How-To Setup Hierarchy in xCAT 2

10/23/10, 10:30:09 AM

Table of Contents

1.0 Introduction

In large clusters ,it is desirable to have more than one node (the Management Node (**MN**)) handle the installation and management of the compute nodes. We call these additional nodes **service nodes (SN)**. You can have one or more service nodes set up to install & manage groups of compute nodes. With xCAT, you have the choice of either having each service node install a distinct set of compute nodes, or (**on Linux)** having a pool of service nodes, any of which can respond to an installation request from a compute node.

****Important** The xCAT servicenodes must be at the same release/ version as the xCAT Management Node. This version can be checked by running nodels -v.**

```
nodels -v
Version 2.3 (svn r4039, built Mon Aug 24 13:51:18 EDT 2009)
```

This document explains the basics for setting up a hierarchical cluster in xCAT. The user of this document should be very familiar with setting up a xCAT non-hierarchical cluster, and preferably have one that they would like to move to a hierarchical cluster. The document will cover only the additional steps needed to make the cluster hierarchical.

The service nodes need to communicate with the xCAT database on the Management Node and be able to run xCAT commands to install the nodes. The Management Node must run either PostgreSQL, DB2 or MySQL . PostgreSQL, DB2 and MySQL allows a remote client to be set up on the service node, such that the service node can access (read/write) to the database on the Management Node. The default SQLite database does not support this ability. We will show you how to setup and migrate to the new database. The initial install of xCAT will setup SQLite first, so there is always need for migration. The service node will be installed with the xCAT code and other required rpms such as PosgreSQL, perl, rpm. It will require all the dependencies that were setup on the MN. DB2 must be installed and configured manually on the Service Node(s).

There is a different install package xCATsn for the service node vs. xCAT for the Management Node.

2.0 Setup the MN Hierarchical Database

Before proceeding to setup hierarchy on your cluster with this document, you should first read xCATtop for information on downloading and installing xCAT on your Management Node.

The table setup described below is applicable for AIX or Linux unless otherwise indicated, but our example will be Linux.

Once your Management Node is installed with xCAT, you need to setup either MySQL, DB2 or PostgreSQL. MySQL and PostgreSQL are available with the Linux OS. We package MySQL for you in our AIX dependency package. DB2 is an IBM product and must be purchased.

Instructions for setting up a MySQL database for xCAT are found in the xCAT2.SetupMySQL documentation.

Instructions for setting up a PostgreSQL database for xCAT are found in the xCAT2Postgresql documentation.

Instructions for setting up a DB2 database for xCAT are found in the xCAT2SetupDB2 documentation.

If you are on AIX, you should also use the following "Using xCAT Service Nodes with AIX".

2.1 Setup Service Node Access to the Database

Even if you have your database up and running on the Management Node, you probably now need to give the Service Node access to the database, now that you have defined your service node. Each database has a different method to do this. Check the above database documentation, for allowing remote host/ client access to the database.

For the mysqlsetup script, use the mysqlsetup -f < filename> option. See man mysqlsetup.

2.2 Define the Management Node in the Policy Table

As of 2.4 this is automatically done during the initial install of xCAT. Check the policy table to ensure it is already there. If it is not follow these instructions.

Run this command to get the correct management node name known by ssl:

```
openssl x509 -text -in /etc/xcat/cert/server-cert.pem -noout|grep Subject:
this will display something like:
Subject: CN=mgt.cluster
```

Update the policy table with mgt.cluster output from the command:

```
chtab priority=5 policy.name=<mgt.cluster> policy.rule=allow
```

Note: this name must be an MN name that is known by the service nodes.

Make sure the site table has at least the following settings (using tabdump, tabedit, chtab):

```
#key,value,comments,disable
"xcatiport","3002",,
"xcatdport","3001",,
"master","mn20",,
```

where mn20 is the hostname of the management node as known by the service nodes.

Verify the policy table contains at least:

```
#priority,name,host,commands,noderange,parameters,time,rule,comments,disable
"1","root",,,,,,"allow",,
"2",,,"getbmcconfig",,,,"allow",,
"3",,,"nextdestiny",,,,"allow",,
"4",,,"getdestiny",,,,"allow",,
"5","mn20",,,,,,"allow",,
```

2.1 Define the service nodes in the database

This document assumes that you have previously defined your compute nodes in the database and have them setup such that you can run hardware command, such as rpower and can install them. We are

now going to show you how to setup hierarchy to the cluster. You will also be adding all the relevant database data for the service nodes (SN) such that the SN can be installed and managed from the Management Node(MN). In addition you will be adding the information to the database that will tell xCAT which service nodes (SN) will service which compute nodes (CN).

