Virtualization Standards for Business Continuity: Part 6

This is the sixth of a series of articles defining the policies, guidelines, standards, and procedures that provide the foundation of a virtualized environment, thus enabling business continuity, disaster recovery, and high availability, with an emphasis toward Return On Investment (ROI).

The focus of this article will be to describe a virtual SCSI adapter slot numbering standard that will help the system administrator organize and manage these adapters and associated storage on the VIO servers and client LPAR's. As mentioned in the previous article of this series, one of the most difficult tasks in a virtualized AIX environment is the identification and management of virtual SCSI adapters and storage, with their associated client LPAR's. These tasks are more difficult in large environments such as the p590 where there will likely be multiple VIO servers and dozens, if not hundreds of client LPAR's on a single frame. The initial configuration of the virtual SCSI adapters and storage is fairly straight forward, though not simple. The most difficult tasks are adding and deleting storage from existing client LPAR environments and the VIO servers presenting that storage. The system administrator must be able to accurately identify the associations between the virtual SCSI adapters and client LPAR's on all VIO servers. Adding storage to the wrong virtual SCSI adapter on a VIO server will probably not be a catastrophic error, however deleting storage from the wrong adapter on a VIO server will almost certainly be catastrophic to the applications running on the client LPAR.

A Korn Shell script called "mapluns", published in "*AIX Update Journal, May 2006, Issue 127, and June 2006, Issue 128*", provides the system administrator with the ability to accurately identify the associations between virtual SCSI adapters, storage devices, storage frames, LUN ID's, PVID's, client LPAR's, and VIO servers. It will also span multiple VIO servers across multiple frames to provide these associations in high availability environments where storage devices may be configured across multiple client LPAR's.

In the previous article in this series, a range of slot numbers was selected for use as virtual SCSI adapters. This range of slot numbers was between 10 and 499. To illustrate an example virtualized configuration, this article will assume a dual VIO server arrangement on each frame. The purpose of the dual VIO server configuration is to ensure the client LPAR's always have at least one active path to network and storage communications. As was described in previous articles in this series, an even/odd numbering scheme will be used with the virtual SCSI adapter slot numbers to segment them between the VIO servers. Even numbered slot numbers will only be used on VIO servers with even numbered node names, and odd numbered slot numbers will only be used on VIO servers with odd numbered node names. This methodology will assist the system administrator with identifying associations between virtual SCSI adapters and VIO servers.

To illustrate the slot numbering scheme, the example VIO server names from previous articles will be used as shown in table 5.1:

VIO Server Name	***VIO Server Location***	***Managed System Name***	***Frame***
dalapvio00	First VIO Server node on the 1st frame	Server-9119-590-SN12A345B	A
dalapvio01	Second VIO Server node on the 1st frame	Server-9119-590-SN12A345B	A
dalapvio02	First VIO Server node on the 2nd frame	Server-9119-590-SN67D890E	B

VIO Server Name	*VIO Server Location*	*Managed System Name*	*Frame*
dalapvio03	Second VIO Server node on the 2nd frame	Server-9119-590-SN67D890E	B

Table 5.1: Example VIO Server names

For the remainder of this article, the p590 frame identified by the managed system name "Server-9119-590-SN12A345B" will be referred to as frame "**A**" and the p590 frame identified by the managed system name "Server-9119-590-SN67D890E" will be referred to as frame "**B**".

To best illustrate the need for standardized virtual SCSI adapter slot numbering, a couple of example high availability (HA) configurations of client LPAR's will be discussed. The first HA example will consist of an active client LPAR and a hot-standby client LPAR whose only function is to wait for a fail over. The second example will consist of an Active-Active pair of HACMP nodes. The cluster pair nodes will be assumed to exist on separate hardware frames A and B (as shown in table 5.1), and will be serviced by dual VIO servers on each frame (again as shown in table 5.1). Table 5.2 shows the example VIO server nodes and associated HA nodes as defined for illustrating the virtual SCSI adapter slot numbering standards.

