

xCAT Stalite Cookbook

05/19/10, 02:39:23 PM

Table of Contents

1	Introduction.....	2
1.1	Limitations	3
2	Configuring Stalite	3
2.1	Lock Down main image.....	3
2.2	litefile table.....	4
2.3	litetree table.....	5
2.4	stalite table.....	6
2.5	Policy.....	6
2.6	Add Linux Distro Packages.....	7
2.7	Add Third-Party Software.....	7
2.8	Create Stalite Image.....	8
2.8.1	Prepare the “.postinstall” file	8
	Rules for evaluating postinstall file	8
2.8.2	Run “genimage” command	9
2.9	Setting up Post install scripts for stalite.....	9
2.10	Modify stalite image.....	10
2.11	Installing a new Kernel in the stalite image (This is not verified).....	10
2.12	Run liteimg command.....	10
2.13	Set the boot state to “stalite”.....	11
2.14	Install the Node.....	11
3	New Commands.....	12
4	Stalite Directory Structure.....	13
4.1	The noderes Table.....	13
5	Adding/updating software and files for the running nodes.....	13
6	Hierarchy Support	14
6.1	Setup the diskfull service node.....	14
6.2	Generate the stalite image.....	14
6.3	Sync the “/install” directory.....	14
6.4	Set the boot state to “stalite”.....	14
6.5	Install the Node.....	14
7	Advanced Stalite features.....	15
7.1	litetree table.....	15
8	Debugging techniques.....	16

1 Introduction

This document details a design for native xCAT NFS Root. We call the design statelite, because in addition to having an NFS Root implementation, some files can be stored persistently and maintain state through reboots.

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

Statelite offers the following advantages over xCAT's stateless (RAMdisk) implementation:

1. Some files can be made persistent over reboot. This is useful for license files or database servers where some state is needed. However, you still get the advantage of only having to manage a single image.
2. Changes to hundreds of machines can take place instantly, and automatically, by updating one main image. In most cases, machines do not need to reboot for these changes to take affect.
3. Ease of administration by being able to lock down an image. Many parts of the image can be read-only, so no modifications can transpire without updating the main image.
4. Files can be managed in a hierarchical manner. For example: Suppose you have a machine that is in one lab in Tokyo and another in London. You could set table values for those machines in the xCAT database to allow machines to sync from different places based on their attributes. This allows you to have one base image with multiple sources of file overlay.
5. Ideal for virtualization – In a virtual environment, you may not want a disk image (neither stateless nor stateful) on every virtual node as it consumes memory and disk. Virtualizing with the statelite approach allows for images to be smaller, easier to manage, use less disk, less memory, and more flexible.

Disadvantages

1. NFS Root requires more network traffic to run as the majority of the disk image runs over NFS. This may depend on your workload, but can be minimized.
2. NFS Root can be complex to set up. As more files are created in different places ,there are greater chances for failures. We think this flexibility is also one of the great virtues of Statelite. The image can work in nearly any

environment.

1.1 Limitations

Currently, statelite has only been tested on Red Hat and Red Hat Clone environments on x86_64 platforms, and SLES11 (SuSE Linux Enterprise Server 11) on POWER systems. We have no reason to think it won't work Red Hat on POWER and SLES11 on x86_64.

Statelite does not yet support hierarchical clusters (those with service nodes).

Statelite is not currently supported on AIX.

Only one statelite image may be defined for a node or noderange.

2 Configuring Statelite

Getting started with Statelite provisioning requires that you have xCAT set up and running. Before continuing with the rest of this document, the xCAT management node must be set up, and the nodes' hardware control, resource, and type attributes must be defined as described in the [xCAT Top Doc](#).

Please report any errors to the xCAT mailing list (<https://lists.sourceforge.net/lists/listinfo/xcat-user>).

