XCAT 2.x

01/23/2009

Table of Contents

1.0 Overview of xCAT

xCAT (Extreme Cluster Administration Tool) is a toolkit that provides support for the deployment and administration of large cluster environments.

Earlier versions of xCAT have been used to deploy and manage many high end Linux clusters since 1999. The new xCAT version 2.X is a complete rewrite of xCAT that includes many architectural changes and functional enhancements.

xCAT is a scalable distributed computing management and provisioning tool that provides a unified interface for hardware control, discovery, and OS diskful/diskless deployment. Now xCAT 2 is open source on the Source Forge Website , so you can use it with confidence and participate in making it even better.

1.1 xCAT Architecture

xCAT 2 is a complete rewrite of xCAT 1.3 that includes many architectural changes and functional enhancements. All commands are client/server, authenticated, logged and policy driven. XCAT 2 supports roll base authentication. The clients can be run on any OS with Perl, including Windows. All communication is SSL encrypted. The code has been completely rewritten in Perl, and table data is now stored in a relational database and with the plug-in architecture you can chose your database from SQLite, MySQL, PostgreSQL with more options coming.

In the xCAT client/server application, flow between the client and server is controlled by the xCAT daemon (xcatd) on the Management Node. When xcatd receives the command which has been packaged as (XML), it detemines whether the sender has authority to execute the command by evaluating the ACL's in the policy table. The daemon also receives status and inventory information from the nodes as they are deployed. See xCAT 2 Architecture for more details.

xCAT 2 was designed to scale for extremely large clusters. See xCAT on the world's fastest know supercomputer. With the Hierarchical support, a single management node may have any number of stateless or statefull service nodes to increase the provisioning throughput and management of the largest clusters. All cluster services such as LDAP, DNS, DHCP, NTP, Syslog, etc... can be automatically configured throughout the cluster. Outbound cluster management commands such as rpower, xdsh, xdcp, etc...utilize this hierarchy for scalable systems management.

1.1.1 Stateless and Stateful Choices

Stateless nodes are an important concept in xCAT 2. A stateless node is defined as one that has no "state" (configuration changes, software updates, etc.) stored permanently on it. This is extremely useful in a cluster for the following reasons:

- All nodes will have a much greater likelihood of staying consistent. And if the administrator does suspect that a node is out of sync with the rest of the cluster, they can simply reboot it and know that it is back in its original, pristine state.
- If a node experiences a hardware problem, the hardware can be pulled from the rack and replaced with new hardware and the node booted again and it will come up with the same state as before.
- In a provisioning environment, new nodes can be provisioned or moved without the worry of them losing state.

xCAT 2 provides the choice of either stateless or stateful nodes. A stateful node is one that has the OS installed on its local hard disk and therefore, changes to the node (configuration changes, software updates, etc.) can be made over time and those changes will persist.

Stateless nodes in xCAT 2 are implemented by not putting the OS on the local disk of the node. There are 3 choices for stateless:

1. **RAM-root** – The entire OS image is contained in a RAM file system that is sent to the node when it boots. Typical size for a minimal compute node for Linux is 75-160 MB of memory.
2. **Compressed RAM-root** – The OS image is in a compressed tar file. Individual files are extracted and cached when read. File writes are done to the cached copy. Typical size for a minimal compute node for Linux is 30-64 MB of memory.
3. **NFS Hybrid** – This is more accurately called NFS-root with copy-on-write. A minimal boot kernel is sent to the node, which readonly NFS mounts the OS image from the server. Files read are cached in memory. File writes are done to the cached copy. Typical size for a minimal compute node for Linux is 5 MB of memory.

