

System P Hardware Management

03/19/2010, 10:55:42 AM

1.Introduction.....	2
1.1.Terminology.....	2
2.System P hardware discovery.....	3
2.1.Prerequisites.....	3
2.1.1.site table.....	3
2.1.2.Network configuration.....	3
2.1.3.System P Hardware and HMC.....	4
2.1.4.Cleanup BPA/FSP IPs on Service Network and HMC.....	5
2.1.5.Hostname and IP address Planning.....	6
2.2.xCAT DB setup and hardware discovery.....	7
2.2.1.xCAT DB setup with HW Discovery.....	8
2.2.1.1.Setup DHCP server.....	8
2.2.1.2.Power on all the frames and CECs.....	10
2.2.1.3.Update xCAT tables (for new xCAT 2.4 clusters).....	11
2.2.1.4.Discover HMCs/frame/CECs, and define them in xCAT DB.....	14
2.2.2.Permanent ip addresses.....	16
2.3.Limitations.....	19
3.HMC connections to frame/CEC.....	20
3.1.Set proper passwords for BPA/FSP/HMC.....	20
3.2.Run mkhwconn to assign frames/CECs to correct HMCs.....	21
3.3.List frames/CECs from HMC.....	22
3.4.Move frames/CECs to other HMCs (optional).....	22
3.5.rspconfig to update password (optional).....	22
3.6.rspconfig to update frame number (optional).....	23
3.7.rspconfig to update BPA/FSP hostname (optional).....	23

1. Introduction

This cookbook provides information and step-by-step instructions used with hardware discovery by the xCAT Management Node (MN), and setting up connections between HMC to IBM System P machines.

1.1. Terminology

The following terms will be used in this document:

- Frame (BPA) node: in xCAT 2.x, a node with `nodetype` set to *bpa* represents a high end System P server 24 inch frame. For example, here is a frame node:

```
Object name: Server-9458-100-SN0SQIH42-1
groups=bpa,all
hcp=Server-9458-100-SN0SQIH42-1
id=5
mgt=bpa
mtm=9458-100
nodetype=bpa
serial=0SQIH42
```

In above example, the attribute “`nodetype`” indicates this is a frame node, where attribute ‘`id`’ indicates the frame id number; The attribute ‘`mgt`’ indicates the current type of the hardware control point (`hcp`) is *bpa*, and attribute ‘`hcp`’ indicates the `hcp` server node object. For this example the `hcp` is set to itself which is the default environment prior to be managed by any HMC. The attributes ‘`mtm`’ and ‘`serial`’ indicate the machine type model and its serial number of the server.

For lower end System P servers, there is no BPA device contained in a 19 inch frame, so there is no xCAT node object represented for 19 inch System P frame.

- CEC (FSP) node: in xCAT 2.x, has a server node with attribute `nodetype` set to *fsp* which represents a System P CEC. Here is an example of CEC node that exists in a high end System P server:

```
Object name: Server-9118-575-SN0SQIH42
groups=all,fsp
hcp=Server-9118-575-SN0SQIH42
id=6
mgt=fsp
mtm=9118-575
nodetype=fsp
parent=Server-9458-100-SN0SQIH42-1
serial=0SQIH42
```

In above example, the attribute 'nodetype' indicates this is a CEC node; and the 'id' attribute indicates the cage number of this CEC in a 24 inch frame. (This will be set to blank for low end System P machine). The attribute 'parent' indicates the frame BPA server node that this CEC is controlled by. The 'mgt' attribute is initially set to *fsp* and 'hcp' attribute is set to itself as the CEC server node prior to be managed by the HMC.

2. System P hardware discovery

This chapter will introduce how the xCAT MN can discover HMCs, System P frames, and CECs working with xCAT *lssl* command. The System P hardware will be discovered on the xCAT service network, and then added to xCAT database as node attributes.

2.1. Prerequisites

Before performing hardware discovery, users should confirm the following prerequisites:

2.1.1. site table

Make sure the following attributes in "site" table are checked and are properly adjusted to match you xCAT cluster site environment:

- domain
- nameservers
- ntpservers

2.1.2. Network configuration

The xCAT MN needs to be properly connected to the xCAT service subnet which is used with all HMCs, System P frames and CECs being used in the xCAT cluster. This service subnet should be located on a private subnet to allow xCAT MN, DHCP server to communicate with HMCs, BPA (frame), and FSP (CECs) in your cluster. The following is the schematic diagram of the xCAT network topology:

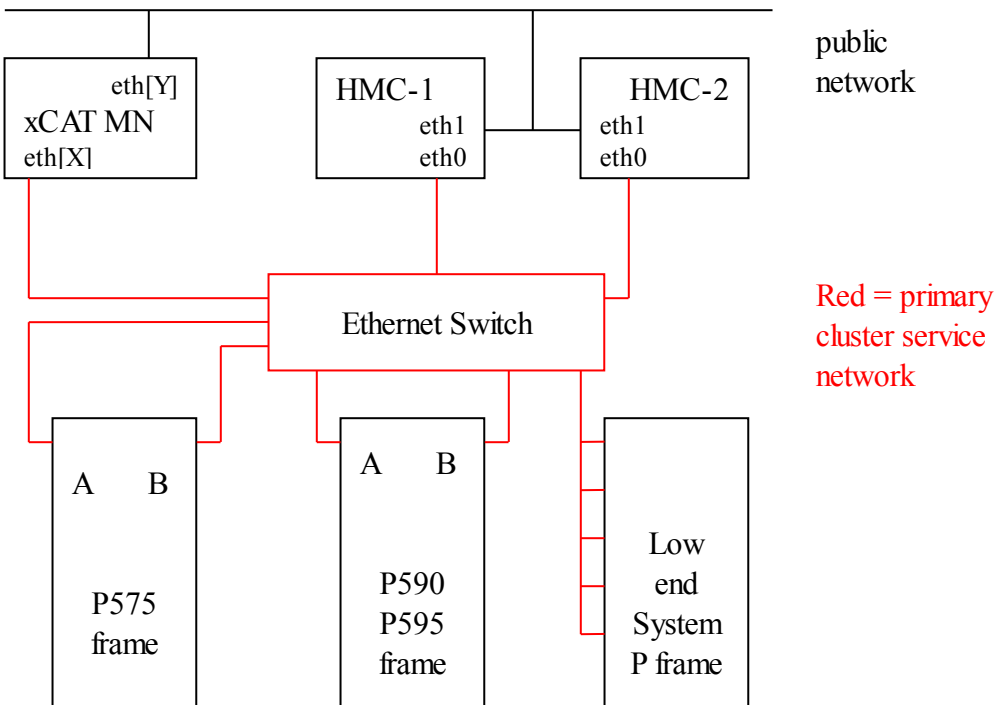


Chart 1.1 Cluster Service network topology

2.1.3. System P Hardware and HMC

- This function currently supports System P hardware (P5,P6,P7) working in xCAT 2.3.4 and xCAT 2.4 releases.
- HMCs should be configured with correct static ip addresses in HW service VLAN, so that they can communicate with xCAT MN. Because the DHCP service is not currently supported on the HMC, (DHCP service normally running on xCAT MN), the DHCP service on HMC should be turned off prior to performing this discovery function. (By default, the DHCP service is disabled for all network interfaces on HMC.)

Reference the HMC website and documentation for more knowledge. The following are minimal steps required to Setup the HMC network for Static IP, and enable SLP and SSH ports working with HMC GUI.

- Open the HMC GUI, Select **HMC Management**, then **Change Network Settings**.
- Select **Customize Network Configuration**, and then **LAN Adapters**.
- Select **Ethernet interface** configured on the service network.
- Click on the **Details** button.
- Select **Basic Settings**, Click on **Open**, and **Specify IP address**. Fill in **IP address**, **Netmask** for HMC static IP on the xCAT service network. Make sure that DHCP Server box is not selected and is blank.

f. Select on **Firewall Settings**, Click on **SLP, Secure Shell**, in the upper window.(You may also want to enable other HMC Firewall settings)

g. Click on the **Allow incoming** button for each required setting.

h. Make sure you Select OK at the bottom of the window to save your updates. Reboot the HMC, and then make sure Network changes are properly working.

- The DHCP service can be run from different server that is connected to the xCAT service VLAN, instead of xCAT MN. In this case, users need to configure the DHCP service manually, and skip the step “Setup DHCP service on MN” (section 2.2).
- The Frame and CEC should be configured as a dynamic IP by default so that the DHCP server can properly assign proper hardware (HW) IP addresses in the xCAT HW service VLAN. If the administrator wants to implement static ip addresses with the BPA/FSP, they must be set in the proper HW VLAN subnet address range specified by the DHCP server. For low end System P servers, only the CEC/FSP need to be configured.

2.1.4. Cleanup BPA/FSP IPs on Service Network and HMC

The xCAT administrator needs to make sure that the BPA/FSP ip addresses and server node names are planned out and are properly defined when working with the xCAT Database and the DHCP environment. There should be no issues if this is a new xCAT System P cluster installation, where the frames and CECs are being specified in the xCAT database and HMC for the first time.

For existing xCAT clusters previously setup with a HMC DHCP server environment where BPA/FSPs are already acknowledged by HMC and xCAT DB, it is important that they use the same existing BPA/FSP network ip addresses and server node names. This includes setting up the DHCP server dynamic address ranges to match the current subnets used by the BPA/FSPs.

If the HW service network requires changes to the BPA/FSP ip addresses, the administrator should plan to cleanup the current BPA/FSP environment. This includes doing cleanup for both the HMC and the xCAT Database for any IP addresses and server node name changes.

For the HMC, the administrator should plan to remove the existing frames and servers that will require new HW IP addresses or server hostnames working in the xCAT HW VLAN. This will allow the xCAT *mkhwconn* command to reinitialize the frame and CECs using the xCAT DB information to make new HW connections to the HMC.

For the xCAT MN, the administrator should review the xCAT DB using *lsdef* and *tabdump* commands to reference any existing HMC/frame/Server node objects that require updates. The xCAT *chdef* command can be used to modify server node attributes, or *rmdef* command if they want to remove the HMC/frame/Server node objects to get to a clean state. It is important that xCAT administrator also clean up Domain Name Service (DNS) and */etc/hosts* file make sure the HMC/frame/Server IP addresses and host names are matching the proper settings required for their xCAT cluster.

2.1.5. Hostname and IP address Planning

For xCAT hardware discovery support, the prerequisite information are the hostnames and ip address allocation being used for all the HMCs, FSPs and BPAs. The mapping between Server hostnames, ip addresses, and physical hardware components need to be defined before proceeding with the xCAT hardware discovery function.

Note: If this is a migration scenario such as migrating from a CSM CRHS environment to xCAT, or is an existing xCAT cluster, where you plan to use the default FSP/BPA host names, i.e., Server-*<Model>-<Type>-<SerialNumber>-<Side>*, then this 2.1.5 "Planning" section is not necessary and can be bypassed by the xCAT administrator.

If the FSPs/BPAs are already using selected static ip addresses, then some steps in the configure DHCP section may not be necessary. All of the HMC connections being used in the xCAT cluster need to be reviewed, and then updated (remove CSM CRHS environment) to make sure the HMC will properly communicate with xCAT MN using the SLP protocol before proceeding with the xCAT hardware discovery.

Each FSP/BPA side of the frame/CEC should have one ip address, and a hostname defined for network communication. The ip addresses for FSPs/BPAs are allocated from the DHCP server which is usually on the xCAT management node (MN). The DHCP service process in the hardware discovery is a little complex, where there are two types of ip addresses allocation mechanisms "random ip addresses" or "permanent ip addresses" that can now be used with xCAT HW discovery.

The DHCP service can be configured to use a dynamic ip range as part of the DHCP configuration file (AIX `/etc/dhcpsd.cnf`, Linux `/etc/dhcpd.conf`), or it can be referenced by a MAC addresses and ip address mapping working with the DHCP leases file (AIX `/etc/db_file_cr`, Linux `/var/lib/dhcpd/dhcp.leases`). When using DHCP dynamic ip range, each DHCP client is initially assigned a random ip address, and this ip address based on it's lease expire information will determine if the DHCP client will require a new hardware ip address. The implementation working with dynamic ip range environment may lead to BPA/FSP ip addresses being changed.

When using the MAC addresses with ip address mapping in DHCP leases file, each DHCP client is assigned a specified permanent HW ip address. If the DHCP client MAC address and ip address mapping is configured in the DHCP leases file, the BPA/FSP will consistently receive the same ip address from the DHCP server.

