical devices on the VIO Server. The steps are included here as a reference for the situations where multiple physical drives or SAN LUNs are mapped to one virtual SCSI server adapter:

1. Login to the client partition

   Identify the vscsi disk drive and display the information about the device using:
   `$ lscfg –l hdisk0`

   The output looks something like:
   hdisk0 U9117.570.10XXXXX-V1-C11-L8100000000000000

   where U9117.570 and 10XXXXX are the machine type.model and serial number
   V1 is the partition ID for this partition

C11 is the slot number for the Virtual SCSI Client adapter
L is the LUN label and 8100000000000000 is the LUN number

Record the slot number (Cx) for the Virtual SCSI client adapter on the partition

Record the Virtual SCSI LUN number

2. Login to the HMC and use the GUI to map the virtual slots for the virtual SCSI adapters from the client partition to the respective VIO Server. This step is eliminated if the same virtual slot numbers have been used between the client partition and the VIO Server.

3. Login to the VIO Server

   Identify the specific adapter for the virtual slot number using:
   ```
   $ lsdev –vpd | grep vhost
   ```

   Display the mapping of the Virtual SCSI Server adapter to the device names:
   ```
   $ lsmap –vadapter vhost0
   ```

   Use the Virtual SCSI LUN number from the client partition to determine the physical backing device on the VIO Server:

	Partition Name	Partition ID	Virtual Slot Number	Mapping	
VIO Client	LPAR_A_1	1	11	Virtual SCSI Device	hdisk0
VIO Server	VIO_A_51	51	11	SVSA	vhost0
				VTD	vtscsi0
				(VSCSI) LUN	0x8100000000000000
				Backing Device	hdisk2 (Note 1)
				Physloc – LUN	Lxxxxxxxxxxxxxxx
VIO Server	VIO_A_52	52	21	SVSA	vhost0
				VTD	vtscsi0
				(VSCSI) LUN	0x8100000000000000
				Backing Device	hdisk2 (Note 1)
				Physloc – LUN	Lxxxxxxxxxxxxxxx

Table 2: Virtual SCSI Disk Mapping

Note 1: The backing device name may be different for the same physical LUN due to the order the devices are discovered on each VIO Server. It is preferable to have the same device names for correlation purposes.

References

Introduction to the Virtual I/O Server Whitepaper
http://www.ibm.com/servers/eserver/pseries/hardware/whitepapers/virtual_io.pdf

Virtual Networking on AIX 5L White Paper
http://www.ibm.com/servers/aix/whitepapers/aix_vn.pdf

Using the VIO Server
http://publib.boulder.ibm.com/infocenter/eserver/v1r3s/topic/iphb1/iphb1.pdf

Virtual I/O Server Commands Reference
http://publib.boulder.ibm.com/infocenter/eserver/v1r3s/topic/iphcg/iphcg.pdf

Introduction to Advanced POWER Virtualization on IBM eServer p5 Servers Redbook
http://www.redbooks.ibm.com/redpieces/abstracts/sg247940.html?Open

Advanced POWER Virtualization on IBM eServer p5 Servers: Architecture and Performance Considerations Redbook
http://www.redbooks.ibm.com/abstracts/sg245768.html?Open

Partitioning Implementations for IBM eServer p5 Servers Redbook
http://www.redbooks.ibm.com/redpieces/abstracts/sg247039.html

POWER5 Virtualization: How to set up the IBM Virtual I/O Server
http://www.ibm.com/developerworks/eserver/library/es-aix-vioserver-v2/

Useful Links

Virtual I/O Server Support Web Site
http://www.software.ibm.com/webapp/set2/sas/f/vios/home.html

SDDPCM for AIX 5L
http://www.ibm.com/support/docview.wss?uid=ssg1S4000201

Host Attachment Script for SDDPCM on AIX 5L
http://www.ibm.com/support/docview.wss?uid=ssg1S4000203

SAN Volume Controller Support for Virtual I/O Server Environments
http://www.ibm.com/support/docview.wss?uid=ssg1S1002575

IBM Corporation
Systems and Technology Group
Route 100
Somers, New York 10589

Produced in the United States of America
April 2006