1.2 xCAT Features

Features provided by xCAT for AIX or Linux clusters include the following:

- Deploying diskless and diskfull nodes.
- Node discovery
- Operating system image management.
- Support for user-provided customization scripts.
- xCAT data store in plug-in relational database (SQLite, MySQL,Postgresl, TBD)
- Hardware control commands for discovering hardware, gathering MAC addresses, VPD, and environments, power control, initiating a network boot, and LPAR creation/deletion.
- Hierarchical support to allow large system to distribute the management of the cluster to service nodes.
- Remote console support.
- Parallel remote shell and remote copy commands.
- Monitoring plug-in infrastructure (RMC, Ganglia)
- Notification infrastructure which lets users monitor xCAT database table changes.
- Predefined conditions, responses and sensors.
- Software and firmware inventory
- Xen support
- Windows support
- GUI for initial cluster setup
- Allow continuous operation during cluster software updates using plug-in job scheduler (LoadLeveler, Moab).
- Automatic setup for syslog, remote shell, DNS, DHCP, and ntp for both the xCAT management node and the cluster nodes.
- Documentation and "**man**" pages.

1.3 xCAT license

xCAT 2 Open Source License: Eclipse Public License

1.4 xCAT support

Support for xCAT may now be purchased. Other support is offered through the xCAT public mailing list. You can also open a new feature request or a bug report using Tracker on Source Forge. Check out the Support section on the xCAT Source Forge Home page.

1.5 xCAT Commands

Note: some of these commands run on Linux and AIX, some are targeted only for AIX or Linux.

1.5.1 Database support

- DB Tables– Complete list of xCAT database tables descriptions.
- chdef - Change xCAT data object definitions.
- chtab - Add, delete or update rows in the database tables.
- dumpxCATdb – dumps entire xCAT database.
- gettab – searches through tables with keys and return matching attributes.
- lsdef - used to display xCAT object definitions which are stored in the xCAT database.
- mkdef – used to create xCAT data object definitions.
- mkrrbc - Adds or deletes BladeCenter management module and switch node definitions in the xCAT cluster database.
- mkrrnodes – adds or deletes nodes in the xCAT cluster database. Allows creation/deletion of many nodes at once.
- nodeadd - Adds nodes to the xCAT cluster database.
- nodech - Changes nodes' attributes in the xCAT cluster database.
- nodels - lists the nodes, and their attributes, from the xCAT database.
- noderm - removes the nodes in the noderange from all database table.
- restorexCATdb – restore the xCAT database.
- rmdef - remove xCAT data object definitions.
- tabdump - isplay an xCAT database table in CSV format.
- tabedit - view an xCAT database table in an editor and make changes.
- tabgrep - list table names in which an entry for the given node appears.
- tabrestore - replaces the contents of an xCAT database table with the contents in a csv file.
- xcatstanzafile - Format of a stanza file that can be used with xCAT data object definition commands.

1.5.2 Hardware Control

- getmacs - Collects node MAC address.
- lsslp - Discovers selected networked services information within the same subnet.
- lsvm - Lists partition profile information for HMC- and IVM-managed nodes.
- nodestat - display the running status of a noderange
- rbeacon - Turns beacon on/off/blink or gives status of a node or noderange.
- rcons - remotely accesses the serial console of a node.
- replaycons - replay the console output for a node
- reventlog - retrieve or clear remote hardware event logs

- rmigrate - Execute migration of a guest VM between hosts/hypervisors .
- rmvm - Removes HMC- and IVM-managed partitions.
- rnetboot - Cause the range of nodes to boot to network.
- rpower - remote power control of nodes
- rscan - Collects node information from one or more hardware control points.
- rsetboot - Sets the boot device to be used for BMC-based servers for the next boot only.
- rspconfig - configures various settings in the nodes' service processors.
- rspreset - resets the service processors associated with the specified nodes
- switchblade - reassign the BladeCenter media tray and/or KVM to the specified blade
- wcons – windowed remote console
- wkill – kill windowed remote consoles

1.5.3 Monitoring

- monadd - Registers a monitoring plug-in to the xCAT cluster.
- moncfg - Configures a 3rd party monitoring software to monitor the xCAT cluster.
- mondecfg - Deconfigures a 3rd party monitoring software from monitoring the xCAT cluster.
- monls - Lists monitoring plug-in modules that can be used to monitor the xCAT cluster.
- monrm - Unregisters a monitoring plug-in module from the xCAT cluster.
- monstart - Starts a plug-in module to monitor the xCAT cluster.
- monstop - Stops a monitoring plug-in module to monitor the xCAT cluster.
- regnotif - Registers a Perl module or a command that will get called when changes occur in the desired xCAT database tables.
- unregnotif - unregister a Perl module or a command that was watching for the changes of the desired xCAT database tables.

