

XCAT 2.x

12/17/2008

Table of Contents

1.0 Overview of xCAT	2
1.1 xCAT Architecture	2
1.2 xCAT Features	3
1.3 xCAT license	3
1.4 xCAT support	3
1.5 xCAT Commands	3
1.5.1 Database support	3
1.5.2 Hardware Control	4
1.5.3 Monitoring	4
1.5.4 Inventory	5
1.5.5 Parallel Commands	5
1.5.6 Deployment	5
1.5.7 Others	6
2.0 Installing a xCAT Management Node	6
2.1 Basic Steps	6
2.2 Databases	7
2.2.1 SQLite	7
2.2.2 PostgreSQL	7
2.2.3 MySQL	7
2.2.4 DB2 (TBD)	7
3.0 Deploying and Maintaining Cluster Nodes	7
3.1 Linux	7
3.1.1 Linux Cookbook	7
3.1.2 BladeCenter (TBD)	7
3.1.3 iDataPlex	7
3.1.4 SLES 10.1	7
3.2 AIX	7
3.2.1 AIX diskfull nodes	7
3.2.2 AIX diskless nodes	8
3.3 Updating the Cluster	8
4.0 Node Discovery	8
5.0 Using Hierarchy	8
5.1 Linux	8
5.2 AIX	8
6.0 Monitoring	8
7.0 Uninstalling xCAT	9
8.0 Migrating from xCAT 1.3 to 2.x	9
9.0 xCAT for CSM Admin	9
10.0 xCAT on Windows	9
11.0 xCAT and Xen	9

12.0 References	9
12.1 xCAT Summary commands and tables	9
12.2 Stateless GPFS	9
12.3 Maui	9
12.4 Torque	9
12.5 Ganglia	9
12.6 LDAP	9
12.7 XCAT Developer Guide (TBD)	10
13.0 Known Bugs	10
14.0 Feature requests	10
15.0 References	10
16.0 Glossary	10

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 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.

1.2 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 OS updates (rollupdates) 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

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
- [psh](#) – parallel remote shell
- [xdcp](#) – concurrently copies files too and from multiple nodes.
- [xdsh](#) – concurrently runs commands on multiple nodes.
- [xdshbak](#)- formats the output of the xdsh 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.
- [mkdsklnode](#) - 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
- [rmdsklnode](#) - 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 postscripts 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.
- [xcoll](#) – Formats command output
- [xpbsnodes](#) - PBS pbsnodes front-end for a noderange.
- [Summary of xCAT Commands](#)

2.0 Installing a xCAT Management Node

2.1 Basic Steps

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

1. Setup your cluster networks. On Linux, 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. Install the MN with the OS.
On Linux, 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. Ensure SELinux is disabled.
On AIX, it is recommended to setup OpenSSH for remote shell connection. All relevant base AIX services should be configured and running. This includes (but is not limited to) bootp, tftp, NFS, and hostname resolution.
3. Configure Cluster-Facing NICS
4. Configure NTP
5. Configure Hostname
6. Configure DNS or some hostname resolution method.
7. Setup basic /etc/hosts file
8. Configure ethernet switches
9. Download and install xCAT 2 on the Management Node.
For Linux, follow instructions in [xCAT2 Cookbook for Linux](#).
For AIX, follow instructions in [xCAT2.1 on AIX](#).
10. Verify xCAT running correctly , by running commands `lsdef -h` , or `tabdump site` .

2.2 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.

2.2.1 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) , if you are not Using Hierarchy (service nodes).

2.2.2 PostgreSQL

Instructions for setting up a PostgreSQL database on Linux, go to the [xCAT2 Cookbook for Linux](#).

2.2.3 MySQL

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

2.2.4 DB2 (TBD)

3.0 Deploying and Maintaining Cluster Nodes

3.1 Linux

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 (TBD)

Instructions for installation of a BladeCenter configuration.

3.1.3 iDataPlex

An example of an iDataPlex configuration, and instructions for installation are contained in the [iDataPlex How-to](#).

3.1.4 SLES 10.1

Some helpful notes on installing SLES 10 SP1 are contained in the “[SLES 10 SP 1 notes](#)”.

3.2 AIX

xCAT supports deployment of diskfull and diskless AIX nodes 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 Cluster

For updating your cluster after the initial install, you can use the [updatenode](#) function in xCAT.

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 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

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 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>

16.0 Glossary