The below tables are setup for Linux. Some attributes are the same for Linux and AIX but to setup the proper table attributes in an AIX cluster refer to the documentation in xCAT2onAIXServiceNodes.

For this example, we have two service nodes: **sn1** and **sn2**. We will call our Management Node: **mn1**. **Note: service nodes are by convention defined with group "service".** Some of the commands in this document will use the group "service" to update all service nodes.

Note: a Service Node's service node is the Management Node; so a service node must have a direct connection to the management node. The compute nodes do not have to be directly attached to the Management Node, only to their service node. This will all have to be defined in your networks table.

2.1.1 Add Service Nodes to the nodelist Table

We assume that you have already defined your compute nodes in the nodelist table. You must also define your service nodes, and by convention we put them in a "service" group. You will define them as you did your compute nodes, except add the group called service. We usually have a group compute* for our compute nodes, to distinguish between the two types of nodes. If you want to use your own group for service nodes, rather than "service", we have setup many defaults in the system based on the group "service". Some of these are the postscripts that will run in the postscripts table. The template, pkglist, etc that will be used. These will have to be modified to use your new group name for "service". It is probably easier just to use "service".

```
chdef -t node -o  sn1,sn2 groups=service,ipmi,all


So for example your nodelist table could look something like this.

 tabdump nodelist
#node,groups,status,appstatus,primarysn,comments,disable
"sn1","service,ipmi,all",,,,,
"sn2","service,ipmi,all",,,,,
"blade01","all,ls21,bc01,blade,compute,compute1",,,,,
"blade02","all,ls21,bc01,blade,compute,compute2",,,,,
"swa01","nortel,switch",,,,,
"bca01","mm",,,,,
```

2.1.2 Add Service Nodes to the nodetype Table

You will have to add your servicenodes to the nodetype table for install as was done with the compute nodes. The profile type of service will allow you to setup the additional rpms required to be installed for the service node. See Build the Service Node Stateless Image.

```
chdef -t group -o service arch=x86_64 os=fedora8 nodetype=osi
 profile=service
```

tabdump nodetype

#node,os,arch,profile,nodetype,comments,disable

"service","fedora8","x86_64","service","osi",,

"compute","fedora8","x86_64","compute","osi",,

2.1.3 Add Service Nodes to the servicenode Table

An entry will be created in the service node table for each service node or the service group. This table describes all the services you would like to setup on that service node or group of service nodes. **All service nodes must be defined in the servicenode table for xCAT**, whether or not you want services setup. If you do not want a service setup then just set the attribute for that service to 0 or "no". Since the **default is "0" or "no"**, it only requires one attribute to be set to define the service node in the table.

You chose the services that you would like started on your service node by setting the attributes in the servicenode table. When the xcatd daemon is started or restarted on the service node, a check will be made by the xCAT code that the services from this table are configured on the SN and running, and will stop and start the service as appropriate.

This check will be done each time the xcatd is restarted on the SN. If you do not wish this check to be done, and the service not to be restarted, use the "reload" option when starting the daemon on the service node.

```
xcatd -r
```

So for example, the following command will setup our service node group (sn1,sn2) to start the nfsserver the DNS nameserver and tftp automatically on the service nodes.

```
chdef  -t group -o service   setupnfs=1  setupnameserver=1 setuptftp=1
```

Note: When using the chdef commands, the attributes in the noderes table are named slightly different than the table attributes in the noderes table. This

```
is to not have a conflict with other tables with the same attribute names.
See  chdef xxxxxx (TODO) for a list of the node attribute valid for chdef.
```

If you do not want any service started on my service nodes, then run the following command to define the service nodes but start no services.

```
chdef  -t group -o service   setupnfs=0
```

If you wanted to start nfs on sn1 and nameserver on sn2, then you could do the following

```
chdef -t node -o sn1 setupnfs=1
chdef -t node -o sn2 setupnameserver=1
```