2.1 Lock Down main image

Any statelite image will need some files to be read/writable for individual nodes. But we encourage you to first lock down the main image. The image will be placed in `/install/netboot/<os>/<arch>/<profile>/rootimg` after running `genimage` (later in this doc).

xCAT, by default, during installation will create on the Management Node an `/etc/exports` entry that exports the `/install` directory read/writeable. Change it to read-only, so it looks as follows:

```
/install *(ro,no_root_squash,sync,no_subtree_check)
```

Restart the NFS server:

```
service nfs restart
```

2.2 litefile table

The litefile table specifies the directories and files on the statelite nodes that should be readwrite, persistent, or readonly overlay. All other files in the statelite nodes come from the readonly statelite image. See [litefile man page](#) for details on the file attributes.

Sample Data for a RedHat statelite setup.

This is the minimal list of files needed, you can add your additional files to the setup.

Notice that all files are in tmpfs, the default for the options field. This gives you an NFS root solution with no persistent storage.

```
#image,file,options,comments,disable
"ALL", "/etc/adjtime",,,,
"ALL", "/etc/fstab",,,,
"ALL", "/etc/inittab",,,,
"ALL", "/etc/lvm/.cache",,,,
"ALL", "/etc/mtab",,,,
"ALL", "/etc/ntp.conf",,,,
"ALL", "/etc/ntp.conf.predhclient",,,,
"ALL", "/etc/resolv.conf",,,,
"ALL", "/etc/resolv.conf.predhclient",,,,
"ALL", "/etc/ssh/",,,,
"ALL", "/tmp/",,,,
"ALL", "/var/account/",,,,
"ALL", "/var/arpwatch",,,,
"ALL", "/var/cache/alchemy",,,,
"ALL", "/var/cache/foomatic/",,,,
"ALL", "/var/cache/logwatch/",,,,
"ALL", "/var/cache/man/",,,,
"ALL", "/var/cache/mod_ssl/",,,,
"ALL", "/var/cache/mod_proxy/",,,,
"ALL", "/var/cache/php-pear/",,,,
"ALL", "/var/cache/systemtap/",,,,
"ALL", "/var/empty/",,,,
"ALL", "/var/db/nscd/",,,,
"ALL", "/var/gdm/",,,,
"ALL", "/var/lib/dav/",,,,
"ALL", "/var/lib/dhcp/",,,,
"ALL", "/var/lib/dhclient/",,,,
"ALL", "/var/lib/php/",,,,
"ALL", "/var/lib/scsi/",,,,
"ALL", "/var/lib/ups/",,,,
"ALL", "/var/lib/random-seed",,,,
"ALL", "/var/lib/iscsi",,,,
"ALL", "/var/lib/logrotate.status",,,,
"ALL", "/var/lib/ntp/",,,,
"ALL", "/var/lib/xen/ntp",,,,
"ALL", "/var/lock/",,,,
"ALL", "/var/log/",,,,
"ALL", "/var/run/",,,,
```

```
"ALL", "/var/tmp/", , , ,
"ALL", "/var/tux/", , , ,
"ALL", "/opt/xcat/", , , ,
"ALL", "/xcatpost/", , , ,
```

Sample Data for SLES11

This is the minimal list of files needed, you can add your additional files to the setup.

```
#image,file,options,comments,disable
"ALL", "/etc/inittab", , , ,
"ALL", "/etc/lvm/.cache", , , ,
"ALL", "/etc/mtab", , , ,
"ALL", "/etc/ntp.conf", "bind", , ,
"ALL", "/etc/ntp.conf.org", "bind", , ,
"ALL", "/etc/ntp.conf.postbackup", "bind", , ,
"ALL", "/etc/resolv.conf", , , ,
"ALL", "/etc/ssh/", , , ,
"ALL", "/etc/sysconfig/", , , ,
"ALL", "/etc/syslog-ng/", , , ,
"ALL", "/tmp/", , , ,
"ALL", "/var/tmp/", , , ,
"ALL", "/var/run/", , , ,
"ALL", "/etc/yp.conf", , , ,
"ALL", "/var/lib/", , , ,
"ALL", "/var/empty/", , , ,
"ALL", "/var/spool/", , , ,
"ALL", "/var/lock/", , , ,
"ALL", "/var/log/", , , ,
"ALL", "/var/cache/", , , ,
"ALL", "/etc/fstab", , , ,
"ALL", "/var/adm/", , , ,
"ALL", "/root/.viminfo", , , ,
"ALL", "/root/.bash_history", , , ,
"ALL", "/opt/xcat/", , , ,
"ALL", "/xcatpost/", , , ,
```

Use the command

```
tabedit litefile
```

to copy/paste the above sample lines into the litefile table.