Frame	*Node Name*	*Function*
A	dalapvio00	VIO Server
A	dalapvio01	VIO Server
A	dalapora00	HACMP Active Node (Cluster 1)
A	dalapora01	HACMP Active Node (Cluster 2)
B	dalapvio02	VIO Server
B	dalapvio03	VIO Server
B	dalapora50	HACMP Hot-Standby Node (Cluster 1)
B	dalapora51	HACMP Active Node (Cluster 2)

Table 5.2: Example HA node names on multiple frames

Notice in table 5.2 the even/odd numbering of the VIO server node names and the same strategy used for the client LPAR nodes. In the first article of this series, it was recommended that equal numbers of even/odd client LPAR node names be created on each frame. This even/odd numbering strategy is used by many of the scripts that will be presented in this series of articles to evenly distribute the network and storage traffic across the VIO servers and associated hardware adapters.

Virtualization Standards for Business Continuity: Part 6

Experience with administering a virtualized environment has revealed that it is a good idea to create all of the virtual SCSI adapters in the slot range on the VIO servers, at the time the VIO servers are created. This practice eliminates the need for modifying the VIO server profiles to add more virtual SCSI adapters every time a new client LPAR is added or deleted. This also eliminates the need for DLPAR'ing virtual SCSI adapters into or out of a running VIO server. However, the greatest benefit of creating all of the virtual SCSI adapters in the 10 – 499 range on each VIO server, is the associated vhost adapters have a consistent numbering sequence. This consistency is extremely valuable to the system administrator when maintaining and supporting the VIO servers and client LPAR's. A Korn shell script called "mkviolpar" was provided in a previous article in this series that performed the task of creating each of the dual VIO servers utilizing the slot numbering and even/odd numbering schemes to be described in this article.

In order to aid the system administrator in tracking and monitoring the virtual SCSI adapters and associated storage, each numbered group of 10 slots is associated with only one client LPAR, even across multiple frames. This is useful when defining high availability clusters and helps to eliminate confusion. By using this standard, the system administrator is always certain that any numbered group of 10 SCSI adapters belongs to one and only one client LPAR, regardless of which frame it may be configured on. The need for this consistency is difficult to illustrate without assuming the client LPAR's will be HACMP clusters, and the associated application storage will be shared between multiple frames. Table 5.3 provides an example slot numbering scheme defining all slots in several numbered groups of 10, using even/odd numbering between the VIO servers, and between multiple frames for high availability. The dual VIO server arrangement per frame ensures the client LPAR's will always have at least one communication path to storage in the event of a scheduled or unscheduled outage to one of the VIO servers. Table 5.3 contains client LPAR's comprising two HACMP clusters consisting of two nodes each. The first HACMP cluster is an Active-HotStandby cluster composed of nodes "dalapora00" and "dalapora50". The second HACMP cluster is an Active-Active cluster composed of nodes "dalapora01" and "dalapora51".

Frame	***VIO Server Name***	***VIO vhost Adapter***	***Slot Number / Status***	***Adapter Function***	***Client LPAR Cluster Node***
A	**dalapvio00**	**vhost0**	**10 / Active**	**Operating System storage**	**dalapora00**
A	**dalapvio00**	**vhost1**	**12 / Active**	**Application Storage**	**dalapora00**
A	dalapvio00	vhost2	14 / Inactive	Application Storage	dalapora00
A	dalapvio00	vhost3	16 / Inactive	Application Storage	dalapora00
A	dalapvio00	vhost4	18 / Inactive	Miscellaneous Storage	dalapora00
A	dalapvio00	vhost5	20 / Active	Operating System Storage	dalapora01
A	dalapvio00	vhost6	22 / Active	Application Storage	dalapora01
A	dalapvio00	vhost7	24 / Inactive	Application Storage	dalapora01
A	dalapvio00	vhost8	26 / Active	Application Storage	dalapora01
A	dalapvio00	vhost9	28 / Inactive	Miscellaneous Storage	dalapora01

