xCAT Cluster Management

11/18/2010, AM 08:23:04

Table of Contents

1.1 Overview of xCAT

xCAT (Extreme Cloud Administration Tool) is a toolkit that provides support for the deployment and administration of large cluster environments. This documents will describes the tools that xCAT provides to manage your cluster.

xCAT 2 is open source on the Source Forge Website , so you can use it with confidence and participate in making it even better.

1.2 xCAT Management Features

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

- Documentation (online and printable) and “man” pages.
- xCAT data stored in plug-in relational database or your choice (SQLite, MySQL,Postgresl, DB2)
    - xCAT commands to access the current database.
    - Node group support.

- Parallel remote commands (remote shell, remote copy and rsync support)
- Software and firmware inventory
- Logging and Auditing
- Monitoring plug-in infrastructure (RMC, Ganglia). Shipped monitoring service xcatmon.
    - Notification infrastructure which lets users monitor xCAT database table changes.
    - Predefined conditions, responses and sensors.
    - Space monitoring (e.g /var size)
- Automatic setup of services such as syslog, ssh, DNS, DHCP, and ntp for both the xCAT management node, service nodes, and the cluster nodes.
- updatenode capability – post-install maintenance of the nodes.
- Support for user-provided customization scripts (to run during pre and post-install).
- User Management
- HA support for the xCAT Management Nodes
- HA support for the xCAT Service Nodes
- Health Checks
- Hardware control commands for discovering hardware (node discovery) , gathering MAC addresses, VPD, and environments, power control, initiating a network boot, and LPAR creation/deletion.
- Deploying diskless, diskfull and statelite nodes.
- Hierarchical support to allow large system to distribute the management of the cluster to service nodes.
- Remote console support.
- Allow continuous operation during cluster software updates (**rolling updates**) using plug-in job scheduler (LoadLeveler, Moab).
- Install and configuration assistance for IBM HPC products (GPFS, LoadLeveler, Parallel Environment, ESSL/PESSL, compilers, RSCT).

1.3 xCAT Documentation

xCAT provides a complete set of documentation covering the Cluster environment you are trying to setup or maintain. The documentation is available on the SourceForge WIKI, and in PDF and HTML format. The documentation online has a "discussion" tab at the top. Use this to leave your comments about how that page could be improved, We welcome the user community to help make our documentation better. Users can also be given permission to edit current docs or supply new ones.

For current documents:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_Documentation

1.4 xCAT Database

xCAT supports a plugable interface which allow you to choose the relational database you wish to use. SQLite being the default when xCAT is installed on the xCAT Management Node for the first time. xCAT current also supports (MySQL, Postgresql,DB2) on AIX and Linux.

xCAT will automatically perform the initial setup of an SQLite Database when the xCAT 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 SQLite does not support remote access to the database. For hierarchy, you need to setup PostgreSQL or MySQL, or DB2.

xCAT has tools to automatically setup all the databases for you. xCAT also provides a full set of database command so that you can access the database without caring which database you are using with xCAT.

For information on the supported databases and the xCAT setup scripts:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Choosing_the_Database

For information on the setup of xCAT commands to access the database:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Listing_%26_Modifying_the_Database

1.5 Parallel Commands

xCAT delivers a set of commands that can be run remote commands (ssh, scp,rsync,ping,cons) in parallel on multiple nodes. In addition the command have the capability of formatting the output from the commands, so the results are easier to process. These command will make it much easier to administer your large cluster.

For a list of the Parallel Commands and their man pages:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_Commands#Parallel_Commands

1.6 Software and Firmware Inventory

xCAT provides a command '**sinv**' that checks the software and firmware configuration in this cluster. This command uses the xdsh parallel command, so it is in itself a parallel command, and thus can be run on multiple cluster nodes at one time and is hierarchical.

See the sinv man page for more information.

1.7 Logging and Auditing

Syslog

xCAT automatically sets up syslog during the install on the xCAT Management Node (MN) and the cluster nodes. It use local4 to register it syslog entries and it logs to the /var/log/messages file.

The setup of the service nodes and compute nodes is to send all syslog messages to the Management Node (MN) to the /var/log/messages .

Auditing

xCAT will not only log all xCAT commands that are run and who runs them to syslog, but also to an auditlog table in the xCAT database.

