

Virtualization Standards for Business Continuity: Part 1

The purpose of this series of articles is to define the policies, guidelines, standards, and procedures that provide the foundation of a virtualized environment enabling business continuity, disaster recovery, and high availability, with an emphasis toward Return On Investment (ROI).

Virtualization is the next evolution in the recentralization of the data center. In the last few years, as networking costs have decreased and support costs have increased, the advantages of a centralized data center has reemerged. These advantages include cost reductions for many IT resources including hardware, software, facilities, and personnel. Within a centralized data center further cost reductions are realized by reducing the amount of hardware that is required to support today's business functions. Reducing the hardware in turn reduces the costs associated with facilities, networking, support and maintenance.

Virtualization in the AIX environment maximizes the utilization of resources and provides the greatest Return On Investment (ROI) for the associated hardware. It permits the administrator to share network and storage resources between multiple systems while reducing the hardware support requirements.

The current trend in system design is to provide a separate system for each application or instance of an application, such as a database, web application, or financial package. In the past a computer system was contained within a single hardware frame, and the frame consisted of the box, CPU, Memory, boards, adapters, disks, etc. The current IBM hardware platforms permit the system administrator to segment a frame into multiple systems called Logical Partitions (LPAR's). These LPAR's can be assigned all or part of the hardware resources contained in the frame, such as CPU, Memory, and I/O adapters. Each LPAR may be used to host any application that normally runs on a standalone machine, high availability cluster, or in a disaster recovery environment.

The concept of Logical Partitioning (LPAR) permits the segmentation of a single computer frame into multiple systems, however without virtualization it requires dedicated hardware resources for networking and communication with disk storage. With virtualization, the system administrator may share hardware resources between multiple LPAR's. The component that permits virtualized sharing of network and disk storage resources is called the VIO Server, and when implemented, is normally configured on one or more LPAR's within a hardware frame.

The remainder of this series of articles will discuss the design and implementation of a large, multiple data center, virtualized environment using VIO servers and client LPAR's. This discussion will focus on enabling business continuity utilizing this technology.

Partition, Node, and Host Name Standards

Virtualization is based on the capability of defining multiple systems within a single computer frame. These system divisions within a frame are referred to as logical partition's (LPAR). The first step in defining an LPAR is to assign it a name. The partition (LPAR) name is usually the same as the system node name. It is important at this time to distinguish between the node name of a system and it's host name. The node name is associated with an instance of an operating system, where the host name is associated with a network adapter. A system may have many different host names, but only a single node name. Also the node name always remains with that instance of the operating system, whereas a host name may float between adapters within a system, or between systems, or across data centers. All partition, host, and node names should be enterprise wide unique values in order to eliminate conflicts during fail-overs, whether they be planned, unplanned, manual, automated, or part of a disaster

Virtualization Standards for Business Continuity: Part 1

recovery effort. The following is an example of an established standard for creating node names and host names.

The node name shall consist of exactly 10 characters as shown in table 1:

<i>Location Code</i>	+	<i>OS Type</i>	+	<i>Environment</i>	+	<i>Application Code</i>	+	<i>Sequence ID</i>
3 char	+	1 char	+	1 char	+	3 char	+	2 char

Table 1

Example information for details of each component of the node name standard is shown in table 2:

<i>Node Name Component</i>	<i>Number of Characters</i>	<i>Example Values</i>
Location Code	3	dal = Dallas bos = Boston nas = Nashville
OS Type	1	a = AIX l = Linux o = OS/400
Environment	1	a = Acceptance Testing d = Development p = Production t = Testing x = Disaster Recovery
Application Code	3	vio = VIO Server nim = NIM Server sap = SAP mqs = MQ Series ora = Oracle db2 = DB2 ifx = Informix
Sequence ID	2	A two character identifier to distinguish multiple instances of a node type. This two character identifier may contain the following characters: 0-9,A-Z,a-z

Table 2

For any single node, one or more host names may be created to identify all of the various network interfaces. Normally, each node will have a host name that is identical to the node name. As an example, assume an AIX node exists in a Dallas data center location, which is a production system