Virtualization Standards for Business Continuity: Part 6

A	**dalapvio01**	**vhost0**	**11 / Active**	**Operating System Storage**	**dalapora00**
A	**dalapvio01**	**vhost1**	**13 / Active**	**Application Storage**	**dalapora00**
A	dalapvio01	vhost2	15 / Inactive	Application Storage	dalapora00
A	dalapvio01	vhost3	17 / Inactive	Application Storage	dalapora00
A	dalapvio01	vhost4	19 / Inactive	Miscellaneous Storage	dalapora00
A	dalapvio01	vhost5	21 / Active	Operating System Storage	dalapora01
A	dalapvio01	vhost6	23 / Active	Application Storage	dalapora01
A	dalapvio01	vhost7	25 / Inactive	Application Storage	dalapora01
A	dalapvio01	vhost8	27 / Active	Application Storage	dalapora01
A	dalapvio01	vhost9	29 / Inactive	Miscellaneous Storage	dalapora01
B	**dalapvio02**	**vhost0**	**10 / Inactive**	**Operating System storage**	**dalapora50**
B	**dalapvio02**	**vhost1**	**12 / Active**	**Application Storage**	**dalapora50**
B	dalapvio02	vhost2	14 / Inactive	Application Storage	dalapora50
B	dalapvio02	vhost3	16 / Inactive	Application Storage	dalapora50
B	dalapvio02	vhost4	18 / Inactive	Miscellaneous Storage	dalapora50
B	dalapvio02	vhost5	20 / Active	Operating System Storage	dalapora51
B	dalapvio02	vhost6	22 / Active	Application Storage	dalapora51
B	dalapvio02	vhost7	24 / Inactive	Application Storage	dalapora51
B	dalapvio02	vhost8	26 / Active	Application Storage	dalapora51
B	dalapvio02	vhost9	28 / Inactive	Miscellaneous Storage	dalapora51
B	**dalapvio03**	**vhost0**	**11 / Active**	**Operating System Storage**	**dalapora50**
B	**dalapvio03**	**vhost1**	**13 / Active**	**Application Storage**	**dalapora50**
B	dalapvio03	vhost2	15 / Inactive	Application Storage	dalapora50
B	dalapvio03	vhost3	17 / Inactive	Application Storage	dalapora50
B	dalapvio03	vhost4	19 / Inactive	Miscellaneous Storage	dalapora50
B	dalapvio03	vhost5	21 / Active	Operating System Storage	dalapora51
B	dalapvio03	vhost6	23 / Active	Application Storage	dalapora51

B	dalapvio03	vhost7	25 / Inactive	Application Storage	dalapora51
B	dalapvio03	vhost8	27 / Active	Application Storage	dalapora51
B	dalapvio03	vhost9	29 / Inactive	Miscellaneous Storage	dalapora51

Table 5.3: Virtual SCSI Adapter Slot Numbering

Table 5.3 shows that slots numbered in the 10-19 range are assigned to all four VIO servers on both frames, even numbered slots on the even numbered VIO servers, and odd numbered slots on the odd numbered VIO servers. The client LPAR's in this first example represent an Active-HotStandby HACMP cluster configuration. The Active HACMP node labeled “dalapora00” would only exist on frame A, and the HotStandby node, labeled “dalapora50”, would only exist on frame B. This arrangement is to provide redundant systems in the event of a scheduled or unscheduled outage. For this purpose, even though the “dalapora00” does not exist on frame B, it's assigned application storage slots are reserved on both frames.

In this example configuration, only one pair of application storage slots (12, 13) are used by the VIO servers to provide access to storage for the client LPAR's. All four VIO servers across both frames are configured to utilize this pair of application storage slots for the purpose of providing both client LPAR's in the cluster access to their shared storage. All application storage would be assigned to the vhost adapters associated with slot 12 on even numbered VIO servers, and slot 13 on odd numbered VIO servers.

When the client LPAR boots, it is presented with the virtual SCSI adapters assigned to it, and the attached disks. Table 5.4 shows the slots associated with each example client LPAR, which correspond to vhost adapters on the VIO servers.