1.5.4 Inventory

- rinv - remote hardware inventory.
- rvitals – retrieves remote hardware vitals information.
- sinv - Checks the software configuration of the nodes in the cluster.

1.5.5 Parallel Commands

- pcons - runs a command on the noderange using the out-of-band console.
- pping - parallel ping.
- ppping – parallel ping between nodes in a cluster.
- prsync – parallel rsync
- pscp – parallel remote copy (supports scp and not hierarchy)
- psh – parallel remote shell (supports ssh and not hierarchy)
- xdcp – concurrently copies files too and from multiple nodes. (scp/rcp and hierarchy)
- xdsh – concurrently runs commands on multiple nodes. (supports ssh/rsh and hierarchy)
- xdshbak- formats the output of the xdsh command.
- xcoll – Formats command output of the psh, xdsh, rinv command

1.5.6 Deployment

- copycds-cdrom - client side wrapper for copycds supporting physical drives.
- copycds - Copies Linux distributions and service levels from CDs to install directory.

- genimage - Generates a stateless image to be used for a diskless install.
- geninitrd - Regenerates the initrd for a stateless image to be used for a diskless install.
- mkdsklsnode - xCAT command to define and initialize AIX/NIM diskless machines.
- mknimimage - xCAT command to create AIX image definitions.
- mknb - creates a network boot root image for node discovery and flashing
- nimnodecust - xCAT command to customize AIX/NIM standalone machines.
- nimnodeset - xCAT command to initialize AIX/NIM standalone machines.
- nodeset - set the boot state for a noderange
- packimage - Packs the stateless image from the chroot file system.
- rbootseq - Persistently sets the order of boot devices for BladeCenter blades.
- rinstall - Begin installation on a noderange
- rmdsklsnode - Use this xCAT command to remove AIX/NIM diskless machine definitions.
- rmnimimage - xCAT command to remove an xCAT osimage definition and the associated NIM resources.
- setupiscsidev - creates a LUN for a node to boot up with, using iSCSI.
- updateSNimage – (No longer used) Adds the needed Service Node configuration files to the install image.
- updatenode - Reruns postsctipts or runs additional scripts on the nodes.
- winstall - Begin installation on a noderange and display in wcons
- xcat2nim - Use this command to create and manage AIX NIM definitions based on xCAT object definitions.

1.5.7 Others

- makedhcp - Creates new dhcp configuration files and updates live dhcp configuration using omapi.
- makedns - sets up domain name services (DNS) from the entries in /etc/hosts.
- makehosts - sets up /etc/hosts from the xCAT hosts table.
- makenetworks - populates the xCAT networks table, using network information from the local system
- noderange – Supported syntax for compactly expressing a list of node names.
- pbstop - Monitors your cluster in a terminal window.
- xcatstart - Starts the xCAT daemon (xcatd) on AIX.
- xcatstop - Stops the xCAT daemon (xcatd) on AIX.
- xCATWorld – Sample client program for xCAT.
- xpbsnodes - PBS pbsnodes front-end for a noderange.
- Summary of xCAT Commands

2.0 Installing a xCAT Management Node

To install the xCAT Management Node (MN), the following steps are taken:

1. Install the MN with the OS.
2. Configure Cluster-Facing NICS
3. Configure NTP
4. Configure Hostname
5. Configure DNS or some hostname resolution method.
6. Setup basic /etc/hosts file

7. Configure ethernet switches

2.1 Install a Linux Management Node with xCAT

2.1.1 Setup Your Networks

xCAT install process will scan and populate certain settings from the running configuration. Having the networks configured ahead of time will aid in correct configuration.

2.1.2 Install the Management Node OS

It is recommended to ensure that dhcp, bind (not bind-chroot), expect, httpd, nfs-utils, vsftpd, and perl-XML-Parser are installed. If the management server will be on the network and RHN activated or yum is pointed to the Fedora repositories, these installs will happen automatically later if not done now.

