

System P Hardware Discovery and HMC Connection Setup

1. Introduction

This cookbook provides step-by-step instructions on hardware discovery and setting up connections from HMC to IBM System P machines.

1.1 Terminology

Following terms will be used in this document:

- Frame (BPA) node: in xCAT 2.x, a node with nodetype '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, attribute nodetype indicates this is a frame node, attribute 'id' indicates the frame number of this frame; attribute 'mgt' equaling 'bpa' and attribute 'hcp' equaling itself mean it has not been managed by any HMC yet; attributes 'mtm' and 'serial' indicate the machine type model and its serial number.

For low end System P servers, since there is no BPA device in a 19 inch frame, no xCAT node object will be represented for the low end System P frame.

- CEC (FSP) node: in xCAT 2.x, a node with nodetype 'fsp' represents a System P CEC. Here is an example of CEC node (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, attribute ‘nodetype’ indicates this is a CEC node; ‘id’ indicates the cage number of this CEC (for low end System P machine, this attribute doesn't make sense, normally is set to blank); ‘parent’ indicates the frame (BPA) node that this CEC belongs to; ‘mgt’ equaling ‘fsp’ and ‘hcp’ equaling itself indicate this CEC node has not been managed by any HMC yet.

2. System P hardware discovery

This chapter will introduce how xCAT MN can discover HMCs, System P frames, and CECs working with xCAT commands. The System P hardware will be discovered on the xCAT service network, and be 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. This service subnet should be located on one large 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:

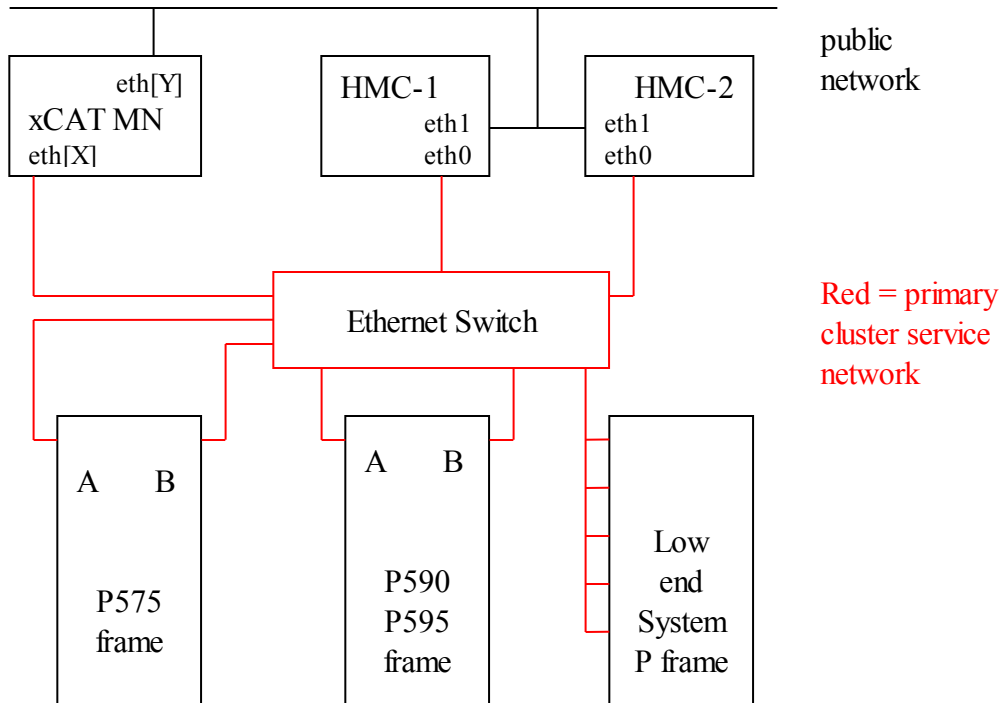


Chart 1.1 Cluster Service network topology

2.1.3 Hardware and HMC

- This function currently supports P5 and 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 will not be supported on the HMC, (DHCP service normally running on xCAT MN), the DHCP service on HMC should be turned off before performing this discovery function. (By default, the DHCP service has been disabled for all network interfaces on HMC.)

Refer following steps to Setup HMC network for Static and enable SLP, SSH 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 another server that is connected to the 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).
 - Frame and CEC should to be configured as DHCP client dynamically from the DHCP server. This is the expected scenario that a cluster admin is facing. (For low end System P servers, only CEC 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 DHCP environment. There should be no implementation issues if this is a new System P cluster installation environment where there are no frames and CECs specified in the xCAT database or HMC.