To show you the results and the other services that could be setup automatically, run

```
tabdump servicenode
#node,nameserver,dhcpserver,tftpserver,nfsserver,conserver,monserver,ldapserver
   ,ntpserver,ftpserver,nimserver,comments,disable
"service","1","1","1","1","1","1",”1”,"1","1",,"Starts all services on all
   service nodes",
```

2.1.4 Add Service Nodes postscripts to the postscripts table

xCAT provides postscripts to automatically install and setup the Service Node (keys, credentials, database access,etc) during the service node install. These postscripts need be added to the postscripts table for the service group, if they are not already there. Here you can also add your own postscripts for customizing the service nodes, by adding them to the table and putting them in /install/ postscripts.

For AIX: Check the following doc:http://xcat.svn.sourceforge.net/viewvc/xcat/xcat-core/trunk/xCAT-client/share/doc/xCAT2onAIXServiceNodes.pdf

```
For Linux:
  chdef -t group -o service postscripts=servicenode,xcatserver,xcatclient

on Linux:
tabdump postscripts
#node,postscripts,comments,disable
"service","servicenode,xcatserver,xcatclient",,
"xcatdefaults","syslog,remoteshell,otherpkgs,syncfiles",,
```

2.1.5 Assigning Nodes to their Service Nodes (updating the noderes table)

The noderes table defines where each node should boot from (xcatmaster), where commands should be sent that are meant for this node (servicenode), and the type of network booting supported (among other things).

The servicenode attribute for a compute node should be set to the hostname of the service node(s) that the management node knows it by. The xcatmaster attribute should be set to the hostname of the service node that the compute node knows it by.

Host name resolution must have been setup in advance, with /etc/hosts, DNS or dhcp to ensure that the names put in this table can be resolved on the Management Node, Service nodes, and the compute nodes.

To keep our example simple, suppose all our nodes in node group compute1 are serviced by sn1 and all the nodes in node group compute2 are serviced by sn2, and the Management node know the service nodes by the names sn1,sn2 and the compute node knows the servicenode by names sn1-c, sn2-c.

Your service nodes must be defined in the noderes table also. The Management Node (mn1) is the xcatmaster and the service node of a service node. There are other attributes in this table for the nodes required for xCAT install explained in other documentation based on your cluster, but they are not relevant to hierarchy.

```
chdef -t node -o compute1 servicenode=sn1 xcatmaster=sn1-c

chdef -t node -o compute2 servicenode=sn2 xcatmaster=sn2-c

chdef -t  node -o sn1,sn2  servicenode=mn1 xcatmaster=mn1
```

```
tabdump noderes
#node,servicenode,netboot,tftpserver,nfsserver,monserver,nfsdir,installnic,
primarynic,cmdinterface,xcatmaster,current_osimage,next_osimage,nimserver,
comments,disable
"compute1",”sn1”,"pxe",,,,,"eth0","eth0",,"sn1-c",,,,,
"compute2",”sn2”,"pxe",,,,,"eth0","eth0",,"sn2-c",,,,,
"sn1","mn1","pxe",,,,,"eth0","eth1",,mn1,,,,,
"sn2","mn1","pxe",,,,,"eth0","eth1",,mn1,,,,,
```

***Note*:** *tabdump -d noderes will give further explanation of each table attribute.*

2.1.5.1 Service Node Pools (not supported in AIX)

As of the xCAT 2.2 release, there is support for Service Node Pools. These are lists of service nodes that support a set of compute nodes. Having a list of service nodes allows backup service node(s) for a compute node when the primary service node is unavailable, or can be used for work-load balancing on the service nodes.

To use Service Node pools, you need to architect your network such that all of the compute nodes and service nodes are on the same flat network. If you don't want the management node to respond to/manage some of the compute nodes, it shouldn't be on that same flat network.