Virtualization Standards for Business Continuity: Part 1

running Oracle, and it is the first node in the sequence. A node name for this node would appear as:

<i>Location Code</i>	+	<i>OS Type</i>	+	<i>Environment</i>	+	<i>Application Code</i>	+	<i>Sequence ID</i>
dal	+	a	+	p	+	ora	+	01

Table 3

This equates to an example node name of “dalapora01”, and usually, would also be used as a host name with an IP address assigned to it. The point being illustrated here is the node name and the host name of a system are separate entities and should be thought of in that way when designing virtualized and clustered systems.

Now that we have established a node and host naming standard, we can use this to create logical partitions (LPAR's) and partition profiles through the Hardware Management Console (HMC). When building a business continuity environment utilizing virtualization, redundant key components is an important consideration. The VIO servers are key components since they will provide client LPAR's access to network and storage devices. Reducing business function outages are the primary concern when designing a business continuity environment. In order to reduce downtime associated with systems that may be providing critical business functions, dual VIO servers are usually configured on each pSeries frame. Dual VIO servers provide the client LPAR's with I/O redundancy in the event of failure of one of the VIO's. They also permit the system administrator to perform system maintenance on each VIO server without requiring an outage on any client LPAR.

Each client LPAR is subsequently configured to have access to the redundant resources provided by both VIO servers on the frame.

When creating dual VIO servers, the node names should comply with our previously defined standard for node and host names. As an example, consider two pSeries frames that will be used to provide LPAR's configured as HACMP cluster nodes, one node of the cluster on each frame. In this scenario, each frame will host dual VIO servers, thus providing each client LPAR with redundant I/O access to the physical resources. For a two node cluster, one client LPAR will be configured on each frame to provide frame redundancy. The VIO servers themselves, are not configured using HACMP because they are designed to work in redundant pairs, and do not require it. Using the previously defined naming standard, dual VIO servers on each frame might have example partition, node and host names as shown in table 4.

<i>VIO Server Name</i>	<i>VIO Server Location</i>	<i>Frame Name</i>
<i>dalapvio00</i>	First VIO Server node on the 1st frame	Server-9119-590-SN12A345B
<i>dalapvio01</i>	Second VIO Server node on the 1st frame	Server-9119-590-SN12A345B
<i>dalapvio02</i>	First VIO Server node on the 2nd frame	Server-9119-590-SN67D890E
<i>dalapvio03</i>	Second VIO Server node on the 2nd frame	Server-9119-590-SN67D890E

Virtualization Standards for Business Continuity: Part 1

Table 4

When implementing virtualization, it is desirable to distribute the network and storage communication traffic equally across dual VIO servers. Since, at the time of this writing, this is not yet an automated process, it must be configured manually. Later in this series of articles, shell scripts will be provided to aid in the distribution of this traffic. These scripts use the “sequence ID” portion of the node name to make decisions regarding how to divide the traffic. Specifically, they determine whether the node name ends in an even or odd number, then select primary and secondary adapters based on the number. The following table shows examples of how storage traffic would be distributed across the VIO servers using this methodology:

<i>Node Name</i>	<i>MPIO hdisk</i>	<i>Primary Path</i>	<i>Secondary Path</i>
dalapora00	hdisk0	dalapvio00	dalapvio01
“	hdisk1	dalapvio01	dalapvio00
“	hdisk2	dalapvio00	dalapvio01
“	hdisk3	dalapvio01	dalapvio00
“	hdisk4	dalapvio00	dalapvio01
“	hdisk5	dalapvio01	dalapvio00
dalapora01	hdisk0	dalapvio01	dalapvio00
“	hdisk1	dalapvio00	dalapvio01
“	hdisk2	dalapvio01	dalapvio00
“	hdisk3	dalapvio00	dalapvio01
“	hdisk4	dalapvio01	dalapvio00
“	hdisk5	dalapvio00	dalapvio01

Table 5

This table shows the primary communication path for the even numbered hdisk's on the even numbered client LPAR's, is through the even numbered VIO server (actually it is through the even numbered virtual SCSI adapter which will be discussed in a later article, but for now assume the even numbered virtual SCSI adapters are associated with the even numbered VIO server). The secondary path for the even numbered hdisk's on the even numbered client LPAR's is through the odd numbered VIO server. The logic is of course reversed for odd numbered hdisk's.