Client LPAR Cluster Node	***Client LPAR Slot Number***	***VIO Server***	***VIO Slot Number / vhost Adapter***	***Adapter Function***
dalapora00	10	dalapvio00	10 / vhost0	Operating System Storage
dalapora00	11	dalapvio01	11 / vhost0	Operating System Storage
dalapora00	12	dalapvio00	12 / vhost1	Application Storage (active)
dalapora00	13	dalapvio01	13 / vhost1	Application Storage (active)
dalapora50	10	dalapvio02	10 / vhost0	Operating System Storage
dalapora50	11	dalapvio03	11 / vhost0	Operating System Storage
dalapora50	12	dalapvio02	12 / vhost1	Application Storage (HotStandby)
dalapora50	13	dalapvio03	13 / vhost1	Application Storage (HotStandby)

Table 5.4: Cluster Nodes, Slot Numbers, and vhost adapters

Since the HACMP configuration is an Active-HotStandby arrangement, the client LPAR “dalapora50”

operates as the hot-standby and only needs access to the storage from the active node, associated with virtual SCSI slot numbers 12, 13.

Table 5.4 also illustrates the consistent vhost adapter numbering sequence between the VIO servers. Observe the slot pairs 12 and 13 are associated with virtual SCSI adapter "vhost1" across all four VIO servers. This consistency is due to the creation of all virtual SCSI adapters in the 10 – 499 range at the time the VIO servers were created.

The next example illustrates an Active-Active HACMP configuration, where both nodes of a cluster provide application services, and each provides fail over capabilities for the other. In this scenario, a pair of virtual SCSI adapters are utilized for application storage on each active node, and since each node acts as a fail over for the other cluster node, each node must also be presented with the application storage virtual SCSI adapters from the other cluster node. Table 5.5 identifies an example slot numbering scheme associated with an Active-Active HACMP cluster where each node provides separate application services.

Frame	*VIO Server Name*	*VIO vhost Adapter*	*Slot Number / Status*	*Adapter Function*	*Client LPAR Allocation*
A	**dalapvio00**	**vhost5**	**20 / Active**	**Operating System Storage**	**dalapora01**
A	**dalapvio00**	**vhost6**	**22 / Active**	**Application Storage**	**dalapora01**
A	dalapvio00	vhost7	24 / Inactive	Application Storage	dalapora01
A	**dalapvio00**	**vhost8**	**26 / Active**	**Application Storage**	**dalapora01**
A	dalapvio00	vhost9	28 / Inactive	Miscellaneous Storage	dalapora01
A	**dalapvio01**	**vhost5**	**21 / Active**	**Operating System Storage**	**dalapora01**
A	**dalapvio01**	**vhost6**	**23 / Active**	**Application Storage**	**dalapora01**
A	dalapvio01	vhost7	25 / Inactive	Application Storage	dalapora01
A	**dalapvio01**	**vhost8**	**27 / Active**	**Application Storage**	**dalapora01**
A	dalapvio01	vhost9	29 / Inactive	Miscellaneous Storage	dalapora01
B	**dalapvio02**	**vhost5**	**20 / Active**	**Operating System Storage**	**dalapora51**
B	**dalapvio02**	**vhost6**	**22 / Active**	**Application Storage**	**dalapora51**
B	dalapvio02	vhost7	24 / Inactive	Application Storage	dalapora51
B	**dalapvio02**	**vhost8**	**26 / Active**	**Application Storage**	**dalapora51**
B	dalapvio02	vhost9	28 / Inactive	Miscellaneous Storage	dalapora51
B	**dalapvio03**	**vhost5**	**21 / Active**	**Operating System Storage**	**dalapora51**

B	**dalapvio03**	**vhost6**	**23 / Active**	**Application Storage**	**dalapora51**
B	dalapvio03	vhost7	25 / Inactive	Application Storage	dalapora51
B	**dalapvio03**	**vhost8**	**27 / Active**	**Application Storage**	**dalapora51**
B	dalapvio03	vhost9	29 / Inactive	Miscellaneous Storage	dalapora51