The site.dhcpinterfaces attribute should be set such that the SNs' DHCP daemon only listens on the NIC that faces the compute nodes, not the NIC that faces the MN. This avoids some timing issues when the SNs are being deployed (so that they don't respond to each other before they are completely ready).
To define a list of service nodes that support a set of compute node(s), in the noderes table, in the service node attribute, put a comma-delimited list of the service nodes. The list will be processed left to right, picking the first service node on the list to run the command. If that service node is not available, then the next service node on the list will be chosen until the command is successful. Errors will be logged. If no service node on the list can process the command, then the error will be returned. You can provide some load-balancing by assigning your service nodes as we do below.

```
chdef -t  node -o compute1  servicenode=sn1,sn2
chdef -t  node -o compute2  servicenode=sn2,sn1

tabdump noderes
#node,servicenode,netboot,tftpserver,nfsserver,monserver,nfsdir,installnic,
primarynic,cmdinterface,xcatmaster,current_osimage,next_osimage,nimserver,
comments,disable
"compute1",”sn1,sn2”,"pxe",,,,,"eth0","eth0",,,,,,,
"compute2",”sn2,sn1”,"pxe",,,,,"eth0","eth0",,,,,,,
"sn1","mn1","pxe",,,,,"eth0","eth1",,mn1,,,,,
"sn2","mn1","pxe",,,,,"eth0","eth1",,mn1,,,,,
```

Note: the noderes table's xcatmaster, tftpserver,nfsserver attributes should be blank for any node entry that has the noderes servicenode attribute set to a pool of service nodes.

Verify your setup by running the following commands:

To list the attributes of the service group, run:

```
lsdef -t group -l service
```

To list the attributes of the compute1 group , run:

```
lsdef -t group -l compute1
```

2.1.6 Setup Site Table

If you are **not** using the NFS-hybrid method of stateless booting your compute nodes, set the installloc attribute to “/install”. This instructs the service node to mount /install from the management node. (If you don't do this, you have to manually sync /install between the management node and the service nodes.)

```
chtab key=installloc site.value=/install
```

If using service node pools:

The site table sharedtftp attribute must be set to 0 or "no". The service nodes should not automount the tftpboot directory from the management node when using pools. The default is 1, the service node will automount the tftpboot directory when not using pools.

```
chdef -t site -o clustersite  sharedtftp=0
```

Note: *If your service nodes are stateless and site.sharedtftp=0 (required), if you reboot any service node when using servicenode pools, any data written to the local /tftpboot directory is lost. You will need to run nodeset for all of your compute nodes again.*

```
[root@xcat20RRmn bladedb]# tabdump site
#key,value,comments,disable
"nameservers","9.129.47.250,9.129.8.1",,
"ntpservers","9.10.228.45, 9.10.225.159",,
"defserialport","0",,
"defserialspeed","19200",,
"installloc","/install",,
"dhcpinterfaces","mn1|eth0;sn1|eth1;sn2|eth1",,
"forwarders","9.129.8.1,9.129.8.2",,
"consoleondemand","yes",,
"xcatdport","3001",,
"xcatiport","3002",,
"tftpdir","/tftpboot",,
"installdir","/install",,
"master","mn1",,
"sharedtftp","0",,
"domain","cluster.net",,
"timezone","America/New_York",,
"nameserver","9.129.8.1,9.129.47.250",,
```

2.1.7 Setup networks Table

All networks in the cluster must be defined in the networks table. When xCAT was installed, it ran makenetworks, which created an entry in this table for each of the networks the management node is on. You will need to update the entry for the network the management node uses to communicate to the service nodes, and create one for each network the service nodes use to communicate to the compute nodes.

You can have xCAT ignore any table entry by setting the **disable** attribute. For example, if you have a public network defined, and you want to disable the entry for the public network (connected to the outside world):

```
chtab net=9.114.88.160 networks.netname=public networks.disable=1
```

Note:*If using service node pools, the networks table dhcpserver attribute can be set to any single service node in your pool. The networks tftpserver, and nameserver attributes should be left blank.*

2.1.8 Setup nodehm/noderes Table for conserver/monserver

These tables will allow you to specify which service node should run the conserver and monserver daemon for the nodes assigned to the service node.

```
chdef -t node -o compute1  conserver=sn1 monserver=sn1

chdef -t node -o compute2  conserver=sn2 monserver=sn2
```

2.1.8.1 Conserver and Monserver and Pools

The support of conserver and monserver with Service Node Pools is still not supported. You must explicitly assign these functions to a service node using the nodehm.conserver and noderes.monserver attribute as above.