Random ip addresses: When the FSPs and BPAs are powered up for the first time, the MAC addresses for the FSPs and BPAs are not known by the DHCP server or the admin. We can only work a dynamic ip range in DHCP configuration file, so each FSP or BPA will get an assigned random ip address. As mentioned, the random ip address for each FSP or BPA could change when the DHCP client on FSP or BPA restarts. If we specify a large enough dynamic ip range to avoid the DHCP ip addresses reuse, the dynamic random ip addresses can work similar as "permanent". Please be aware that using random DHCP ip addresses will increase the maintenance effort, because you can not guarantee BPA/FSP server ip address. Using random DHCP ip addresses opens the possibility that FSPs/BPAs ip addresses may be changed during the FSPs/BPAs server reboot. The FSPs/BPAs ip addresses changing will result in HMC connection being lost, where you have to do some administrator steps to recover. The random DHCP ip addresses solution should be able to

work well for most of the scenarios when the proper dynamic range is used on the HW VLAN.

Permanent ip addresses: The implementation of having permanent ip addresses for BPA/FSP as part of DHCP is clearer and easier to maintain. To achieve the permanent ip addresses, we need to specify the proper MAC address and ip address mapping for each FSP and BPA in the DHCP leases file. The implementation of getting the MAC address for each FSP and BPA can be difficult unless the FSPs and BPAs ip addresses and MAC addresses from DHCP server. We can ping the ip addresses and the mac addresses will be added to the arp cache. The majority of the time we will need to work with *lsslp* command to collect the BPA/FSP MAC address from random IP addresses, propagate the permanent IP addresses in xCAT DB to assign the permanent ip with MAC address mapping into DHCP leases file with *makedhcp*.

Both the "Random ip addresses" or "Permanent ip addresses" can be used with HW discovery, the administrator needs to select the proper solution as the BPA/FSP ip addresses assignment working with DHCP

The HMC server host names should be always be resolvable through name resolution (/etc/hosts, DNS). The FSP/BPA server host name and IP addresses should be specified in /etc/hosts files and/or the xCAT hosts table as part of the hardware discovery process.

2.2. xCAT DB setup and hardware discovery

The xCAT administrator has multiple choices for the xCAT HW discovery implementation. One of these choices involve selecting the random or permanent IP addresses as part of the DHCP configuration setup. The second choice is based on how the xCAT administrator wants to implement the HW server information (node name, IP address, host name, MTMS) into the xCAT DB. This is based on the following:

1) Is this a new or an existing xCAT HW cluster.

If this is an existing xCAT cluster, you want to work with current xCAT 2.3 HW discovery environment commands (*lsslp*) to locate new System P hardware. You can move from our current random to the new supported permanent DHCP environment.

If this is a new xCAT cluster, you can manipulate xCAT tables to position the HW IP address being used with BPA/FSP server nodes. You also have a choice to working with the new permanent or existing random DHCP environment.

2) How large and complex is the xCAT HW cluster.

If this is a small cluster, you should be fine working with random IP addresses for the DHCP environment, and can work directly with HW discovery commands.

If this is a larger and more complex cluster, you have the choice working with permanent or random DHCP environment. You also may want to manipulate the HW server data for BPA/FSP using the new xCAT HW discovery methods working directly with xCAT tables to manipulate the BPA/FSP server information.

We will try specify these different choices in our xCAT DB setup and hardware discovery implementation.

2.2.1. xCAT DB setup with HW Discovery

2.2.1.1. Setup DHCP server

1. networks table

All the FSPs and BPAs need to receive their random ip addresses from the DHCP server. The first step is to create xCAT network table object in the xCAT DB working with `mkdef` command for the service HW VLAN network used by xCAT cluster.

Following is an example that provides random IP subnet "192.168.200.0".

```
#netname,net,mask,mgtifname,gateway,dhcpserver,tftpserver,nameservers,  
ntpservers,logservers,dynamicrange,nodehostname,comments,disable  
"RandomIPRange","192.168.200.0","255.255.255.0","en0","192.168.200.205","192.168.200.  
205","192.168.200.205","192.168.200.205","192.168.200.205","192.168.200.1-  
192.168.200.100",,,
```

In the above example, the HW VLAN subnet is from 192.168.200.1-192,168,200.255 where network interface name `en0` on the xCAT MN is used for HW VLAN for both random and permanent ip address. The "192.168.200.1-192.168.200.100" field indicates the random dynamic ip range that is used by the HW service network. The BPA/FSPs will receive HW IP addresses in this dynamic range from DHCP server 192.168.200.205 , which is also the xCAT MN.

There is an xCAT command `makenetworks` that may be executed working with Linux xCAT MN that can populate xCAT networks table, but this command will not specify the `dynamicrange` field. The xCAT administrator needs to manually set the `dynamicrange` attribute in the HW service network object working with xCAT `chdef` command according to their DHCP requirements .

On AIX xCAT MN, command `makenetworks` is not supported yet. All attributes fields in networks table need to be updated by the xCAT administrator working with xCAT `chdef` command to network object.

2. Stop bootp service and initialize DHCP service(AIX only)

For Linux clusters, there is one xCAT command `makedhcp` that will add the `dhcp` service to `inittab (/etc/inittab)` automatically.

For AIX clusters, there is a `bootp` service daemon on the xCAT MN working with AIX node installations. If the xCAT administrator wants to use the DHCP service instead of `bootp` for xCAT node cluster node installations, they will need to disable the `bootp` service. They will need to enable the `dhcpsd` into `rc.tcpip` to start `dhcp` service during system bootup. The AIX NIM installation will enable `bootp` automatically, so the administrator may need to stop the `bootp` daemon, to properly enable `dhcp` server.