Note: It is recommended that DO NOT change their default options, since some file are necessary for the node to boot up, for example, "/etc/ssh/", "/etc/mtab". Before you change the default option, you must make sure it won't affect the system.

2.3 litetree table

The litetree table controls where the initial content of the files in the litefile table come from, and the long term content of the "ro" files. When a node boots up in statelite mode, it will by default copy all of its tmpfs files from the /.default directory

of the root image, so there is not requirement to setup a litetree table. If you decide that you want some of the files pulled from different locations that are different per node, you can use this table. See Advanced Statelite features.

You can chose to use the defaults and not setup a litetree table.

2.4 statelite table

You may want some files in the image to be stored permanently, to survive reboots. This is done by entering the information into the `statelite` table.

See [statelite man page](#) for description of the attributes.

Note: In the `statelite` table, the `node` or `nodegroups` in the table must be unique; that is a node should appear only once in the table, as itself or as the member of a group. This makes sure that only one `statelite` image can be assigned to a node. See Limitations.

An example would be:

```
"japan",,,"cnfs:/gpfs/state",,,
```

Any nodes in the `japan` node group will have their state stored in the `/gpfs/state` directory on the machine known as `cnfs`. The `image` field should not be updated, currently it is not used.

When the node boots up, then the value of `statemnt` will be mounted to `/.statelite/persistent`. The code will then create the following subdirectory `/.statelite/persistent/<nodename>`

This directory will be the root of the image for this node's persistent files.

To set the “`statemnt`” value, you can use variables in xCAT database, it follows the grammar in the `litetree` table, for example:

```
#node,image,statemnt,comments,disable  
"hv321par05",,,"$noderes.nfssserver:/lite/state/$nodetype.profile",,
```

Note: Do not name your persistent storage directory with the node name, as the node name will be added in the directory automatically. If you do, then a directory named `/state/n01` will have its state tree inside `/state/n01/n01`.

2.5 Policy

Ensure policies are set up correctly. When a node boots up, it queries the xCAT database to get the `lite-files` and the `lite-tree`. In order for this to work, the `command` must be set in the policy table to allow nodes to request it. (This should happen automatically when xCAT is installed, but you may want to verify it if you experience problems booting.)

```
chtab priority=4.7 policy.commands=litefile policy.rule=allow
chtab priority=4.8 policy.commands=litetree policy.rule=allow
```

2.6 Add Linux Distro Packages

If you haven't already done so, copy the packages from the distro media into `/install`. For example:

```
copycds /iso/RHEL5.2-Server-20080430.0-x86_64-DVD.iso
```

Now create a list of distro packages that should be installed into the image. You should start with the base packages in the compute template. These are required. For example:

```
mkdir -p /install/custom/netboot/rh
cd /install/custom/netboot/rh
cp /netboot/rh/compute.pkglist compute.pkglist
```

You can then add more packages to the compute .pkglist by editing compute .pkglist.

2.7 Add Third-Party Software

If you have additional software that you want in the image, you can add it to the distro package directory that was created in the previous step and then rerun `createrepo`. But if you want to keep your distro package directory pristine, with only distro files in it, you can use xCAT's `otherpkgs` support:

Create a directory on your management node named `"/install/post/otherpkgs/<OS>/<ARCH>/"`, and put your third party rpm packages into the directory:

```
mkdir /install/post/otherpkgs/rhels5.3/x86_64
# put RPMs in there
```

Go to the directory `"/install/custom/netboot/<OS>/"` and add the names of the RPM packages you want in the image into the `<profile>.otherpkgs.pkglist` file.

```
cd /install/custom/netboot/rh
vi compute.otherpkgs.pkglist
```

Note: if some of these RPMs will need write permissions to files on the node, add those files or directories into the "litefile" table.