For more information on syslog and auditlog:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Syslog_and_auditlog

1.8 Monitoring

There are two monitoring infrastructures in xCAT 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:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Monitoring_an_xCAT_Cluster

1.9 Automatic setup of Services

xCAT when it is installed on the xCAT Management Node, and when it installs the service nodes or cluster node automatically sets up the required services such as ssh, dns, dhcp, nfs,etc. It determines this setup based on attributes put in the database for the node. If the node is an nfs server, then nfs will be setup.

By default, during install XCAT sets up ssh root passwordless access between all nodes. We give you the option to limit the nodes that will be setup for node to node passwordless ssh access. ssh is still always setup with root passwordless access from the Management nodes and Service Nodes to the nodes and between Service Nodes.

1.10 Updatenode

The updatenode command has many features to allow you to manage and update your nodes after install. Some of these are (adding and updating software packages, sync'ing configuration files, re-running post and preinstall scripts, correcting broken ssh keys and credentials)

For more information check the updatenode doc:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Using_Updatenode

and the man page:

http://xcat.sourceforge.net/man1/updatenode.1.html

1.11 User-provided customization scripts

xCAT automatically runs a few postscripts and postbootscripts that are delivered with xCAT to set up the nodes. You can also add your own scripts to further customize the nodes.

The 'Using and Creating Postscripts' document describes how to setup your own scripts:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Using_%26_Creating_Postscripts

1.12 User management

xCAT only sets up root during install to access ssh without password to the cluster, access the database and run xCAT commands. You can add other userids to have there priviledges by following the instructions outlined in the User Management Document:
https://sourceforge.net/apps/mediawiki/xcat/index.php?title=User_Management

1.13 HA for the xCAT Management Node

The following documentation explains the procedure for setting up a HA Backup Management Node:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Setting_Up_an_HA_Mgmt_Node

1.14 HA for the xCAT Service Nodes

TBD

1.15 Health Checks

TBD

1.16 Hardware Control Commands

xCAT supports for hardware discovery by the xCAT Management Node (MN), and setting up connections between HMC to IBM System P machines.

For detailed instructions on using these commands, refer to the document on 'xCAT System P Hardware Management' :
https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_System_p_Hardware_Management

1.17 Deploying Stateless, Statefull and Statelite Nodes

xCAT supports three types of node install:

Stateless nodes are defined as one that has no "state" (configuration changes, software updates, etc.) stored permanently on it.

Statefull nodes have 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.

Statelite nodes provides an efficient and flexible diskless solution , because most of the OS image is NFS mounted read-only, but a configurable list of directories and files can be read-write. The read-write files can either be persistent across reboots, or volatile (restoring to pristine state after reboot).

For Linux Deployment

If installing Statefull or Stateless, refer to the following document:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_pLinux_Clusters

If installing Statelite:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_Linux_Statelite

For AIX Deployment

If installing Statefull, refer to the following documents:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_AIX_RTE_Diskfull_Nodes
https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_AIX_mksysb_Diskfull_Nodes

If installing Stateless:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_AIX_Diskless_Nodes

If installing Statelite:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=XCAT_AIX_Cluster_Overview_and_Mgmt_Node

1.18 Hierarchical Support

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.

The service nodes need to communicate with the xCAT database on the xCAT 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, PostgreSQL or DB2 Database be set up instead of the SQLite default database on the xCAT Management Node. 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.

For more information on setting up a Hierarchical Cluster

For Linux:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Setting_Up_a_Linux_Hierarchical_Cluster

For AIX:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Setting_Up_an_AIX_Hierarchical_Cluster

1.19 Remote Console Support

xCAT provides two commands rcons and wcons for remote access to the serial console of a node. This is invaluable for install monitoring
See the rcons and wcons man page for more information.

1.20 Rolling Updates

The xCAT rolling update support allows you to update the OS image on a subset of cluster nodes at a time, such that the remaining nodes in the cluster can still be running jobs.

The following document gives details on the setup required to use this function:

https://sourceforge.net/apps/mediawiki/xcat/index.php?title=Rolling_Update_Support

1.21 Install and configuration assistance for IBM HPC products (GPFS, LoadLeveler, Parallel Environment, ESSL/PESSL, compilers, RSCT).

TBD