



Using FlexBoot for Booting RHEL 6.4 from an iSCSI Target

User Manual

Rev 1.0

www.mellanox.com

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT ("PRODUCT(S)") AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES "AS-IS" WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies
350 Oakmead Parkway Suite 100
Sunnyvale, CA 94085
U.S.A.
www.mellanox.com
Tel: (408) 970-3400
Fax: (408) 970-3403

Mellanox Technologies, Ltd.
Beit Mellanox
PO Box 586 Yokneam 20692
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

© Copyright 2013. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, CORE-Direct®, InfiniBridge®, InfiniHost®, InfiniScale®, MLNX-OS®, PhyX®, SwitchX®, UFM®, Virtual Protocol Interconnect® and Voltaire® are registered trademarks of Mellanox Technologies, Ltd.

Connect-IB™, ExtendX™, FabricIT™, Mellanox Open Ethernet™, Mellanox Virtual Modular Switch™, MetroX™, MetroDX™, ScalableHPC™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Contents

Preface	4
1 Using FlexBoot for Booting RHEL 6.4 from an iSCSI Target	5
1.1 Configuring the iSCSI Target Machine	5
1.2 Installing RHEL6.4 on a Remote Storage over iSCSI	6
1.3 SAN-Bootting the Diskless Client with FlexBoot	10
Appendix A: Sanity Checks	12
A.1 Sanity Check: iSCSI Login	12
A.2 DHCP Configuration for PXELINUX with FlexBoot	13
A.3 pxelinux.cfg/default	14
A.4 DHCP Configuration for iSCSI Boot with FlexBoot (PXE SAN Boot).....	14
Appendix B: SAN-Bootting with FlexBoot in CHAP Environment	15

Preface

About this Manual

This User Guide provides guidance for booting diskless PXE clients to RHEL6.4 from remote storage over iSCSI, using Mellanox Network technology card and boot agent (FlexBoot) over Ethernet Network.

The setting-up and configuration of the iSCSI target and PXE server as described within this document are merely recommendations and provided as is, and may be applied according to requirements of the environment in which the actual work takes place.

The actions described in this paper are focused on Mellanox Network technology solution.

Document Conventions

The following lists conventions used in this document.



NOTE: Identifies important information that contains helpful suggestions.



CAUTION: Alerts you to the risk of personal injury, system damage, or loss of data.



WARNING: Warns you that failure to take or avoid a specific action might result in personal injury or a malfunction of the hardware or software. Be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents before you work on any equipment.

1 Using FlexBoot for Booting RHEL 6.4 from an iSCSI Target

Below are instructions on how to provision a diskless system (the client) with a fresh RHEL 6.4 installation to a remote storage (IE: a LUN partition on an iSCSI target) and then SAN-Boot (iSCSI boot) the client using Mellanox PXE boot agent (FlexBoot). The iSCSI configuration in this document is very basic (no CHAP authentication, no multipath I/O) and demonstrates basic PXE SAN Boot capability.

In this document, IET is used as the iSCSI Target.

In the example below:

- sq030 (a Linux host) - performs the role of the PXE server and an iSCSI target and
- sq070 (a host) - performs the role of the PXE client and a sanity check environment
- sq030 – The OS is configured with IP 12.7.6.30, and the DHCP IP configuration for the diskless client is 12.7.6.70
- sq030 - has a 21.5GB disk partition on its local HDD, called /dev/cciss/c0d0p9, which will be used as the storage volume to host the client's OS. It also serves as DHCP, TFTP, and NFS server for PXE clients
- Both servers have a Mellanox ConnectX-3 based 10GE NIC equipped with PXE boot capabilities via Expansion ROM called FlexBoot of version 3.4.149

1.1 Configuring the iSCSI Target Machine

➤ To configure the iSCSI target:

1. Download the IET target software from:

<http://sourceforge.net/projects/iscsitarget/files/iscsitarget/1.4.20.2/>

2. Install iSCSI target and additional required software on target server

```
[root@sq030 ~]# yum install kernel-devel openssl-devel gcc rpm-build
[root@sq030 ~]# cd /tmp/
[root@sq030 tmp]# tar xzvf
./autodirect/QA/venters/iscsitarget-1.4.20.2.tar.gz
[root@sq030 tmp]# cd iscsitarget-1.4.20.2/
[root@sq030 iscsitarget-1.4.20.2]# make && make install
```