2.1.9 Verify the Tables

To verify that the tables are set correctly, run lsdef on a service node, compute1, compute2:

```
lsdef service,compute1,compute2
```

3.0 Install or Stateless Boot the Service Nodes on Linux

The service node must contain not only the OS, but also the xCAT software and it's dependencies. In addition, a number of files are added to the service node to support the PostgreSQL, DB2 or MySQL database access from the service node to the Management node, and ssh access to the nodes that the service nodes services. The following sections explain how to accomplish this.

The below installation process for service nodes is for Linux. For the installation of service nodes in an AIX cluster refer to the documentation in xCAT2onAIXServiceNodes.

3.1 Build the Service Node Stateless Image

On Linux, we recommend that you use stateless service nodes, but if you want to have diskfull, statefull service nodes instead, skip this section and follow section 3.2, Set Up the Service Nodes for diskfull Installation.

Note: this section assumes you can build the stateless image on the management node because the service nodes are the same OS and architecture as the management node. If this is not the case, you need to build the image on a machine that matches the service node's OS/architecture.

1. Check the service node packaging to see if it has all the rpms required. We ship a basic requirements lists to bring up a service node, but you may want to add additional operating system packages.

```
cd /opt/xcat/share/xcat/netboot/fedora/
vi service.pkglist
vi service.exlist
```

Edit service.exlist and verify that nothing is excluded that you want on the service nodes.

While you are here, edit compute.pkglist and compute.exlist, adding and removing parckages as necessary. Ensure that the pkglist contains bind-utils so that name resolution will work during boot.

Note:*If you change the files, you should copy them to the /install/custom/netboot/fedora directory, so that the next install of xCAT will not wipe out your changes. The code will look in the /install/custom directory first.*

2. Add additional packages including xCAT and xCAT dependencies required on the Service Nodes. The Service Nodes require the same dependency packages needed for the Management Node.

   Notice we are installing xCATsn, the servicenode meta package.

   use the otherpkgs function to install the additional software:
   (see xCAT2-updatenode documentation):

   First, go to the xCAT Download site in the section titled **RPMs in tarball - download tarball ...** Download the level of xCAT tarball you desired. Go to the xCAT Dependencies Download page and get the latest xCAT dependency tarball. Download and copy the files to the Management Node (MN) and untar them:

```
  mkdir -p /install/post/otherpkgs/fedora8/x86_64/xcat
  cd /install/post/otherpkgs/fedora8/x86_64/xcat

  tar jxvf core-rpms-snap.tar.bz2
  tar jxvf xcat-dep-2*.tar.bz2
```

   Second, add rpm names into the service.<osver>.<arch>.otherpkgs.pkglist file. In most cases, this file is already created under /opt/xcat/share/xcat/netboot/<os> directory. If it is not, you can create your own by referencing the existing ones.

```
  vi /install/custom/netboot/fedora/service.ferdora8.x86_64.otherpkgs.pkglist
```

   add the following:

```
   -OpenIPMI-tools
   -tftp-server
   xcat/xcat-core/xCATsn
   xcat/xcat-dep/fedora8/x86_64/conserver
```

   where the '-' means remove the rpm first before installing other rpms. For SLES, the packages to be removed is perl-doc.

3. Run image generation:

Note: we assume you have run copycds to place the OS image in /install/fedora8/...

```
rm -rf /install/netboot/fedora8/x86_64/service
cd /opt/xcat/share/xcat/netboot/fedora/
./genimage -i eth0 -n tg3,bnx2 -o fedora8 -p service
```

4. For SLES 11, make sure that atftp rpm from xCAT is installed in the image not the atftp from the OS.

```
rpm –root  /install/netboot/fedora8/x86_64/service/rootimg  -qa | grep atftp
```

If the rpm is from then OS then remove and install the xCAT atftp (atftp-0.7.1.*rpm) into the otherpkgs directory with the xCAT rpms as you did above and run genimage again.

```
rpm –root  /install/netboot/fedora8/x86_64/service/rootimg -e <atftp rpm>
   --nodeps

cp atftp-0.7.1.*rpm /install/post/otherpkgs/fedora8/x86_64
```

add the rpm name into:

```
  /install/custom/netboot/fedora/service.otherpkgs.pkglist

run genimage to pick up the additional package.

cd /opt/xcat/share/xcat/netboot/fedora/
./genimage -i eth0 -n tg3,bnx2 -o fedora8 -p service
```