2.1.3 Ensure that SELinux is Disabled

/etc/sysconfig/selinux should contain:

```
SELINUX=disabled
```

If this change had to be made, reboot the system.

2.1.4 Prevent DHCP client from overwriting DNS configuration

Find the /etc/sysconfig/network-scripts/ifcfg-* files relevant to any NICs that are DHCP configured, and put “PEERDNS=no” into them.

2.1.5 Configure Cluster-Facing NICS

Configure the cluster-facing NICs. An example /etc/sysconfig/network-scripts/ifcfg-eth1:

```
DEVICE=eth1
ONBOOT=yes
BOOTPROTO=static
IPADDR=11.16.0.1
NETMASK=255.255.0.0
```

2.1.6 Configure Hostname

/etc/sysconfig/network should have HOSTNAME=(desired hostname).

2.1.7 Configure DNS Resolution

/etc/resolv.conf should at least point to its own DNS (which will get set up later). For example:

```
search cluster
```

```
nameserver 11.16.0.1
```

2.1.8 Set Up Basic hosts file

Ensure lines like the following is in /etc/hosts for each compute node, service node, bmc, mm,etc in your cluster on the Management Server:
This is important for using makedns.

```
127.0.0.1               localhost.localdomain localhost
::1                     localhost6.localdomain6 localhost6
11.16.9.227             xcat20mn.clusters.com xcat20mn
11.16.9.252             xcatmn
11.16.9.226             cu03sv
11.16.9.225             cu03svbmc
### Global CVLAN
192.168.0.1             mm
192.168.0.2             mm20
### Service Node CVLAN
192.152.101.254         cu03svgate
192.152.101.1           cu03sv-eth1
192.152.101.101         cu03cpbmc
192.152.101.102         cu03cp
11.16.9.146             c115v2mm01
11.16.9.212             c115bl4mm01
11.16.9.171             c115bv3mm01
```

2.1.9 Restart Management Node

Though it is possible to restart the correct services for all settings except SELinux, the simplest step would be to reboot the management server at this point.

2.1.10 Configure Ethernet Switches

xCAT can use the ethernet switches for discovery. In general, this requires that the user in advance set up an ip address and basic snmp functionality. Allowing the snmp version 1 community string “public” read access will allow xCAT to communicate without further customization. It is also recommended that spanning tree be set to portfast or edge-port for faster boot performance. Please see the relevant switch documentation as to how to configure these items.

2.1.11 Download Linux Distro ISOs and Create Repository for installs

1. For example, to install Fedora 8 on your cluseter. Get Fedora ISOs and place in a directory, for example /root/xcat2:

   ```
   mkdir /root/xcat2
   cd /root/xcat2
   export BASEURL=ftp://download.fedora.redhat.com/pub/fedora/linux/releases/8
   ```

```
wget $BASEURL/Fedora/x86_64/iso/Fedora-8-x86_64-DVD.iso
wget $BASEURL/Fedora/ppc/iso/Fedora-8-ppc-DVD.iso
```

2. Create YUM repository for Fedora RPMs (not needed on SLES):

```
mkdir /root/xcat2/fedora8
mount -r -o loop /root/xcat2/Fedora-8-x86_64-DVD.iso /root/xcat2/fedora8

cd /etc/yum.repos.d
mkdir ORIG
mv fedora*.repo ORIG

Create fedora.repo with contents:

[fedora]
name=Fedora $releasever - $basearch
baseurl=file:///root/xcat2/fedora8
enabled=1
gpgcheck=0
```

   On SLES, get access to the SLES RPMs and run "zypper sa <url>" to point to them.

3. Install createrepo (not needed on SLES):

```
yum install createrepo
```

2.1.12 Remove the package tftp-server

xCat ships the package atftp* package with its dependencies. Ths package conflicts with the tftp-server that is by default installed with the OS.
In order to install the xCAT successfully, you have to remove the tftp-server first.

