

System P Hardware Discovery and HMC Connection Setup

01/07/10, 10:55:42 AM

1. Introduction

This cookbook provides information and some 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:

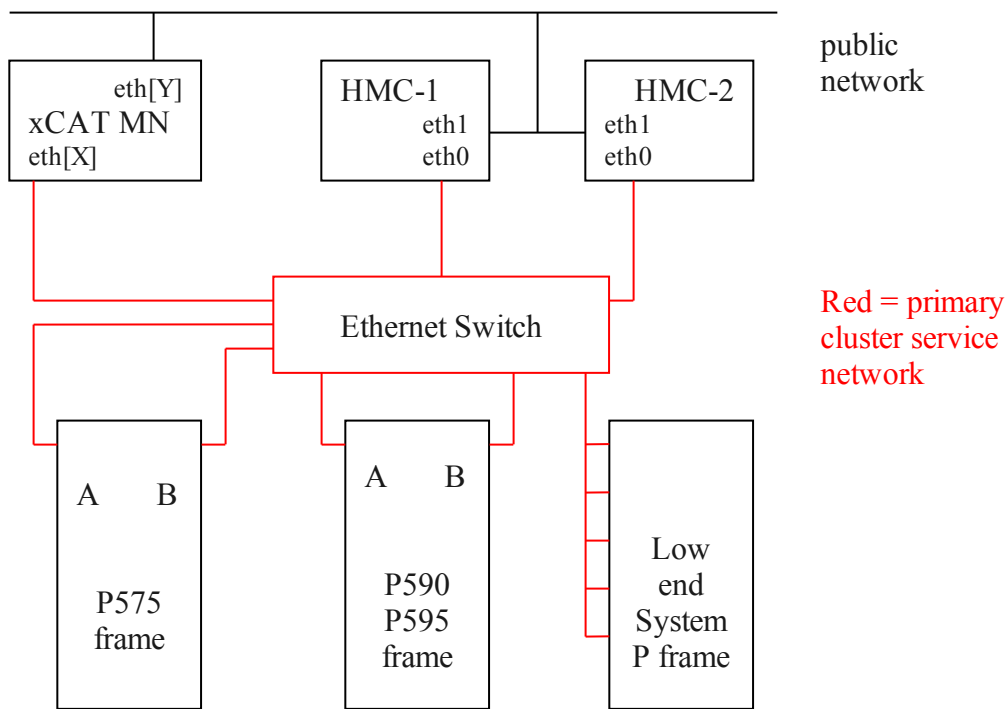


Chart 1.1 Cluster Service network topology

2.1.3 System P Hardware and HMC

- This function currently supports System P5/P6 hardware in xCAT 2.3.
- HMCs should be configured with correct static IP addresses in service subnet, 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.

- a. Open the HMC GUI, Select **HMC Management**, then **Change Network Settings**.
 - b. Select **Customize Network Configuration**, and then **LAN Adapters** .
 - c. Select **Ethernet interface** configured on the service network.
 - d. Click on the **Details** button.
 - e. 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 subnet, 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 to be configured as a dynamic IP so that the DHCP server can properly assign proper IP addresses in the xCAT service subnet. If admin wants to use static IP's with the BPA/FSP, it must be set in the proper dynamic address range specified with the DHCP server. For low end System P servers, only CEC/FSP have 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 properly understood working with the xCAT Database and the DHCP environment. There should be no implementation issues if this is a new System P cluster installation environment where the frames and CECs specified in the xCAT database or HMC for the first time.

For existing xCAT clusters previously setup with standard HMC DHCP 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 nodenames. This includes setting up the DHCP server dynamic address ranges to match the current subnets used by the BPA/FSPs.

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

For the HMC, the administrator should plan to remove the existing Frames and Servers that will require new IP addresses, working in the new subnet. This will allow the xCAT "mkhwconn" command to reinitialize the frame and CECs used by 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.

2.2 Setup DHCP service on xCAT MN.

The DHCP environment is required to properly acknowledge the dynamic IP addresses used with System P servers from the xCAT MN. The xCAT administrator may manually setup the DHCP environment on a separate server in the xCAT service network or use the xCAT MN.

The xCAT DB requires network tables to be created using the "mkdef" command for each subnet that is being used by the xCAT cluster. This includes the network definition used for xCAT service network along with any xCAT cluster VLANs used by nodes and LPARs.

The following is an example that provides information in regard to an xCAT service networks table working with subnet 192.168.200.0 with a subnet mask of 255.255.255.0 .

```
#netname,net,mask,mgtifname,gateway,dhcpserver,tftpserver,nameservers,  
ntpserver,logserver,dynamicrange,nodehostname,comments,disable  
, "192.168.200.0", "255.255.255.0", "en2", "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", ,,,
```

In the above example, "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 (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 fill the field "dynamicrange". The xCAT administrator needs to set the "dynamicrange" attribute in the network definitions 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.

For AIX clusters, there is a "bootp" service used on the xCAT MN working with node installations. If the xCAT administrator wants to use DHCP service instead of "bootp" for xCAT cluster node installations, they will need to disable the bootp service.

Disable the bootp service (for AIX only, optional)

a) Stop bootp daemon:

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

kill the bootp process

b) Stop bootp from restarting by commenting "bootps" line in /etc/inetd.conf file:

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

c) Stop and restart the inetd subsystem:

```
refresh -s inetd
```

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

Run "makedhcp -n" to add a dynamic IP pool.