The following are the steps used by the xCAT administrator to disable the bootp and enable dhcpd daemon for HW discovery and NIM node installation. The xCAT administrator can toggle between bootp and dhcpd if they want to use bootp as part of the AIX node installation. You just need to be aware that dhcp server daemon for HW discovery will not be active whenever the bootp daemon is active

Disable the bootp service (for AIX only, optional)

a) Stop bootp from rebooting by commenting bootps line in /etc/inetd.conf file:

```
#bootps dgram udp wait root /usr/sbin/bootpd bootpd /etc/bootptab
```

b) Stop and restart the inetd subsystem:

```
refresh -s inetd
```

c) Stop bootp daemon:

```
ps -ef | grep bootp
```

```
kill the bootp process
```

d) Start dhcpd server from rebooting by uncomment dhcpd lines in /etc/rc.tcpip

```
# Start up the DHCP Server
```

```
start /usr/sbin/dhcpd "$src_running"
```

e) Stop and restart the tcpip group

```
stopsrc -g tcpip
```

```
startsrc -g tcpip
```

3. Write static IP address for the network interface running for DHCP service

It is necessary to write a static IP address for the ethernet network interfaces used with DHCP on the xCAT Management Node. This is required so the DHCP server could provide service automatically after DHCP server (xCAT MN) reboot. If the xCAT MN uses a dynamic ip as part of the , this can cause DHCP server issues after reboot.

The following examples represent the DHCP network interface ent0 being used .

For Linux Clusters, write static IP address into network configuration file.

For RHEL, the network configuration is in file like /etc/sysconfig/network-scripts/ifcfg-eth0:

```
DEVICE=eth0
```

```
BOOTPROTO=static
```

```
HWADDR=00:14:5E:5F:20:90
```

```
IPADDR=192.168.200.1
NETMASK=255.255.255.0
NETWORK=50.0.0.0
ONBOOT=yes
```

For SLES, the network configuration is in file like `/etc/sysconfig/network/ifcfg-eth0`:

```
DEVICE=eth0
BOOTPROTO=static
HWADDR=00:14:5E:5F:20:90
IPADDR=192.168.200.1
NETMASK=255.255.255.0
STARTMODE=onboot
```

For AIX Clusters, issue the following commands to write static IP address for network interface:

```
mktcpip -a 192.168.200.1 -i en0 -m 255.255.255.0 -h xCAT_MN
```

4. Generate dhcp configuration file for service nodes and compute nodes

There is xCAT command `makedhcp` that can be used to help to create the DHCP configuration file, and establish the DHCP service daemon working on the xCAT MN.

Run "`makedhcp -n`" to create a new DHCP dynamic IP pool configuration.

Command `makedhcp` with flag `-n` will create the dhcp service configuration file based on attributes found in the xCAT site and networks tables. In this configuration file, the dynamic address range IP pool is created based on the field `dynamicrange` in networks table.

Even if there are no definitions listed in networks table, the `makedhcp` command will try to generate generate DHCP service for all active subnets found on xCAT MN, even though there may be no dynamic IP ranges specified.

The xCAT administrator should examine the DHCP configurations files on the xCAT MN `/etc/dhcpd.conf` (Linux) or `/etc/dhcpsd.cnf` (AIX) after execution of `makedhcp`. If you don't want the DHCP service to serve a subnet, you have to manually comment or remove the subnet definitions in `/etc/dhcpd.conf` (Linux) or `/etc/dhcpsd.cnf` (AIX).

See `makedhcp` man page for details of this command.

2.2.1.2. Power on all the frames and CECs

The system administrator will want to power on the frames/CECs manually after the CE has physically completed the System P server installation and checkout. Since the DHCP server is now configured with a dynamic ip range in the HW service subnet, then all the FSPs and BPAs will get floating HW IP addresses from the DHCP server.

For xCAT clusters, they want to use HW discovery working with existing lsslp command structure that was provided in xCAT 2.3. The administrator can bypass section 2.2.1.3 “Update xCAT tables” with nodelist and switch tables and go directly to section 2.2.1.4 “Discover HMCs/Frame/CECs and define them in xCAT DB.

If you are creating a new large xCAT cluster, you may want to read through the section 2.2.1.3 Update xCAT tables to see if this HW discovery methods meets your needs.

2.2.1.3. Update xCAT tables (for new xCAT 2.4 clusters)

Some information that was mentioned in the “Planning” section can be used to implement into xCAT DB. This includes the system specified node names, the mapping between the ip addresses and hostnames, and the mapping between the hostnames and the physical hardware server components. There are multiple ways to provide the FSP/BPA data in the xCAT DB. We will represent multiple solutions that the xCAT administrator may use to propagate the BPA/FSP data.

For many existing xCAT clusters, they will want to bypass this section, and work directly with the xCAT DB with lsslp command structure. and not work with nodelist and switch tables.

1. Update the nodelist table

The “**nodelist**” table stores all the server hostnames and node group information for all the nodes. Before proceeding with the hardware discovery process, the "nodelist" table needs to be updated to include all the FSPs and BPAs hostnames information.

```
• nodelist table  
#node,groups,status,appstatus,primarysn,comments,disable  
frame1a,"all,bpa,bpa-a"  
frame1b,"all,bpa,bpa-b"  
... ..  
frame16a,"all,bpa,bpa-a"  
frame16b,"all,bpa,bpa-b"  
f1c1,"all,fsp"  
f1c2,"all,fsp"  
... ..  
f16c16,"all,fsp"
```

In the nodelist table, each node may have different status and appstatus in the future, so we could not use "groups" in nodelist table, each node has to be listed separately in the nodelist table. It is time consuming to input all the nodes into nodelist table, the mkdef command can be used to add all the nodes into nodelist table:

```
mkdef frame1a-frame16a groups="all,bpa,bpa-a"  
mkdef frame1b-frame16b groups="all,bpa,bpa-b"
```

```
mkdef flc1-flc16 groups="all,bpa,bpa-b"  
... ..  
mkdef fl6c1-fl6c16 groups="all,bpa,bpa-b"
```