2.1.13 Remove the package OpenIPMI-tool on RHEL

The xCAT software ships its own ipmi-tool* package with its dependencies. This package conflicts with the OpenIPMI-tool packages on RHEL, which must be removed before xCAT is installed.

2.1.14 Download and Install xCAT 2 From an MN That Has Internet Access

Follow the instructions at the xCAT Download site in the section titled "**RPMs in directories, structured for YUM download** ". This allows you to get a repo file to put on your management node that will point YUM to the proper place to get the xCAT RPMs. Then follow the instructions in section "**Open Source Package Dependencies that xCAT Requires"** to download the correct repo file for the xcat-dep packages**.**

For SLES, follow the instructions on the download page to point zypper to the online repositories.

2.1.15 Download and Install xCAT 2 For an MN That Does Not Have Internet Access

Follow the instructions at the xCAT Download site in the section titled **RPMs in tarball - download tarball ...** Download the open source dependencies tarball from the section **Open Source Package Dependencies that xCAT Requires** . Copy the files to the management node into /root/xcat2 and untar:

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

On Red Hat related distros point YUM to the local repositories for xCAT and its dependencies:

```
cd /root/xcat2/xcat-dep/rh5/x86_64
./mklocalrepo.sh
cd /root/xcat2/core-snap
./mklocalrepo.sh
```

Or on SLES, do:

```
zypper sa file:///root/xcat2/core-snap
zypper sa file:///root/xcat2/xcat-dep/sles10/x86_64
```

2.1.16 Get the Requisite Packages From the Distro

xCAT depends on several packages that come with the linux distro. Make those packages available to the install process. The method of doing that differs with each distro. A few examples are given here.

2.1.16.1 Fedora

If your management node has access to the internet, you can simply create a file called **/etc/yum.repos.d/fedora-internet.repo** that contains:

```
[fedora-everything]
name=Fedora $releasever - $basearch
failovermethod=priority
#baseurl=http://download.fedora.redhat.com/pub/fedora/linux/releases/
   $releasever/Everything/$basearch/os/
mirrorlist=http://mirrors.fedoraproject.org/mirrorlist?repo=fedora-
   $releasever&arch=$basearch
enabled=1
gpgcheck=1
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-fedora file:///etc/pki/rpm-gpg/RPM-GPG-
   KEY
```

If you management node does not have internet access, then download the necessary fedora RPMs, copy them to the MN, and create a local YUM repository:

```
cd /root/xcat2/xcat-dep/rh5/x86_64
```

```
export
   BASEURL=http://download.fedora.redhat.com/pub/fedora/linux/releases/8/Everything
   /x86_64/os/Packages/

wget $BASEURL/perl-Net-SNMP-5.2.0-1.fc8.1.noarch.rpm
wget $BASEURL/perl-XML-Simple-2.17-1.fc8.noarch.rpm
wget $BASEURL/perl-Crypt-DES-2.05-4.fc7.x86_64.rpm
wget $BASEURL/net-snmp-perl-5.4.1-4.fc8.x86_64.rpm
wget $BASEURL/ksh-20070628-1.1.fc8.x86_64.rpm
wget $BASEURL/perl-IO-Socket-INET6-2.51-2.fc8.1.noarch.rpm
wget $BASEURL/dhcp-3.0.6-10.fc8.x86_64.rpm
wget $BASEURL/syslinux-3.36-7.fc8.x86_64.rpm
wget $BASEURL/mtools-3.9.11-2.fc8.x86_64.rpm
wget $BASEURL/expect-5.43.0-9.fc8.x86_64.rpm
wget $BASEURL/perl-DBD-SQLite-1.12-2.fc8.1.x86_64.rpm
wget $BASEURL/perl-Expect-1.20-1.fc8.1.noarch.rpm
wget $BASEURL/perl-IO-Tty-1.07-2.fc8.1.x86_64.rpm
wget $BASEURL/scsi-target-utils-0.0-1.20070803snap.fc8.x86_64.rpm
wget $BASEURL/perl-Net-Telnet-3.03-5.1.noarch.rpm

createrepo .
```

Note: if using Fedora 9, you must have at least this version of net-snmp-perl: net-snmp-5.4.1-19.fc9