5. Prevent DHCP from starting up until xcatd has had a chance to configure it:

```
chroot /install/netboot/fedora8/x86_64/service/rootimg chkconfig dhcpd off
chroot /install/netboot/fedora8/x86_64/service/rootimg chkconfig dhcrelay off
```

6. Edit fstab:

```
cd /install/netboot/fedora8/x86_64/service/rootimg/etc/
cp fstab fstab.ORIG
```

Put in fstab:

```
proc            /proc                proc         rw 0 0
sysfs           /sys                 sysfs        rw 0 0
devpts          /dev/pts             devpts       rw,gid=5,mode=620 0 0
service_x86_64  /                    tmpfs        rw 0 1
none            /tmp                 tmpfs        defaults,size=10m 0 2
```

```
none            /var/tmp              tmpfs       defaults,size=10m 0 2
```

7. IF using NFS hybrid mode, export /install read-only in service node image:

```
cd /install/netboot/fedora8/x86_64/service/rootimg/etc
echo '/install *(ro,no_root_squash,sync,fsid=13)' >exports
```

8. Pack the image

```
packimage -o fedora8 -p service -a x86_64
```

9. To diskless boot the service nodes

```
nodeset service netboot
rpower service boot
```

3.1.1 Update Service Node Stateless Image

To update the xCAT software in the image at a later time:

```
yum --installroot=/install/netboot/fedora8/x86_64/service/rootimg update
   '*xCAT*'
packimage -o fedora8 -p service -a x86_64
```

OR

```
use updatenode otherpkgs function.
see section 3.1.2 for downloading the xCAt tarballs.

Be sure to gen and repack the image:

cd /opt/xcat/share/xcat/netboot/fedora/
./genimage -i eth0 -n tg3,bnx2 -o fedora8 -p service

packimage -o fedora8 -p service -a x86_64
```

Finally diskless boot with the new xCAT software:

```
nodeset service netboot
rpower service boot
```

Note: *The service nodes are set up as NFS-root servers for the compute nodes. Any time changes are made to any compute image on the mgmt node it will be necessary to sync all changes to all service nodes. In our case the /install directory is mounted on the servicenodes, so the update to the compute node image is automatically available.*

3.2 Set Up the Service Nodes for diskfull Installation

Note: *If you are using diskless service nodes as described above, skip this section.*

First, go to the xCAT Download site in the section titled **RPMs in tarball - download tarball ...** Download the level of xCAT tarball you desired. Go to the xCAT Dependencies Download page and get the latest xCAT dependency tarball.
Download and copy the files to the Management Node (MN) and untar them:

```
mkdir -p /install/post/otherpkgs/fedora8/x86_64/xcat
cd /install/post/otherpkgs/fedora8/x86_64/xcat

tar jxvf core-rpms-snap.tar.bz2
tar jxvf xcat-dep-2*.tar.bz2
```

Second, add rpm names into the service.<osver>.<arch>.otherpkgs.pkglist file. In most cases, this file is already created under /opt/xcat/share/xcat/netboot/<os> directory. If it is not, you can create your own by referencing the existing ones.

```
vi /install/custom/install/fedora/service.ferdora8.x86_64.otherpkgs.pkglist
```

add the following:

```
-OpenIPMI-tools
-tftp-server
xcat/xcat-core/xCATsn
xcat/xcat-dep/fedora8/x86_64/conserver
```

where the '-' means remove the rpm first before installing other rpms. For SLES, the packages to be removed is perl-doc.

Third, add the otherpkgs postscript to the service node postscripts table entry.

```
chdef -t group -o service
   postscripts=otherpkgs,servicenode,xcatserver,xcatclient
```

Finally install the service node:

```
nodeset service install
rpower service boot
```

3.2.1 xCAT Service Node RPMS and dependencies

Download the latest xCAT for your release from the following location:

http://xcat.sourceforge.net/yum/download.html

Note: you will be installing the xCAT Service Node rpm xCATsn* metapackage on the Service Node, not the xCAT* Management Node metapackage. Do not install both.

Download the latest xCAT dependencies for your release from the following location:

https://sourceforge.net/projects/xcat/files/

3.2.2 Update Service Node Diskfull Image

Obtain the new xCAT and xCAT dependencies rpms. Follow the same steps that were followed in Set Up the Service Nodes for diskfull Installation.