2. Update mapping between the hostnames and hardware components

1). Use “vpd” table information mode

For the high end servers such as POWER 595/575 that exist in 24 inch frames, only the BPA MTMS frame information is necessary, since the CECs (FSPs) will be automatically located by the BPA. In regard for the System P low end servers that exist in 19 inch frames such as POWER 520, the FSP MTMS information is needed. System administrators have options to find the MTMS information for all the BPAs or FSPs. One is to copy the BPA or FSP MTMS information from the front/rear of the Frame and CEC to place this information manually. We recommend that the administrator execute *lsslp* command to discover all the required hardware information. The command below will create all the MTMS output into the stanza file.

```
Run lsslp -s BPA -z -i 10.0.0.1 > /stanza/file/BPApath
```

or

```
Run lsslp -s FSP -z -i 10.0.0.1 > /stanza/file/FSPpath
```

The vpd method uses the MTMS and parent information to determine which hardware components that the server hostname is working from. For high end servers environment, the "vpd" table needs to be updated to include the BPA/FSP MTMS and “side” information for redundant BPA/FSP environments. The expectation is that the majority of System P systems will reference only one FSP for each CEC, then the FSP side information is not required in vpd table. For the low end servers environment, the "vpd" table needs to be updated to only include the FSP MTMS information. All the BPA and FSP MTMS information for the frame/CECs can be referenced and the administrator can make updates to the stanza file */stanza/file/path*. The command below will write the attributes in stanza file */stanza/file/path* into xCAT DB.

```
cat /stanza/file/BPApath > mkdef -z
```

or

```
cat /stanza/file/FSPpath > mkdef -z
```

The vpd table will like the following after writing the MTMS information into xCAT DB.

- vpd table:

```
#node,serial,mtm,side,asset,comments,disable  
"frame1a","99200G1","9A00-100","A",,,  
"frame1b","99200G1","9A00-100","B",,,  
... ..
```

```
"frame16a","99410D1","9A00-100","A",,,  
"frame16b","99410D1","9A00-100","B",,,
```

Note: The frames' MTMS information can be got from the BPA stanza file.

For System P low end servers, the FSP MTMS information is found from the FSP stanza file that was created by *lsslp* command.

For high end CECs, the xCAT "ppc" table also needs to be updated to include the "cage id" information and the BPA "parent" information for each FSP. For low end CECs, the "ppc" table does not need to be updated, following command can help to write the "cage id" into ppc table:

```
chdef flc1 id=1 parent=frame1a  
chdef flc2 id=2 parent=frame1a  
...  
chdef fl6c2 id=2 parent=frame16a
```

The ppc table will look like:

- ppc table:

```
#node,hcp,id,pprofile,parent,supernode,comments,disable  
"flc1",,"1",,"frame1a",,,  
"flc2",,"2",,"frame1a",,,  
... ..  
"fl6c1",,"1",,"frame16a",,,  
"fl6c2",,"2",,"frame16a",,,
```

2). Use the switch table port mode

The switch port method can be used with larger System P xCAT clusters where the administrator has full control of the HW VLAN switch. The administrator specifies which physical network ports are being used to connect to BPA and FSP network connections. It uses the physical ethernet switch connection to determine which hardware component BPA/FSP server is being connected. The "switch" table then is updated to reference the association between the switch port and the BPA/FSP connection. For high end CECs, only BPA frame hostnames need to be added into switch table. For low end CECs, only CEC fsp hostnames need to be added in.

For example:

- switch table

```
#node,switch, port,vlan,interface,comments,disable
"frame1a","10.0.0.254","1",,,,
"frame1b","10.0.0.254","2",,,,
... ..
"frame16a","10.0.0.254","31",,,,
"frame16b","10.0.0.254","32",,,,
```

For high end CECs (FSP) servers, the xCAT "ppc" table needs to be updated to referenced the cage id information being used for each FSP in the target frame. For low end CECs, the "ppc" table does not need to be updated.

- ppc table

```
#node,hcp,id,pprofile,parent,supernode,comments,disable
"f1c1",,"1",,"frame1a",,,
"f1c2",,"2",,"frame1a",,,
... ..
"f16c1",,"1",,"frame16a",,,
"f16c2",,"2",,"frame16a",,,
```

2.2.1.4. Discover HMCs/frame/CECs, and define them in xCAT DB.

The xCAT administrator needs to select if they want to use “random” or “permanent” DHCP environment. The random method is less complex, but also less stable in the support structure for the administrator. This is because we only need to discover the BPA/FSP once, and use the random floating ip addresses working only in the dynamic range for the HW service VLAN. This section will provide the implementation working with random DHCP environment.

The permanent method is more complex for the administrator, but is less stable for the BPA/FSP support structure. We first need to locate the floating ip addresses being used in the dynamic range, but, them we also allocate specified subnet range for the permanent ip addresses in the HW service VLAN.

The xCAT command *lsslp* is used with the HW discovery working the HMC/frame/CECs, to reference the hardware and network information from the DHCP server . It can then write the discovered information into xCAT DB . It can generate output in different format, including RAW, XML and stanza format. We recommend working with the *-z* flag to create stanza files, so the administrator can review the HW data prior to placing in the xCAT DB. The *lsslp* command does support the *-w* flag which can directly update the HW discovery data directly into the xCAT DB if the administrator does not need to make any changes.

See man page of *lsslp* for details.

1. Use stanza file

- 1). Issue *lsslp* to locate the HMC information and write into a HMC stanza file. You will need to execute the *-m* (multicast) flag to reference later supported HMC V7R35x/V7R71 levels.

```
lsslp -m -s HMC -z -i 10.0.0.1 > /hmc/stanza/file
```

- 2). Review the HMC stanza file and make modifications if necessary.

You will need to include the “username” and “password” attributes being used by the target HMC node. Make sure that the HMC host name and ip address is resolvable in the xCAT cluster name resolution (/etc/hosts, DNS).

Write the HMC stanza information into xCAT DB with xCAT command *mkdef*.

```
cat /hmc/stanza/file | mkdef -z
```

- 3). Issue *lsslp* command to reference the BPA (Frame) information and write into the BPA stanza file. If you are using new xCAT 2.4 “vpd” or “switch” table methods, you can use the optional attributes “switchport” or “vpd” with *lsslp* command. If you specify the attribute “-updatehosts”, it will update the xCAT “hosts” table with the server node hostname and ip address.

```
lsslp -s BPA -z -i 10.0.0.1 -M {switchport|vpd} --updatehosts > /bpa/stanza/file
```

- 4). Review the BPA stanza file and make modifications if necessary. You may want to update the hostnames and/or ip addresses to match your planned xCAT configuration.