Continue now at step: 2.1.17Install xCAT and Dependencies on the Management Node

2.1.16.2 RHEL

To make the necessary RHEL RPM prereqs available to the xCAT install process, mount the RHEL DVD or ISO and then create a repo file in /etc/yum.repos.d that points to it. For example, create a file call rhel-dvd.repo with the following contents
:

```
name=RHEL 5.3 from DVD
baseurl=file:///media/cdrom/Server
enabled=1
gpgcheck=0
```

2.1.16.3 SLES

Mount the SLES DVD or ISO and then run:

```
zypper sa file:///media/cdrom/?????
```

2.1.17 Install xCAT and Dependencies on the Management Node

```
yum clean metadata
yum install xCAT.x86_64
```

Or on SLES, do:

```
zypper install xCAT
```

2.1.18 Test xCAT Installation

```
source /etc/profile.d/xcat.sh
tabdump site
```

2.1.19 Update xCAT Software after Installation

If you need to update the xCAT 2 rpms later:

- If the management node does not have access to the internet: download the new version of http://xcat.sf.net/yum/core-rpms-snap.tar.bz2 and untar it in the same place as before
- (If the management node has access to the internet, the yum command below will pull the updates directly from the xCAT site.)

Then run:

```
yum update '*xCAT*'
```

Or on SLES, do:

```
zypper update -t package '*xCAT*'
```

Note: If you have a service node stateless image in a hierachical configuration , don't forget to update the image with the new xCAT rpms to keep the service node at the same level as the management Node.

2.2 Install an AIX Management Node with xCAT

2.2.1 Setup xCAT Management Node

- Follow AIX documentation and procedures to install and configure the base AIX operating system. (Typically by using the product media.)
- Apply the latest software updates and fixes if needed.
- Install the latest versions of OpenSSL & OpenSSH from the AIX Expansion Pack. This software can also be downloaded from the following sites.

  OpenSSH:

  http://sourceforge.net/projects/openssh-aix

  OpenSSL:

  https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=aixbp

Since these are installp file sets you should run /usr/sbin/updtvpkg to make sure that the RPM reflection of what was installed by installp is updated. This makes it possible for RPM packages with a dependency on OpenSSL to recognize that the dependency is satisfied.

NOTE: For easier downloading without a web browser, you may want to download and install the wget tool from the AIX Toolkit for Linux.

2.2.2 Download and install the prerequisite Open Source Software (OSS)

- Download the dep-aix-2.1.tar.gz tar file from http://xcat.sourceforge.net/aix/download.html and copy it to a convenient location on your xCAT management node.
- Unwrap the tar file.
- gunzip dep-aix-2.1.tar.gz
- tar –xvf dep-aix-2.1.tar
- Read the README file.
- Run the instoss script (contained in the tar file) to install the OSS packages.

2.2.3Verify that the following basic AIX services needed

- **inted**

  inetd includes services such as telnet, ftp, bootp, and others. Edit the /etc/inetd.conf file to turn on all services that are needed. Ftp and bootp are required for pSeries node installations. Stop and restart the inetd service after any changes:

  stopsrc -s inetd

  startsrc -s inetd

- **NFS**

  NFS is required for all NIM installs. Ensure the NFS daemons are running:

  lssrc -g nfs

  If any NFS services are inoperative, you can stop and restart the entire group of services:

  stopsrc -g nfs

  startsrc -g nfs

  There are other system services that NFS depends on such as inetd, portmap, biod, and others. A handy trouble-shooting guide listing the required services can be found at:

http://publib.boulder.ibm.com/infocenter/systems/index.jsp?topic=/com.ibm.aix.commadmn/doc/commadmndita/nfs_troublesh.htm

- **DNS (name resolution)**

  Name resolution is required by xCAT. You can use a simple /etc/hosts mechanism, but setting up DNS would also work. You will need to have an accurate /etc/hosts file to use the xCAT makedns command. See Set Up Basic hosts file.