Update the new xCAT service node rpms:

```
updatenode service -S
```

3.3 Monitor install and boot

```
wcons service      # make sure DISPLAY is set to your X server/VNC or
     rcons <one-node-at-a-time>    # or do rcons for each node
tail -f /var/log/messages
```

3.4 Test Service Node installation

- ssh to the service nodes. You should not be prompted for a password.
- Check to see that the xcat daemon xcatd is running.
- Run some database command on the service node, e.g tabdump site, or nodels, and see that the database can be accessed from the service node.
- Check that /install and /tftpboot are mounted on the service node from the Management Node, if appropriate.
- Make sure that the Service Node has Name resolution for all nodes, it will service.

4.0 Diagnostics

- **root ssh keys not setup** -- If you are prompted for a password when ssh to the service node, then check to see if /root/.ssh has authorized_keys. If the directory does not exist or no keys, on the MN, run xdsh service -K, to exchange the ssh keys for root. You will be prompted for

the root password, which should be the password you set for the key=system in the passwd table.

- **XCAT rpms not on SN** --On the SN, run rpm -qa | grep xCAT and make sure the appropriate xCAT rpms are installed on the servicenode. See the list of xCAT rpms in Set Up the Service Nodes for diskfull Installation. If rpms missing check your install setup as outlined in Build the Service Node Stateless Image for diskless or Set Up the Service Nodes for diskfull Installation for diskfull installs.
- **Error finding the database/starting xcatd** -- If on the Service node when you run tabdump site, you get "Connection failure: IO::Socket::SSL: connect: Connection refused at /opt/xcat/lib/perl/xCAT/Client.pm line 150." Then restart the xcatd daemon and see if it passes by running the command: service xcatd restart. If it fails with the same error, then check to see if /etc/xcat/cfgloc file exists. It should exist and be the same as /etc/xcat/cfgloc on the MN. If it is not there, copy it from the MN to the SN. The run service xcatd restart. This indicates the servicenode postscripts did not complete successfully. Check to see your postscripts table was setup correctly in Add Service Nodes postscripts to the postscripts table.
- **Error accessing database/starting xcatd credential failure**-- If you run tabdump site on the servicenode and you get "Connection failure: IO::Socket::SSL: SSL connect attempt failed because of handshake problemserror:14094418:SSL routines:SSL3_READ_BYTES:tlsv1 alert unknown ca at /opt/xcat/lib/perl/xCAT/Client.pm line 150.", check /etc/xcat/cert. The directory should contain the files ca.pem and server-cred.pem. These were suppose to transfer from the MN /etc/xcat/cert directory during the install. Also check the /etc/xcat/ca directory. This directory should contain most files from the /etc/xcat/ca directory on the MN. You can manually copy them from the MN to the SN, recursively. This indicates the the servicenode postscripts did not complete successfully. Check to see your postscripts table was setup correctly in Add Service Nodes postscripts to the postscripts table. Again service xcatd restart and try the tabdump site again.
- **Missing ssh hostkeys --** Check to see if /etc/xcat/hostkeys on the SN, has the same files as /etc/xcat/hostkeys on the MN. These are the ssh keys that will be installed on the compute nodes, so root can ssh between compute nodes without password prompting. If they are not there copy them from the MN to the SN. Again, these should have been setup by the servicenode postscripts.

5.0 Migrating a Management Node to a Service Node

If you find you want to convert an existing Management Node to a Service Node, you need to work with the xCAT team. It is recommended for now, to backup your database, setup you new Management Server, and restore your database into it. Take the old Management Node and remove xCAT and all xCAT directories, and your database. See Uninstalling xCAT2 documentation , and then follow the process for setting up a SN as if it is a new node.

6.0 Install or Stateless Boot the Service Nodes on AIX

For the installation and update of service nodes in an AIX cluster, refer to the documentation in xCAT2onAIXServiceNodes.

7.0 References

The xCAT advanced Cookbook gives you a detailed example of setting up hierarchy on a Linux Cluster.

The xCAT2onAIXServiceNodes gives you a detailed example of setting up hierarchy on an AIX cluster.

The xCAT2-updatenode doc gives you details on using updatenode to update your install and sync your configuration files on the node.