Write the BPA stanza information into xCAT DB with xCAT command *mkdef*

```
cat /bpa/stanza/file | mkdef -z
```

- 5). Issue *lsslp* to get FSP information and write into the FSP stanza file. If you are using the vpd or switch table method, you can use the optional attributes “switchport” or “vpd” with the *lsslp* command. If you specify the attribute “updatehosts”, it will also update the xCAT hosts table with the server host name and ip address.

```
lsslp -s FSP -z -i 10.0.0.1 -M {switchport|vpd} --updatehosts > /fsp/stanza/file
```

- 6). Review the FSP stanza file and make modifications if necessary. You may want to update the hostnames and/or ip addresses to match your planned xCAT configuration.

Write the BPA stanza information into xCAT DB with xCAT command *mkdef*.

```
cat /fsp/stanza/file | mkdef -z
```

2. Update xCAT database directly with *lsslp*

You should only write directly into the xCAT DB if you are certain that the BPA/FSP server data specified by *lsslp* command is correct. This is used by experienced xCAT administrators, or if they are adding new System P servers into an existing xCAT cluster. The *lsslp -w* flag will update existing xCAT DB data if the new HW discovery finds any BPA/FSP node contentions. The *lsslp -n* flag is used to locate only new found System P BPA/FSP servers during HW discovery. It can work with *-z* stanza, or the *-w* flag but will only reference or add new BPA/FSP servers into the xCAT DB. The xCAT administrator always has the option to use the *chdef* command to add or modify attributes to the HMC/BPA/FSP node objects.

- 1). Issue `lsslp -w` flag to locate HMC information and write into xCAT DB directly. The admin will need to execute the `chdef` command to add missing username and password attributes. Make sure that the HMC host name and IP address is properly made available in the xCAT cluster name resolution (`/etc/hosts`, DNS).

```
lsslp -m -s HMC -w -i 10.0.0.1
```

- 2). Issue `lsslp -w` flag to locate BPA information and write into xCAT DB directly. If this is an existing xCAT cluster, you can use `-n` flag to reference and add only new BPA nodes into the xCAT DB.

```
lsslp -s BPA -n -w -i 10.0.0.1
```

- 3). Issue `lsslp -w` flag to locate FSP information and write into xCAT DB directly. If this is an existing xCAT cluster, you can use `-n` flag to reference and add only new FSPnodes into the xCAT DB.

```
lsslp -s FSP -n -w -i 10.0.0.1
```

2.2.2. Permanent ip addresses

To avoid conflicting between the "random ip addresses" and "permanent ip addresses", two separate subnets are recommended. The following is an example that provides random IP subnet "192.168.200.0" and permanent IP subnet "10.0.0.0" in the xCAT networks table:

Brian ... Having 2 different subnets will not work, because with most xCAT clusters there is only one active ethernet connection to the HW VLAN.. We need to have a solution where there is only one physical HW VLAN being used for both random and permanent IP addresses in the same HW VLAN ..

I will place an alteration below to work with only one subnet.

```
#netname,net,mask,mgtifname,gateway,dhcpserver,ftpserver,nameservers,  
ntpserver,logserver,dynamicrange,nodehostname,comments,disable  
"RandomIPRange","192.168.200.0","255.255.255.0","en0","192.168.200.205","192.168.200.205",  
"192.168.200.205","192.168.200.205",,"192.168.200.205","192.168.200.1-192.168.200.255",,  
"PermanentIPAddress","10.0.0.0","255.255.255.0","en1","10.0.0.1","10.0.0.1","10.0.0.1","10.0.0.1",,  
"10.0.0.1",,,
```

In the above example, network interface name `en0` is used for random ip address, "192.168.200.1-192.168.200.255" field indicates the dynamic IP range used by the service network. The BPA/FSPs will receive dynamic IP addresses in this range from DHCP server 192.168.200.205 (which is the xCAT MN).

Network interface name en1 is used for permanent ip address, "10.0.0.1" is the DHCP server, which is also the xCAT MN. There is no dynamicrange field put in this subnet since the BPA/FSPs should not receive dynamic IP addresses in this range.

This is Brian's suggested HW VLAN. If we want to specify different subnet ranges that would be ok by me . But we need to stay with one HW VLAN.

```
#netname,net,mask,mgifname,gateway,dhcpserver,tftpserver,nameservers,  
ntpserver,logserver,dynamicrange,nodehostname,comments,disable  
"HWVLAN192","192.168.200.0","255.255.255.0","en0","192.168.200.205","192.168.200.  
205","192.168.200.205","192.168.200.205",,"192.168.200.205","192.168.200.1-  
192.168.200.100",,  
"
```

In the above example, network interface name en0 is used for both random and permanent ip address, The HWVLAN192 subnet is using the full ip range subnet 192.168.200.1-192.168.200.255" . The "dynamicrange" field is working with "random" dynamic ip addresses using "192.168.200.1-192.168.200.100". The "permanent" ip range will then use the remaining ip address range "192-168.200.101-192-168.200.255" used by the HW service VLAN. The BPA/FSPs will first receive floating dynamic IP addresses in this range from xCAT MN DHCP server 192.168.200.205. The xCAT administrator will then assign "permanent" ip addresses for the BPA/FSP server nodes working with permanent ip addresses working between 192-168.200.101-192.168.200.255.

The "permanent" ip range support is only available with xCAT 2.4 and later releases. There is a new attribute "--makedhcp" added to xCAT *lsslp* command, and code changes to *makedhcp* command that allows the xCAT administrator to update the DHCP server configuration to now reference the BPA/FSP as permanent ip address settings. This will place an association of the BPA/FSP server hostname, MAC address, and the permanent ip address in the DHCP leases (AIX */etc/db_file.cr*, Linux */var/lib/dhcpd/dhcpd.leases*) files.

There is new xCAT support added to *lsslp -resetnet* command to help update the BPA and FSP interfaces to work with new permanent ip address. This BPA/FSP update activity will reference this new BPA/FSP permanent ip address specified in the xCATDB and DHCP leases file. It will then directly communicate with the current BPA/FSP (dynamic ip address). and update the ip address to now match the permanent ip address without needing to physically power down the frame or CEC. The xCAT administrator does have the option to physically power down the frame or CEC, and when it powers up, it will also reference the permanent ip address that is specified in the DHCP configuration file.

The following is the implementation to be used working with the permanent ip address. It will specify the xCAT commands and administrator activity.

1. Update xCAT DB working with stanza files and makedhcp command

This will specify the command and administrator activity working with BPA and FSP stanza files, for the BPA/FSP permanent ip addresses

1). The discovery of HMC is not used with the BPA/FSP permanent ip address. You should reference all the HMCs based on earlier administrator activity.