Using these traffic distribution scripts and extrapolating this node naming standard to dozens of LPAR's on a p590 frame, it becomes apparent that on LPAR's with even numbered node names, “hdisk0” will always have a primary communication path through the even numbered VIO server. Since “hdisk0” usually contains the “rootvg” volume group, it is NOT desirable to have the primary

Virtualization Standards for Business Continuity: Part 1

path of “hdisk0” for all LPAR's on a frame going through the same VIO server. Therefore it is recommended when configuring LPAR's between two frames, do not use all even numbered node names for the LPAR's on one frame, and all odd numbered node names on the other frame. Otherwise, the result of using the traffic distribution scripts, would be that all LPAR's with even numbered node names would use the VIO server with the even numbered node name as the primary path for “hdisk0”. Additionally, all LPAR's with odd numbered node names would use the VIO server with the odd numbered node name as the primary path for “hdisk0”, which is probably undesirable, as shown in table 6.

<i>Frame Name</i>	<i>Node Name</i>	<i>hdisk0 Primary Path</i>	<i>hdisk0 Secondary Path</i>
Server-9119-590-SN12A345B	dalapora00	dalapvio00	dalapvio01
“	dalapora02	dalapvio00	dalapvio01
“	dalapora04	dalapvio00	dalapvio01
“	dalapora06	dalapvio00	dalapvio01
Server-9119-590-SN67D890E	dalapora01	dalapvio03	dalapvio02
“	dalapora03	dalapvio03	dalapvio02
“	dalapora05	dalapvio03	dalapvio02
“	dalapora07	dalapvio03	dalapvio02

Table 6

A more desirable configuration is to evenly distribute the “hdisk0” traffic across the dual VIO servers, which can be easily achieved if both even and odd numbered node names are used on each frame. Table 7 shows a desirable node name configuration for a group of eight Oracle servers across two p590 frames and the distribution of “hdisk0” traffic across dual VIO servers. The primary and secondary paths of all subsequently numbered “hdisk's” are also distributed evenly as previously discussed.

<i>Frame Name</i>	<i>Node Name</i>	<i>hdisk0 Primary Path</i>	<i>hdisk0 Secondary Path</i>
Server-9119-590-SN12A345B	dalapora00	dalapvio00	dalapvio01
“	dalapora01	dalapvio01	dalapvio00
“	dalapora02	dalapvio00	dalapvio01
“	dalapora03	dalapvio01	dalapvio00
Server-9119-590-SN67D890E	dalapora04	dalapvio02	dalapvio03
“	dalapora05	dalapvio03	dalapvio02

Virtualization Standards for Business Continuity: Part 1

<i>Frame Name</i>	<i>Node Name</i>	<i>hdisk0 Primary Path</i>	<i>hdisk0 Secondary Path</i>
“	dalapora06	dalapvio02	dalapvio03
“	dalapora07	dalapvio03	dalapvio02

Table 7

Of course the distribution of storage and network traffic can be manually configured across the dual VIO servers, however this takes a lot of time and effort. Also the shell scripts designed to perform this task can be configured to reverse the logic of how the traffic is distributed, so the administrator can specify the primary and secondary paths, but this means the administrator must keep track of how each LPAR is configured so they can make a determination of how to configure new LPAR's. It is much easier and more efficient to implement a node naming structure that can be used to automate at least a portion of this configuration process, and relieve the administrator from having to monitor and track this information.

In the previously defined node naming standard the sequence ID is defined as a two character identifier, and the list of possible digits are “0-9”, “A-Z”, and “a-z”. This is not quite accurate, to insure the traffic distribution scripts operate properly, the last character of the node and host name must be a digit between 0 and 9.