Create the .repolist file for your own repository in the same directory as above. The file name is: `<PROFILE>.<OS>.<ARCH>.repolist`. The file format is:

```
Type|URL|name
```

For example, if the repository is local, you can add the repository information as:

```
cd /install/custom/netboot/rh
vi compute.rh.repolist
plaindir|file:///install/post/otherpkgs/rhel5.3/x86_64|
```

otherpkgs

```
    If there's a remote repository, you can add the repository information as:  
    rpm-md|http://xcat.sourceforge.net/yum/xcat-dep/rhel5.3/x86_64|  
xcat-dep  
    rpm-md|http://xcat.sourceforge.net/yum/devel/core-snap|core-snap
```

2.8 Create Statelite Image

2.8.1 Prepare the ".postinstall" file

By editing the .postinstall file for your profile, you can automate the customization of the root image when the root image is generated by the "genimage" command.

XCAT supplies a basic *postinstall file with some **required** setup already in the script. Make sure there is at least one *postinstall file that will be used when you run genimage. The one shipped will setup fstab and rcons to work properly.

The *postinstall files can be found in "/opt/xcat/share/xcat/netboot/<OS>/" directories. If you need to create your own *postinstall file, first you should copy the basic setup supplied in the appropriate *postinstall files. You can add more postinstall process, if you want. If you do modify the script, you should save it in /install/custom/netboot/<os>, so it will not be overlaid with the next install.

Rules for evaluating postinstall file

There is a rule (2.4 release or later) for which * postinstall files will be selected to be used by genimage.

Use these basic rules to edit the correct file in the "/opt/xcat/share/xcat/netboot/<OS>/" directories. The rule allows you to customize you image down to the profile, os and architecture level, if needed.

You will find *postinstall files of the following forms:

- 1) <profile>.<os>.<arch>.postinstall
- 2) <profile>.<arch>.postinstall
- 3) <profile>.<os>.postinstall
- 4) <profile>.postinstall

genimage will select the *postinstall file from the list, if it exist, in the order 1-4. This means, if "<profile>.<os>.<arch>.postinstall" is there, it will be used. If there is no such a file, then the "<profile>.<arch>.postinstall" file will be used. If there's no such a file, then the "<profile>.<os>.postinstall" file will be used. If there is no such file, then it will use "<profile>.postinstall".

The *postinstall file is actually a bash script, you can add your own code to the basic function supplied in the file.

2.8.2 Run “genimage” command

After all our tables are set up, it is time to create a base image. In this example, the operating system is RedHat 5.3, and we will name our image “compute”, so that all our compute (group) nodes will use that image.

```
cd /opt/xcat/share/xcat/netboot/rh
./genimage -o rhels5.3 -p compute -i eth0
-n mlx4_core,mlx4_en,igb,bnx2
-m statelite
```

The genimage command will do several things:

1. It will create an image in

```
/install/netboot/<os>/<arch>/<profile>/rootimg
/install/netboot/rhels5.3/x86_64/compute/rootimg
```

2. It will create a ramdisk and kernel that can be used to boot the initial node.

One new option to genimage named “--permission” is added for statelite mode, you can assign user-definable permission for “/.statelite” directory. The default permission is 755, for example:

```
./genimage -o rhels5.3 -p compute -i eth0 -n
mlx4_core,mlx4_en,igb,bnx2 -m statelite --permission 777
```

would make the permission 777 instead of the 755 default

2.9 Setting up Post install scripts for statelite

The setup to run post install scripts for statelite is the same as for an type of install image in xCAT. The first kind of postscript is executed at genimage time, it is executed again the image itself on the MN . You need to create a file called \$profile.postinstall and put it under /install/custom/netboot/\$os directory. The \$profile.postinstall will get called in chroot env against the image before genimage is done. See section Prepare the “.postinstall” file .