2). Issue *lsslp* command to reference the BPA (Frame) information and write into the BPA stanza file . If you are using new xCAT 2.4 “vpd” or “switch” table methods, you can use the optional attributes “switchport” or “vpd” with *lsslp* command. If you specify the attribute “-updatehosts”, it will update the xCAT “hosts” table with the server node hostname and ip address.

```
lsslp -s BPA -z -i 10.0.0.1 -M {switchport|vpd} --updatehosts > /bpa/stanza/file
```

3). Review the BPA stanza file and make modifications if necessary. You should update the hostnames and ip addresses to match your permanent ip addresses as planned xCAT configuration.

Write the BPA stanza information into xCAT DB with xCAT command *mkdef*

```
cat /bpa/stanza/file | mkdef -z
```

4). Issue *lsslp* to get FSP information and write into the FSP stanza file. If you are using the vpd or switch table method, you can use the optional attributes “switchport” or “vpd” with the *lsslp* command. If you specify the attribute “updatehosts”, it will also update the xCAT hosts table with the server host name and ip address.

```
lsslp -s FSP -z -i 10.0.0.1 -M {switchport|vpd} --updatehosts > /fsp/stanza/file
```

5). Review the FSP stanza file and make modifications if necessary. You may want to update in your xCAT configuration.

Write the BPA stanza information into xCAT DB with xCAT command *mkdef*.

```
cat /fsp/stanza/file | mkdef -z
```

The administrator should now validate that the BPA and FSP server node objects have the proper hostname, new permanent ip address, MAC address and other interfaces (dynamic ip address) specified using *lsdef -l* command. You can also use the xCAT *tabdump* command to reference some xCAT tables (vdp, ppc, hosts, nodelist, mac) to see if they contain the permanent ip addresses. The xCAT administrator can execute the *chdef* command to place the proper BPA/FSP server data in the xCAT DB.

6). Issue the *makedhcp -a* command to place the permanent ip addresses, server host names, and MAC address of the BPA and FSP using the xCAT DB into the DHCP leases files AIX */etc/db_file.cr*, or Linux */var/lib/dhcpd/dhcpd.leases*

```
makedhcp -a
```

7). Issue *lsslp -resetnet* to reset physical BPA and FSP to now reference the new permanent ip addresses. This command will reference the “otherinterfaces” attribute to be used to communicate

with the existing BPA/FSP “dynamic ip” . It will then update the BPA or FSP to now use the permanent ip address specified in xCAT DB and DHCP configuration.

```
lsslp -s BPA -resetnet  
lsslp -s FSP -resetnet
```

2. Update DHCP and reset FSP/BPA directly

This will specify the xCAT *lsslp* command working with “--makedhcp” and “--resetnet” attributes for the BPA/FSP permanent ip addresses . It is not recommended to use the -w flag unless with these new attributes unless you are certain that the vpd or switchport tables properly have the proper BPA/FSP data (permanent ip address, server hostname, MAC, address, and other interfaces specified in the xCAT DB .

- 1). The discovery of HMC is not used with the BPA/FSP permanent ip address. You should reference all the HMCs based on earlier administrator activity.
- 2). Execute *lsslp* command with --makedhcp and --resetnet attributes to update the DHCP configuration for the permanent ip address, server name and MAC address with BPA and FSP . This will use the information from xCAT DB and update the DHCP leases files AIX */etc/db_file.cr*, or Linux */var/lib/dhcpd/dhcpd.leases*.

It will then communicate directly to the BPA and FSP using current dynamic ip from “otherinterfaces” where it updates to the permanent ip address based on the xCAT DB and the DHCP leases file. It will reset of the BPA and FSP to use the new permanent ip address without doing a power down.

```
lsslp -s BPA --makedhcp --resetnet  
lsslp -s FSP --makedhcp --resetnet
```

- 3). The *lsslp* command does have the capability to do many HW discovery activities in the same command execution. It is not recommended, but is possible to use -w to update the xCAT DB when using the vpd or switch port tables (--switchport/vpd). You must make sure that all the required BPA/FSP data is properly identified in the target vpd or switchport tables. The *lsslp* command then can update the DHCP configuration (--makedhcp), and reset the BPA/FSP (--resetnet) at one time.

```
lsslp -s BPA -w -i 192.168.200.205 -M {switchport|vpd} --makedhcp --resetnet  
lsslp -s FSP -w -i 192.168.200.205 -M {switchport|vpd} --makedhcp --resetnet
```

2.3. Limitations

- In a cluster that contains a large number of P5 575 machines, the “lsslp” command may not be able to discover all machines. You can reduce this scaling issue with *lsslp* by using the “-t” (retry times) and “-c” (timeout value) flags . For an example:

```
lsslp -s FSP -i 192.168.200.205 -t 5 -c 3000,3000,3000,3000,3000
```

See “lsslp” man page for the details.

- For HMC with V7R350 release, we had experienced some HMC discovery issues "*lsslp -m*" in different layer2/layer3 ethernet switch environments. In this case, the xCAT admin may have to manually create the HMC node objects using xCAT command "*mkdef*".
- If you run xCAT command *lsslp* with flag "-w" to auto discover BPA/FSP and create BPA/FSP nodes in xCAT DB, there are some types of BPA/FSP that cannot respond the user-defined BPA/FSP system names to xCAT. This causes the node name created by *lsslp* to not be consistent to the system name that is known by HMC. This limitation will not block most functions of xCAT. If system admins want to sync the user-defined system names used by the HMC to xCAT DB, please run *rscan* with -u option to update the FSP/BPA node names in the xCAT database. The *rscan -u* command should only be executed after the running of the *mkhwconn* command (described in section 3).
- The *lsslp -resetnet* is currently only supported when working with -s BPA/FSP. We are looking to enhance the *-resetnet* support to work with selected node ranges in a future xCAT release.

3. HMC connections to frame/CEC

The xCAT administrator can setup the xCAT cluster to connect the BPA or FSP to selected HMCs that are located in their xCAT HW service VLAN. They can also setup a security environment with passwords used with the HMC, BPA, and FSP.