3. Create the IQN in the ietd configuration file.

```
[root@sq030 ~]# vi /etc/iet/ietd.conf
## Optional: for CHAP authentication, uncomment the following #lines
#IncomingUser joe secret
#OutgoingUser jack 12charsecret
Target iqn.2013-10.qalab.com:sq030.prt9
    Lun 0 Path=/dev/cciss/c0d0p9,Type=fileio,IOMode=wb
    ImmediateData Yes
    MaxConnections 1
    InitialR2T Yes
```



The local Hard Disk partition assigned to the LUN (/dev/cciss/c0d0p9 in the example above) must not contain any valuable data, as this data will be destroyed by the installation process taking place later in this procedure

4. Edit the /etc/sysconfig/iscsi-target file as follow.

```
OPTIONS="-c /etc/iet/ietd.conf --address=12.7.6.30"
```

5. Start the iSCSI target service.

```
[root@sqa030 ~]# /etc/init.d/iscsi-target start
```

Perform a sanity check by connecting to the iSCSI target from a remote PC on the 10GE network link. For further information, please refer to section [Sanity Checks](#).

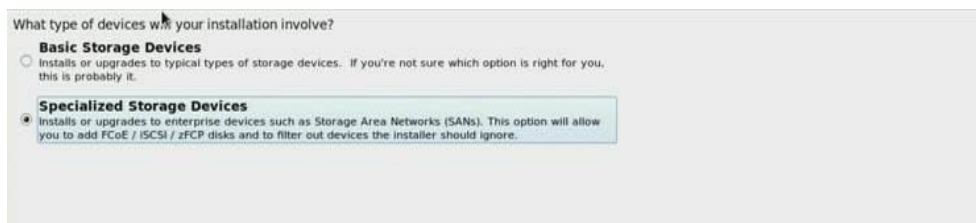
1.2 Installing RHEL6.4 on a Remote Storage over iSCSI

1. Reboot the diskless client and perform a PXE boot with FlexBoot.

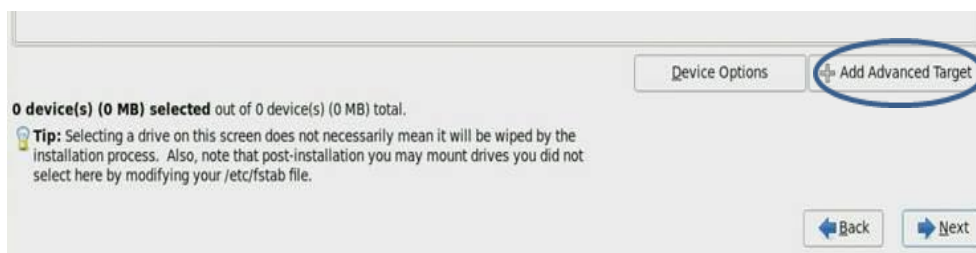
This is not an iSCSI boot, rather a regular PXE initiated network deployment of RHEL6.4. In the DHCP server configuration, the PXELINUX (pxelinux.0) and a RHEL 6.4 distribution media will be provided for network installation. For further information, please refer to sections A.2 [DHCP Configuration for PXELINUX with FlexBoot](#) and [pxelinux.cfg/default](#). The clients' HDD was removed beforehand; therefore the RHEL installer (A.K.A Anaconda) will ask to locate a HDD. The Anaconda's built-in iSCSI discovery will be used to connect to the iSCSI target LUN partition. For further information, please refer to:

https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html-single/Installation_Guide/index.html#ISCSI_disks