The second kind of postscript is the script that runs on the node during node deployment time. During init.d timeframe, /etc/init.d/gettyset calls /opt/xcat/xcatdsklspost that in the image. /opt/xcat/xcatdsklspost uses wget via ftp to get all the postscripts under mn:/install/postscripts and copy them to /xcatpost directory. It uses openssl or stunnel to connect to the rmcd on the mn to get all the postscript names for the node from the postsctipts table. It then calls the postscripts for the node one by one. Among them, there is one called "remoteshell" from the "xcatdefaults" row that setup the ssh keys and other credentials between the mn and the node.

See the [TopDOC Appendix D](#) for setting up postinstall scripts.

2.10 Modify statelite image

Since the files that were now just created will be the default for all the files listed in the litefile table, you can edit the image directly by visiting the root tree in:

```
/install/netboot/<os>/<arch>/<profile>/rootimg  
cd /install/netboot/rhels5.3/x86_64/compute/rootimg
```

You can run `chroot` to make additional changes in the image, for example, `yum/zypper updates/install` using the `-installroot` flag.

For Red Hat it is necessary is to copy some passwd files and ssh settings from xCAT into the image. *Note: don't run the following commands for SLES.*

```
cd /opt/xcat/share/xcat/netboot/add-on/statelite  
./add_passwd /install/netboot/rhel5.3/x86_64/compute/rootimg  
./add_ssh /install/netboot/rhel5.3/x86_64/compute/rootimg
```

2.11 Installing a new Kernel in the statelite image (This is not verified)

Obtain you new kernel and kernel modules on the MN, for example here we have a new SLES kernel.

Copy the kernel into /boot :

```
cp vmlinux-2.6.32.10-0.5-ppc64 /boot
```

Copy the kernel modules into /lib/modules/<new kernel directory>

```
xcatlinuxmn:/lib/modules # ls -al  
total 16  
drwxr-xr-x 4 root root 4096 Apr 19 10:39 .  
drwxr-xr-x 17 root root 4096 Apr 13 08:39 ..  
drwxr-xr-x 3 root root 4096 Apr 13 08:51 2.6.32.10-0.4-ppc64  
drwxr-xr-x 4 root root 4096 Apr 19 10:12 2.6.32.10-0.5-ppc64
```

Run `genimage` to update the statelite image with the new kernel

```
cd /opt/xcat/share/xcat/netboot/sles
```

```
./genimage -i eth0 -n ehea -o sles11.1 -p compute -m statelite -k 2.6.32.10-0.5-ppc64
```

then after a `nodeset` command and `netboot..` the `uname -a` shows the new kernel.

2.12 Run liteimg command

The `liteimg` command will modify your statelite image (the image that `genimage` just

created) by creating a series of links. Once you are satisfied your image contains what you want it to, run “`liteimg <os>-<arch>-<profile>`”, or “`liteimg -o <os> -a <arch> -p <profile>`”. For example:

```
liteimg rhels5.3-x86_64-compute
or
liteimg -o rhels5.3 -a x86_64 -p compute
```

This creates 2 levels of indirection, so that files can be modified while in their image state as well as during runtime. For example, a file like “`$imageroot/etc/ntp.conf`” with default option in the “`litefile`” table, will have the following operations done to it:

In our case `$imageroot` is `/install/netboot/rhels5.3/x86_64/compute/rootimg`

```
mkdir -p $imageroot/.default/etc
mkdir -p $imageroot/.statelite/tmpfs/etc
mv $imgroot/etc/ntp.conf $imgroot/.default/etc
cd $imgroot/.statelite/tmpfs/etc
ln -sf ../../../../default/etc/ntp.conf .
cd $imgroot/etc
ln -sf ../.statelite/tmpfs/etc/ntp.conf .
```

When finished, the original file will reside in `$imgroot/.default/etc/ntp.conf`. `$imgroot/etc/ntp.conf` will link to `$imgroot/.statelite/tmpfs/etc/ntp.conf` which will in turn link to `$imgroot/.default/etc/ntp.conf`

Note: If you make any changes to your `litefile` table after running `liteimg` then you will need to rerun `liteimg` again. This is because files and directories need to have the two levels of redirects created.