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

If the service network requires changes to the BPA/FSP IP addresses, the administrator should plan to cleanup the current CEC/FSP environment. This includes doing cleanup for both the HMC and the xCAT Database for any IP and server 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 any changed attributes, or `rmdef` command if they want to remove the HMC/frame/Server objects to get to a clean state. It is important that xCAT administrator also clean up the `/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.

- Setup the xCAT networks table, and add a dynamic IP range.
Here is an example of 'networks' table:

```
#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, the field "192.168.200.1-192.168.200.255" indicates the dynamic IP range used by the service network. The frames and CECs will receive dynamic IP addresses in this range from DHCP server (xCAT MN).

On Linux xCAT MN, issuing xCAT command "makenetworks" can populate xCAT networks table, but this command will not fill the field "dynamicrange". The xCAT administrator needs to set the dynamicrange manually, according their DHCP requirements working with xCAT "chdef" command.

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

- Stop bootp service (for AIX only, optional)

On AIX, the bootp service is running by default on the xCAT MN. You must first stop the bootp daemon working with the following:

- i. Stop bootp daemon:

```
ps -ef | grep bootp  
kill the bootp process
```

- ii. Stop bootp from restarting by commenting out the following line in the /etc/inetd.conf file:

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

- iii. Stop and restart the inetd subsystem:

```
refresh -s inetd
```

- 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 IP pool is created based on the field "dynamicrange" in networks table.

Note, even if there are no definitions listed in networks table, the makedhcp command will generate DHCP service for all subnets found on xCAT MN, even though there is not dynamic IP ranges. 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.

(Note: ignore the steps specified for frame/BPA if you only have low end System P servers in your xCAT cluster.)

- Power on (or reset if they are running already) HMCs/frames/CECs manually. They will get new dynamic IP addresses from xCAT service DHCP server if the current IP addresses are outside of the dynamic IP address range .
- Run lsslp from xCAT MN to discover the HMCs/frames/CECs. For example ('192.168.200.246' is the IP address used on MN to connect to service network).

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

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

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

(Note: To discover HMC with V7R350 or later release, please run lsslp with flag "-m", which will have lsslp to discover node with multicast instead of broadcast. Refer man page of lsslp for details of flag "-m".)

The "lsslp" command will discover HMCs/frames/CECs in service network and will define them into xCAT DB automatically when using the -w flag. The following xCAT tables will be changed to define the discovered nodes:

ppc, vpd, nodehm, Nodelist, nodetype

Since the SLP response from HMCs/frames/CECs may not contain the hostname of target. And even if there is a hostname, it may be conflicted to the definition in hosts table, or name resolution (/etc/hosts or DNS). For example, a frame with IP 192.168.200.1 can be given a hostname named bpa1 from its ASM webgui. But in local /etc/hosts, 192.168.200.1 indicates a hostname c130f1. The logic that lsslp used to determine the discovered node hostnames is as following (a little bit complicated):

- 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. 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 hostname 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, “lsslpl” will specify a factory-default format name like “Server-<mtm>-SN<sn>” for BPA and FSP type nodes. (for example, "Server-9118-575-SN02012EB).

The “lsslpl” command with `-w` option will also update local `/etc/hosts` file on the xCAT MN, if the IP addresses cannot be resolved by `/etc/hosts` or DNS.

Note:

- For high end System P servers, FSP nodes should be discovered after BPA nodes have been properly defined. The lsslpl 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 ‘lsslpl -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 ‘lsslpl -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.

3. Setup connections from HMC to frame/CEC

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 database (“ppc”, “nodehm” tables and FSP/BPA node objects),

For an example, run “mkhwconn FSP1 -p HMC1 -P <HMC passwd>” will result FSP node “FSP1” to be connected to and managed by HMC node “HMC1”.

See mkhwconn man page for details of this command

Note:

- o Only one HMC can be specified and active in xCAT DB at a time.
- o 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 want 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 (nodehm and ppc tables, FSP/BPA node object). You then can run “mkhwconn with “-t” flag to setup multiple BPA/FSP node to create a new connection to second HMC,. Following is an example:

- i. 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
```

3.3 List frames/CECs from HMC

There is the “lshwconn” command that will provide the current frame/CEC connection data that is located on a HMC, BPA, or FSP. This information currently provides the FSP/BPA server node, the FSP/BPA IP address, and the connection status.

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 frame/FSP connection from the current HMC. You can then use the “`mkhwconn`” command to generate a new connection to a second HMC.

- Run `rmhwconn` to detach selected frames/CECs from the HMC that is currently managing them. For example, if the attribute “`hcp`” attribute is set to HMC node “`hmc1`”, running “`rmhwconn frame1`” will remove the frame node “`frame1`”, as well as any CECs in “`frame1`”, from current “`hmc1`”. The `rmhwconn` will also reset the xCAT database to be in a state prior to 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 3.2.