Table 5.5: Active-Active cluster slot numbers

Notice from tables 5.3 and 5.5 the operating system slot numbers for client LPAR cluster node pairs, “dalapora00/dalapora50” and “dalapora01/dalapora51” are reused between the nodes of a cluster on opposite frames. This appears to violate the policy of assigning unique slot numbers to each client LPAR. The justification for this is because business continuity standards described here assume the operating system storage does not fail over between cluster nodes, therefore the operating system slot numbers can be used for different client LPAR's between frames. Furthermore it makes sense to assign these slot numbers to client LPAR's comprising a cluster pair. It makes it easier for the system administrator to identify the vhost adapters used by a cluster node pair.

Client LPAR	***Client LPAR Slot Number***	***VIO Server***	***VIO Slot Number / vhost Adapter***	***Adapter Function***
dalapora01	20	dalapvio00	20 / vhost5	Operating System Storage
dalapora01	21	dalapvio01	21 / vhost5	Operating System Storage
dalapora01	22	dalapvio00	22 / vhost6	Application Storage (active)
dalapora01	23	dalapvio01	23 / vhost6	Application Storage (active)
dalpaora01	26	dalapvio00	26 / vhost8	Application Storage (standby)
dalapora01	27	dalapvio01	27 / vhost8	Application Storage (standby)
dalapora51	20	dalapvio02	20 / vhost5	Operating System Storage
dalapora51	21	dalapvio03	21 / vhost5	Operating System Storage
dalapora51	22	dalapvio02	22 / vhost6	Application Storage (standby)
dalapora51	23	dalapvio03	23 / vhost6	Application Storage (standby)
dalapora51	26	dalapvio02	26 / vhost8	Application Storage (active)
dalapora51	27	dalapvio03	27 / vhost8	Application Storage (active)

Table 5.5: Active-Active cluster slot numbers

In the example Active-Active HACMP cluster, each node is configured with virtual SCSI adapters providing access to operating system storage, and application storage for it's primary application and the application served by the opposite node.

The Business Continuity policies, guidelines, standards, and procedures to be learned from this article

are as follows:

Policies:

- All systems shall be created according to a business continuity mentality, meaning that reasonable effort will be expended to eliminate single points of failure.
  - Dual VIO Servers on each Frame.
  - Multiple paths to storage presented to each Client LPAR.
  - Multiple systems for manual or automated HACMP fail overs.
- Even / odd numbering schemes shall be implemented to simplify and automate system administration tasks, and to evenly distribute network and storage traffic across multiple resources.
- Operating system storage does not fail over during a scheduled or unscheduled outage.
- Paging space does not fail over during a scheduled or unscheduled outage.
- Application storage space fails over during a scheduled or unscheduled outage.

Guidelines:

- Virtual SCSI adapter slots assigned for operating system storage can be reused between cluster pair nodes on separate frames, since operating system storage does not fail over during a scheduled or unscheduled outage.
- All virtual SCSI adapters using slot numbers in the range of 10-499 should be created at the time the VIO server is created, thus providing consistent slot number / vhost adapter sequencing on all VIO servers.

Standards:

- This article illustrates the even/odd numbering sequence for host names on multiple frames.
- This article provides a standardized slot numbering scheme for virtual SCSI adapters provided by a VIO server.
- This article details multiple HACMP cluster scenarios built in virtualized AIX environments.

Procedures:

- The “mkviolpar” Korn Shell script, provided in part 2 of this series of articles (*AIX Update Journal, February 2006, Issue 136*) is programmed to implement the slot numbering standard described in this article. A Korn shell script called “mklpar” to create the client LPAR's, will be presented in a future article, and will also implement the client LPAR side of this slot numbering scheme.

The next article in this series will provide a naming standard for virtual target devices (VTD), and will introduce the concept of resource groups. The resource group will be the basis of the business continuity, disaster recovery, and high availability naming structures recommended by this series of articles.

Virtualization Standards for Business Continuity: Part 6

Dana French
President, Mt Xia, Inc.
http://www.mtxia.com
615.556.0456