The partition, node and host naming standard defined in this article can be expanded for use with HACMP to identify the various network adapters required by a cluster. Normally, a node in an HACMP cluster has multiple network adapters for the purpose of providing redundant communication paths in the event of a failure of one or more of the adapters. During the configuration of the cluster each network adapter is assigned one or more IP addresses, and each address is associated with a host name. The IP addresses are referred to by there purpose in cluster such as boot, standby, persistent, heartbeat, or service address. The names given to these addresses usually reflect this purpose by attaching a suffix to the end of the node name.

Table 8 lists example host names used to identify all of the various IP addresses assigned to an HACMP cluster node with multiple network adapters. These host names illustrate a naming scheme, which is part of an HACMP configuration methodology, developed for use with the virtualization standards as described in this series of articles.

<i>Host Name</i>	<i>Description</i>
<i>dalapora01</i>	“Service” host name that is activated when HACMP is active on the node, but does not fail over. This “service” host name always remains on the node of the same name as long as HACMP is active.
<i>dalapora01-bt01</i> <i>dalapora01-bt02</i>	Host names associated with a “boot” IP address that is assigned to each network interface at boot time.
<i>dalapora01-pr01</i>	Host name associated with a “persistent” IP address that is assigned to a network interface at boot time, but does not fail over. This host name is always active on a node regardless of whether HACMP is running or not.

Virtualization Standards for Business Continuity: Part 1

<i>dalapora01-rg01</i>	“Service” host name that is associated with a resource group that provides application resources. This is the host name that will fail over between network adapters, HACMP nodes, and across data centers.
<i>dalapora01-sb01</i> <i>dalapora01-sb02</i>	Host name associated with a “standby” IP address that is assigned to each network interface at boot time, but does not fail over. The IP address associated with this host name is replaced during an IPAT take-over operation.
<i>ddcapora01-hb01</i> <i>ddcapora01-hb02</i>	Host name associated with the “heartbeat” IP addresses, one per network adapter. If these IP addresses are manually configured by the HACMP administrator, then a host name should be assigned to each address. The methodology described in this series of articles does not use these host names.

Table 8

As can be observed in Table 8, the node “ddcapora01” may have many host names, some reside permanently on the node, some are only available when HACMP is active, and still others float between the network adapters, HACMP nodes, or even data centers. The assignment of HACMP IP addresses and host names will be discussed in greater detail in a later article.

Notice the division character between the node name and the suffix is a dash “-”, not an underscore “_”. Although the underscore character is supported by some DNS vendors, it is not RFC compliant.

Therefore, the underscore character should never be used in a host name.

The policies, guidelines, standards, and procedures set forth in this article for your consideration are as follows:

Policies:

- All partition names, node names, and host names shall be enterprise wide unique values.
- The last character of all partition, node, and host names shall be a digit between 0 and 9.
- An even/odd numbering sequence shall be used with VIO Server host names to help the system administrator better identify VIO components.
- Do not use the underscore (_) character in partition, node, or host names, use the dash (-) instead.

Guidelines:

- VIO Server LPAR's should be created using a script to ensure consistency and adherence to standards.
- When defining LPAR's on the HMC, use the same name for the partition, profile, node, and host name.
- When defining nodes of an HACMP cluster, the host name which matches the node name of a system in the cluster should NOT fail over to any other node in the cluster.

Virtualization Standards for Business Continuity: Part 1

Standards:

- This article defined a naming standard for logical partitions, nodes, hosts, and network adapters.

Procedures:

- When defining client LPAR's on multiple frames, do not place all LPAR's with even numbered node names on one frame and odd numbered on another, place both even and odd on each frame.

The next article in this series will discuss creating dual VIO server LPAR's using a standardized method, and provide a shell script for this purpose. This method will establish a recommended configuration for the VIO server LPAR and insure that all VIO servers are build in a consistent, repeatable manner.

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