  Make sure to verify and update the contents of both /etc/hosts (you may wish to use the xCAT makehosts command to do this after your nodes have been defined) and /etc/resolv.conf (for domain name and external name servers). See Configure DNS Resolution.

2.2.4 Download and Install the xCAT software

- Download the latest xCAT for AIX tar file from http://xcat.sourceforge.net/aix/download.html and copy it to a convenient location on your xCAT management node.
- Unwrap the xCAT tar file. For example,
  - *gunzip core-aix-2.1.tar.gz*
  - *tar –xvf core-aix-2.1.tar*
- Run the **instxcat** script (contained in the tar file) to install the xCAT software. The post processing provided by the xCAT packages will perform some basic xCAT configuration. (This includes initializing the SQLite database and starting **xcatd** daemon processes.)
- Execute the system profile file to set the xCAT paths. This file was updated during the xCAT post install processing. (". */etc/profile").* (**Note**: Make sure you don't have a .profile file that overwrites the "PATH" environment variables.)

2.2.5Verify the xCAT install

- Run the "*lsdef –h*" to check if the xCAT daemon is working. If you get a correct response then you should be ok.
- Check to see if the initial xCAT definitions have been created. For example, you can run "*lsdef -t site -l*" to get a listing of the default site definition. You should see output similar to the following.

Setting the name of the site definition to 'clustersite'.

Object name: clustersite

domain=abc.foo.com

installdir=/install

tftpdir=/tftpboot

master=7.104.46.27

useSSHonAIX=no

xcatdport=3001

xcatiport=3002

2.3 Setup Alternate Databases

XCAT supports a plugable interface which allow you to choose the relational database you wish to use. The following are the currently supported databases, with SQLite being the default when xCAT is installed on the Management Node for the first time.

2.3.1 SQLite

XCAT will automatically perform the initial setup of an SQLite Database when the Management Node is first installed. This database is sufficient for small to moderate size systems (less than 1000 nodes for Linux, 300 for AIX) , if you are not using hierarchy (service nodes). SQLite cannot be used for hierarchy, because the service nodes require access to the database from the service node and this SQLite does not support remote access to the database. For hierarchy, you need to setup PostgreSQL or MySQL, see below.

2.3.2 PostgreSQL

Instructions for setting up a PostgreSQL database on Linux, go to the xCAT2 Cookbook for Linux.

2.3.3 MySQL

Instructions for setting up a MySQL data base for xCAT on AIX or Linux are found in the xCAT2.1 MySQL setup documentation.

2.3.4 DB2 (TBD)

3.0 Deploying and Maintaining Cluster Nodes

3.1 Linux Nodes

xCAT supports deployment of diskfull and diskless Linux nodes in the cluster.

3.1.1 Linux Cookbook

The xCAT2 Linux Cookbook provides information on setting up diskfull and diskless Linux clusters.

3.1.2 BladeCenter How-To

Instructions for installation of a BladeCenter configuration are contained in the BladeCenter How-to.

3.1.3 IdataPlex How-To

An example of an iDataPlex configuration, and instructions for installation are contained in the iDataPlex How-to.

3.1.4 SLES 10.1 notes

Some helpful notes on installing SLES 10 SP1 are contained in the "SLES 10 SP 1 notes".

3.2 AIX Nodes

xCAT supports deployment of diskfull and diskless AIX nodes and mksysb installs in the cluster.

3.2.1 AIX diskfull nodes

xCAT will install AIX standalone nodes using the NIM (Network Installion Management) "rte" method. NIM is an AIX tool that enables a cluster administrator to centrally manage the installation and configuration of AIX and optional software on machines within a networked environment. XCAT has features that will add you to automatically run the necessary NIM commands. The "Installing AIX on Standalone Nodes " how-to will guide you in this process.

3.2.2 AIX diskless nodes

xCAT supports deploying AIX diskless nodes using NIM. The "Booting AIX Diskless Nodes " how-to describes the process for deploying AIX diskless nodes.

3.3 Updating the xCAT Cluster

You may need to update the nodes in the cluster (OS or applications) updating after the initial install. There are several ways to update depending on the scope of the update. You can use the updatenode function in xCAT to install additional software on the nodes. Other options are to use nimnodecust, rolling update, xdsh , update the diskless image and reboot etc.