2.13 Set the boot state to “`statelite`”

You can now deploy the node:

```
nodeset <noderange> statelite=rhels5.3-x86_64-statelite-compute
```

Or just run:

```
nodeset <noderange> statelite
```

This will create the necessary files in `/tftpboot` for the node to boot correctly.

2.14 Install the Node

Finally, reboot the node so that it boots up in `statelite` mode.

For `x86_64` platform:

```
rpower <noderange> boot
```

Nodeset will have generated the appropriate PXE file so that the node boots from the nfsroot image. This file will look similar to the following:

```
#statelite rhel5.3-x86_64-compute
DEFAULT xCAT
LABEL xCAT
  KERNEL xcat/netboot/rhel5.3/x86_64/compute/kernel
  APPEND initrd=xcat/netboot/rhel5.3/x86_64/compute/initrd.gz
NFSROOT=172.10.0.1:/install/netboot/rhel5.3/x86_64/compute
STATEMNT=cnfs:/gpfs/state XCAT=172.10.0.1:3001 console=tty0
console=ttyS0,115200n8r
```

For POWER platform:

Run the following command to boot up the nodes into statelite mode:

```
rnetboot <noderange>
```

Nodeset will have generated the appropriate yaboot.conf-MAC-ADDRESS file so that the node boots off the nfsroot image. This file will look similar to the following:

```
#statelite rhel5.3-ppc64-compute
timeout=5
image=xcat/netboot/rhel5.3/ppc64/compute/kernel
  label=xcat
  initrd=xcat/netboot/rhel5.3/ppc64/compute/initrd.gz
  append="NFSROOT=192.168.11.108:/install/netboot/rhel5.3/ppc64/c
ompute STATEMNT= XCAT=192.168.11.108:3001 "
```

You can then use rcons or wcons to watch the node boot up.

3 New Commands

The following commands are in /opt/xcat/bin:

litefile <nodename>

Shows all the statelite files that are not to be taken from the base of the image.

litetree <nodename>

Shows the NFS mount points for a node.

litefile <image name>

Shows the stateless files that will be used for a node image.

liteimg <image name>

Creates a series of symbolic links in an image that is compatible with statelite

booting.

4 Stalite Directory Structure

Each stalite image will have the following directories:

```
/.stalite/tmpfs/  
/.stalite/persistent/<nodename>  
/.stalite/mnt # where directory tree is mounted from.  
/.default/  
/etc/init.d/stalite
```

All files that are symbolic links, will link to `/.stalite/tmpfs`.

`tmpfs` files that are persistent link to `/.stalite/persistent/<nodename>/`
`/.stalite/persistent/<nodename>` is the directory where the node's individual storage will be mounted to.

`/.default` is where default files will be copied to from the image to `tmpfs` if the files are not found in the `litetree` hierarchy.

4.1 The noderes Table

`noderes.nfsserver` attribute can be set for the NFSroot server. If this is not set, then the default is the Management Node.

`noderes.nfsdir` – can be set. If this is not set, the the default is `/install`

5 Adding/updating software and files for the running nodes

This needs work (do we shutdown the nodes before doing this?)

Because most of system files for the nodes are NFS mounted on the Management Node with read-only option, installing or updating software and files should be done to the image. The image is located under `/install/netboot/<os>/<arch>/<profile>/rootimg` directory.

To install or update an rpm, do the following:

```
1. rpm --root /install/netboot/<os>/<arch>/<profile>/rootimg rpm_name
```

```
2. Restart the software application on the nodes.
```

```
xdsh <noderange> restart_this_software
```

It is recommended to follow section 2.7 (Addig third party software) to add the new rpm to the `otherpkgs.pkglist` file so that the rpm will get installed into the new

image next time the image is rebuilt.

Note: The newly added rpms are not shown when running `rpm -qa` on the nodes although the rpm is installed. It will shown next time the node is reboot.