3.1. Set proper passwords for BPA/FSP/HMC

The passwords used with BPA/FSP userid 'HMC' needs to be set correctly in xCAT table "ppcdirect" or table "passwd", if the cluster is not going to use the default passwords. Here is an example of table "ppcdirect",

```
#hcp,username,password,comments,disable
"Server-9125-F2A-0262672","HMC","abc123",,
"my_frame_grp1","HMC","abcdefg",,
```

The default passwords of BPA/FSP for userid 'HMC' can also be set in table "passwd" which is used with xCAT cluster, for an example:

```
#key,username,password,comments,disable
"fsp","HMC","ibm4you",,
```

The passwords used with the HMC nodes working with userid "hscroot" is located in the xCAT table "ppchcp" . For an example

```
#hcp,username,password,comments,disable
"c76v1hmc02","hscroot","abc123",,
```

3.2. Run *mkhwconn* to assign frames/CECs to correct HMCs

The *mkhwconn* command allows the xCAT administrator to properly setup the FSP/BPA connection between the HMC and frames/CECs. This command will make the proper connections on the target HMC if the frame/CEC is not already connected. It also makes the proper updates to the xCAT DB working with the “ppc” and “nodehm” tables and with the target FSP/BPA node objects.

For an *mkhwconn* example, run

```
mkhwconn FSP1 -p HMC1 -P <HMC passwd>
```

will result with FSP node *FSP1* to be connected by HMC node *HMC1*.

See *mkhwconn* man page for details of this command

Note:

- Only one HMC can be specified and active in xCAT DB at a time.
- For those high-end System P servers (IH and H servers), the CEC/FSP nodes in the selected frame/BPA are controlled by the BPA node. In this case, the FSP nodes are not allowed to be assigned directly to an HMC node. The BPA nodes should be assigned to the HMC node, and the “*mkhwconn*” will automatically assign all the FSP nodes found in the selected frame to the HMC.

If you need to assign multiple FSP/BPA server nodes to multiple HMC nodes at a given time, you need to first define the BPA/FSP servers to the target HMC connections in the xCAT database. Make sure the xCAT DB tables “nodehm”, “ppc”, and password “ppcdirect” are properly aligned. The FSP/BPA node object must have the “mgt”. “hcp” set to the target HMC, the parent and password for HMC userid is made available. The xCAT administrator can update the xCAT DB attributes for BPA/FSP node objects by using the *chdef* command. You then can run *mkhwconn -t* to setup multiple BPA/FSP nodes to make connections to multiple HMC based on the xCAT DB.

The following is an example of activating a second HMC for BPA/FSP:

Run *chdef* to change the “mgt” (node management type) to “hmc”, and change “hcp” to the corresponding HMC node.

```
chdef frame_group1, fsp_group2 mgt=hmc hcp=hmcl.
```

Here the *frame_group1* includes the selected BPA nodes, and *fsp_group2* includes FSP nodes that do not belong to the BPA nodes in *frame_group1*.

Run *mkhwconn* to have a second *hmcl* be the second and current HMC connections for *frame_group1* and *fsp_group2*.

```
mkhwconn frame_group1, fsp_group2 -t
```

Note: There can be only one active HMC defined in the xCAT DB working with a BPA/FSP server node at this time. The xCAT administrator can toggle between the two HMCs by using the `chdef` command working with the “`mgt`” and “`hcp`” attributes for the target BPA/FSP node.

3.3. **List frames/CECs from HMC**

There is the `lshwconn` command that will provide the current frame/CEC connection data that is specified on a target HMC. This information currently provides the FSP/BPA server node, the FSP/BPA IP address, and the connection status of the BPA/FSP used for the target HMC node.

Run `lshwconn <HMC node>` to locate all BPA/FSP servers on HMC.

See “`lshwconn`” man page for the details.

3.4. **Move frames/CECs to other HMCs (optional)**

There are times when you will want to remove frame/CECs from one HMC and then connect them to a different HMC. You can use the “`rmhwconn`” command to remove the BPA/FSP connection from the current HMC. You can then execute the “`mkhwconn`” command to generate a new connection to a second HMC.

For example, if the attribute “`hcp`” attribute is set to HMC node `hmc1`, working with BPA server node `frame1` running “`rmhwconn frame1`”

This will remove the frame `frame1`, as well as any CECs located in the `frame1` from current HMC `hmc1`. The `rmhwconn` will also reset the xCAT database to be in a state prior to initial HMC connection (`mkhwconn`). The BPA/FSP server node attributes “`mgt`” are set back to `fsp` or `bpa`, and “`hcp`” now points back to parent FSP/BPA server node instead of the HMC node.

Run `mkhwconn` to assign frames/CECs to another HMC, as described in Section 3.2.

3.5. **rspconfig to update password (optional)**

The xCAT admin can run the `rspconfig` command to modify the HMC, admin, and general userid passwords on the BPA/FSP servers. The BPA/FSP servers are pre-set by System P manufacturing using default passwords.

You can use the same password logic for all the System P frames and CECs in your xCAT cluster, or specify unique passwords for HMC, admin, and general for selected BPA or FSP server node. You can only execute one BPA/FSP userid one at a time with the `rspconfig` command in xCAT2.4. The following contains the `rspconfig` changing the HMC userid password used with the BPA and FSP.

```
rspconfig <bpa> HMC_password=abc123
```

```
rspconfig <fsp> HMC_password=abc123
```

3.6. *rspconfig* to update frame number (optional)

The xCAT administrator can run the *rspconfig* command to specify the frame number information when working with 24 inch frames that contain the BPA logic. This information is helpful for large System P clusters where many frames are being used. The *rspconfig* command will allow the xCAT admin to list the current frame number, or can set BPA server node to a specific frame number. The admin can work with the “ppc” table to setup the frame number or execute one BPA server at a time.

```
rspconfig <bpa> frame (list current frame number)
```

```
rspconfig <bpa> frame=4 (change BPA to now be frame 4)
```

3.7. *rspconfig* to update BPA/FSP hostname (optional)

The xCAT administrator can now run the *rspconfig* command to update the hostnames in the BPA or FSP to match the new hostnames that were created in the xCAT DB “hosts table”. The *rspconfig* command can work with selected BPA/FSP nodes based on data found in the xCAT “hosts” table.

```
rspconfig <bpa_node> hostname=*
```

```
rspconfig <fsp_node> hostname=*
```