4.0 Node Discovery

One of the significant features of xCAT 2 is the node discovery approach. It ultimately performs the role of associating node MAC addresses with IP based on some physical cue (ethernet port or Bladecenter slot). It has the same goal as getmacs fulfilled historically, except it is node initiated, has more context to enable accommodation of more complex configurations, and automated.

The xCAT wiki section on Node Discovery contains information on the setup required to use the Discovery function.

5.0 Using Hierarchy

In large clusters it is desirable to have more than one node (the Management Node) handle the installation and management of the compute nodes. We call these additional nodes service nodes.

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, if you are using Linux, having a pool of service nodes, any of which can respond to an installation request from a compute node. This document will cover the former case (distinct sets).

The service nodes need to communicate with the xCAT 2 database on the Management Node and run xCAT commands to install the nodes. The service node will be installed with the xCAT code and requires that the either MySQL or PostgreSQL Database be set up instead of the SQLite Default database. These databases allows a remote client to be set up on the service node such that the service node can access (read/write) the database on the Management Node.

5.1 Linux

For a Linux cluster, setting up hierarchical support is documented in xCAT2 Cookbook for Linux.

5.2 AIX

For an AIX cluster, setting up hierarchical support is documented in (TBD).

6.0 Monitoring

There are two monitoring infrastructures in xCAT 2.0. The ***xCAT Monitoring Plug-in Infrastructure*** allows you to plug-in one or more third party monitoring software such as Ganglia, RMC, SNMP etc. to monitor the xCAT cluster. The ***xCAT Notification Infrastructure*** allows you to watch for the changes in xCAT database tables.

How to enable and use the xCAT Monitoring infrastructure is documented in the . "xCAT 2.0 Monitoring How-to"

7.0 Uninstalling xCAT on Linux

Steps for removing xCAT from your Management Node are documented in Uninstalling xCAT2

8.0 Migrating from xCAT 1.3 to 2.x

If you are planning to migrate from xCAT 1.3 to 2.x, read the xCAT Quick Install Tutorial.

9.0 xCAT for CSM Admin

The Redbook "xCAT 2 Guide for the CSM System Administrator" describes the xCAT architecture, Quick deployment, and contains CSM to xCAT transition scenarios.

10.0 xCAT on Windows

Directions for installing a Windows 2008 Enterprise Server (x86 / x86_64) node with xCAT 2.1 are being developed. The latest notes are available on the SourceForge xCAT wiki.

11.0 xCAT and Xen

xCAT can be configured to work with the Xen hypervisor to install and manage virtual compute nodes. Installing and setting up Xen for xCAT is documented in the xCAT & Xen How-to.

12.0 References

12.1 xCAT Summary Commands and Database Tables

Select here for Summary of Commands.

Select here for Summary of Database Tables.

12.2 Stateless GPFS

GPFS is a premier cluster filesystem. You can run it stateless on nodes. "How to setup a Stateless GPFS Cluster" documents the process.

12.3 Maui

Installing and setting up MAUI for xCAT is documented in the xCAT How-to for MAUI.

12.4 Torque

Installing and setting up Torque for xCAT is documented in the xCAT2 Cookbook for Linux.

12.5 Ganglia

Installing and setting up Ganglia for xCAT is documented in the xCAT How-to for Ganglia.

12.6 LDAP

LDAP can be used for user Management in xCAT. Installing and setting up LDAP for xCAT is documented in the xCAT How-to for LDAP.

12.7 XCAT Developer Guide (TBD)

This space will point to helpful hints and techniques for developers who would like to write code for xCAT.

13.0 Known Bugs

https://sourceforge.net/tracker/?group_id=208749&atid=1006945

14.0 Feature requests

https://sourceforge.net/tracker/?group_id=208749&atid=1006948

15.0 References

- xCAT web site: http://xcat.sf.net/
- xCAT wiki: http://xcat.wiki.sourceforge.net/
- xCAT mailing list: http://xcat.org/mailman/listinfo/xcat-user