To create or update a file for the nodes, just modify the file in the image and restart any application that uses the file.

6 Hierarchy Support

In the statelite environment, the service node needs to provide NFS service for the compute node with statelite, the service nodes must to be setup with diskfull installation.

6.1 Setup the diskfull service node

To setup one diskfull service node, please refer to the document "[How to setup Hierarchy in xCAT 2](#)". Since statelite is a kind of NFS-hybrid method, you should remove the "installloc" attribute in the site table. This makes sure that the service node does not mount the /install directory from the management node on the service node.

6.2 Generate the statelite image

The chapter 2 has described how to generate the statelite image for your own profile.

NOTE: if the NFS directories defined in the litetree table are on the service node, it is better to setup the NFS directories in the service node following the chapter 2.3.

6.3 Sync the "/install" directory

The command "prsync" is used to sync the "/install" directory to the service nodes. Run the following:

```
cd /
prsync install <sn>:/
```

<sn> is the hostname of the service node you defined.

Since the "prsync" command will sync all the contents in the "/install" directory to the service nodes, the first time will take a long time. But after the first time, it will take very short time to sync.

NOTE: if you make any changes in the "/install" directory on the management node, and the changes can affect the statelite image, you need to sync the "/install" directory to the service node again.

6.4 Set the boot state to "statelite"

Please follow the instructions in Set the boot state to "statelite".

6.5 Install the Node

Please follow the instructions in Install the Node.

7 Advanced Stalite features

7.1 litetree table

The litetree table controls where the initial content of the files in the litefile table come from, and the long term content of the “ro” files. When a node boots up in stalite mode, it will by default copy all of its tmpfs files from the `/.default` directory of the root image, so there is not requirement to setup a litetree table. If you decide that you want some of the files pulled from different locations that are different per node, you can use this table.

See [litetree man page](#) for description of attributes.

For example, a user may have two directories with a different `/etc/motd` that should be used for nodes in two locations:

```
10.0.0.1:/syncdirs/newyork-590Madison/rhels5.4/x86_64/compute/etc/motd
10.0.0.1:/syncdirs/shanghai-11foo/rhels5.4/x86_64/compute/etc/motd
```

You can specify this in one row in the litetree table:

```
1,,10.0.0.1:/syncdirs/$nodepos.room/$nodetype.os/$nodetype.arch/
$nodetype.profile
```

When each stalite node boots, the variables in the litetree table will be substituted with the values for that node to locate the correct directory to use. Assuming that `/etc/motd` was specified in the litefile table, it will be searched for in all of the directories specified in the litetree table and found in this one.

You may also want to look by default into directories containing the node name first:

```
$noderes.nfsserver:/syncdirs/$node
```

The litetree prioritizes where node files are found. The first field is the priority. The second field is the image name (ALL for all images) and the final field is the mount point.

Our example is as follows:

```
1,,$noderes.nfsserver:/stalite/$node
2,,cnfs:/gpfs/dallas/
```

The two directories `/stalite/$node` on the node’s `$noderes.nfsserver` and the `/gpfs/dallas` on the node `cnfs` contain root tree structures that are sparsely populated with files that we want to place in those nodes. If files are not found in the first directory, it goes to the next directory. If none of the files can be found in the litetree hierarchy, then they are searched for in `/.default` on the local image.

8 Debugging techniques

1. When a node boots up in statelite mode, there is a script that runs called statelite that is in the root directory of \$imageroot/etc/init.d/statelite. This script is not run as part of the rc scripts, but as part of the pre-switch root environment. Thus, all the linking is done in this script. There is a “set -x” near the top of the file. You can uncomment it and see what the script runs. You will then see lots of mkdirs and links on the console.
2. You can also set the machine to shell. Just add the word “shell” on the end of the pxeboot file of the node in the append line. This will make the init script in the initramfs pause 3 times before doing a switch_root.
3. When all the files are linked they are logged in /.statelite/statelite.log on the node. You can get into the node after it has booted and look in the /.statelite directory.