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 are no dynamic IP ranges.

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.

2.3 Discover HMCs/frame/CECs, and define them into xCAT DB.

The System P discovery allows the xCAT MN to discover any HMC, frame/BPA, or CEC/FSPs found on the xCAT service network. You may ignore the steps specified for frame/BPA if you only have low end System P servers in your xCAT cluster.

Make sure that the DHCP service is currently active in your xCAT cluster. You will want to allocate static IP addresses for HMC, and dynamic IP addresses for BPA and FSPs. You can make sure the HMCs are first active, and then “Power up” or manually reset (power off/power on) frames/CECs in an orderly fashion. The BPA/FSPs will get new dynamic IP addresses from xCAT service DHCP server if the current IP addresses were previously outside of the new dynamic IP address range.

The xCAT command “`lsslp`” executed on the xCAT MN uses SLP protocols to do the discovery for all HMCs, BPAs, and FSPs found in the xCAT service network. The xCAT administrator should first execute “`lsslp`” to examine all the System P servers and HMCs.

The following is an example for `lsslp` command from xCAT MN to discover the HMCs/frames/CECs where ‘192.168.200.246’ is the IP address used for the service network.

```
lsslp -i 192.168.200.246 -s HMC
```

```
lsslp -i 192.168.200.246 -s BPA
```

```
lsslp -i 192.168.200.246 -s FSP
```

Note: To discover an HMC with V7R350 or later release, you need to execute “`lsslp`” with flag “`-m`”, which will have `lsslp` discover HMC node with *multicast* instead of broadcast which is the default. For example:

```
lsslp -i 192.168.200.246 -s HMC -m
```

Refer man page of “`lsslp`” for more details.

After doing the discovery of the HMC/BPA/FSP, you need to update the xCAT DB and the “`/etc/hosts`” file on the xCAT MN to allocate proper server node host names for the selected HMC/BPA/FSP IP addresses. The xCAT administrator can manually do the updates for HMC, BPA, FSP using the xCAT “`mkdef`” command OR

The xCAT administrator can use the xCAT “`lsslp`” with the `-w` flag which works with `/etc/hosts` file and will define the BPA/FSP server nodes to xCAT DB automatically. The “`lsslp`” with `-w` option will not overlay any existing server node attributes in the xCAT DB. It will also update local `/etc/hosts` file on the xCAT MN, if the IP addresses cannot be resolved by `/etc/hosts` or DNS.

The following xCAT tables will be changed to define the discovered nodes:

ppc, vpd, nodehm, Nodelist, nodetype

The SLP response from HMCs/frames/CECs may not contain the host name of target IP, or the host name may be conflicted with the definition in hosts table, or name resolution (/etc/hosts or DNS). For example, a BPA with IP 192.168.200.1 can be given a host name *bpa1* from its ASM webgui, but in local /etc/hosts, 192.168.200.1 indicates a host name *c130fl*. The logic used by lsslp command to specify node host names may be complicated and is the following:

- 1) If an IP address can be found in the xCAT table hosts table (in either simple explicit format, or a regular expression.), use the corresponding nodename. That means this node has already been defined in the xCAT DB. The definitions in hosts table will be used as the first priority.
- 2) If it cannot be found in xCAT DB, but can be resolved by local /etc/hosts or remote DNS. Use the host name in /etc/hosts or DNS. (Regarding if /etc/hosts has higher priority than DNS, it depends on your settings in resolv.conf.)
- 3) If it cannot be resolved in either Step1 or 2, but there is a “name” attribute in SLP response message (for HMC it is always true; for frame it may be set, for CEC, the ‘name’ attribute is always blank). Use it as the node name.
- 4) If it cannot be determine in SLP response message, “lsslp” will specify a factory-default format name like “Server-<mtm>-SN<sn>” for BPA/FSP type nodes. (for example, "Server-9118-575-SN02012EB).

Note:

- For high end System P servers, FSP nodes should be discovered after BPA nodes have been properly defined. The “lsslp” can then update the *ppc.parent* attribute correctly for the FSP nodes controlled by the BPA/frame. (For FSP nodes, *ppc.parent* should be set to BPA node name. See man page of *ppc* table for details.)
- For low end system P servers, BPA is not used, and will not be discovered.

Here is an example of frame/BPA node definition created by ‘lsslp -w’:

```
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
```

Here is an example of CEC/FSP node definition that created by ‘lsslp -w’:

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

2.4 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 192.168.200.246 -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 have to manually create the HMC server node object 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).

3. Setup connections from HMC 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 service network. They can also specify the proper security environment used with passwords to 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 password. 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 assign all the FSP nodes in the selected frame to the HMC automatically.

If you need to assign FSP/BPA server nodes to multiple HMC nodes at a given time, you need to first define the second HMC connection into the xCAT database where `nodehm` and `ppc` tables, and FSP/BPA node object are updated with the `chdef` command. You then can run “`mkhwconn`” with “`-t`” flag to setup multiple BPA/FSP node to create a new connection to second HMC.

The following is an example of activating a second HMC:

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=hmc1.
```

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 `hmc1` be the second and main 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 HMC, BPA, or FSP. 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 onnection (`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 3.2.