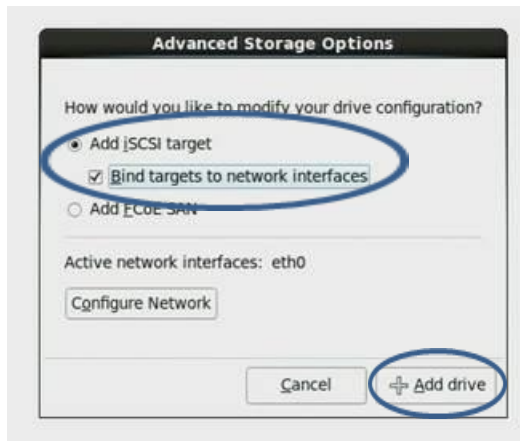
2. Select the Network Interface for the installation process once prompted (after the installer is loaded and starts running on the client). Select the same interface which was used by FlexBoot during PXE boot stage.
3. Select the type of Installation Media access. In this example, we use NFS, which also requires us to enter the NFS server name, and the directory path to the installation media on the NFS.
4. Select **Specialized Storage Devices**.



5. Click on the + **Add Advanced Target** button.



6. In the Advanced Storage Options window perform the following:
 - a. Select the **Add iSCSI target** option.
 - b. Check the **Bind targets to network interfaces** checkbox.
 - c. Click **+Add drive** button.



7. Enter the IP address of iSCSI target.

Optionally, you may choose to enter a customized Initiator Name and select the necessary CHAP authentication of choice. Refer to [SAN-Booting with FlexBoot in CHAP Environment](#). In the example below, *iSCSI Initiator Name* is left with the default value given by the installer and *iSCSI discovery authentication* is left with *No authentication*.



8. Check the relevant Node Name to log in.

If as a result of the discovery, multiple Node Names are found, select the one that is relevant to you.



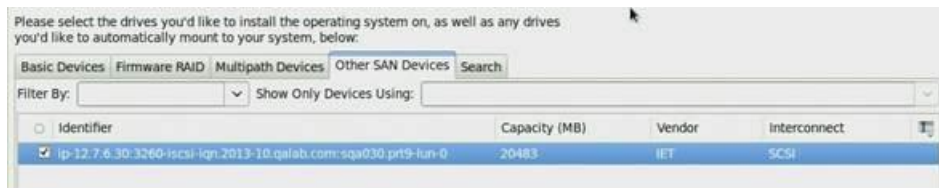
9. Click **Login**.

A successful login is mandatory to proceed. A failure at this stage is probably a result of a target or network configuration error and recovery/troubleshooting that is out of the scope of this document.

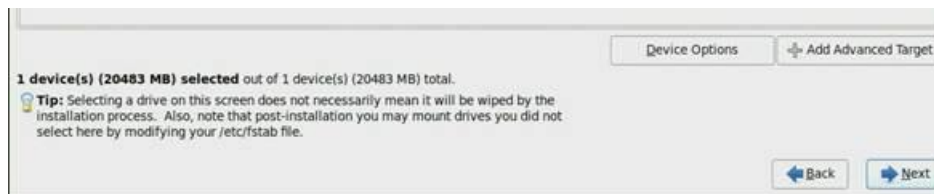


10. Make sure a new storage LUN appears in the **Other AN Devices** tab.

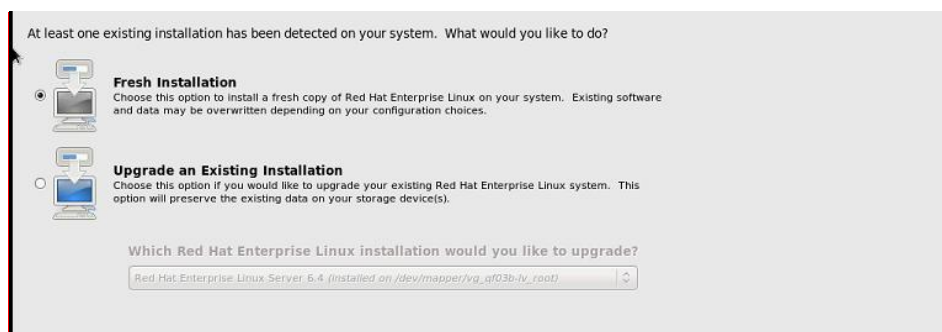
A successful LUN discovery is mandatory to proceed. A failure at this stage is probably a result of a target or network configuration error and recovery/troubleshooting that is out of the scope of this document.



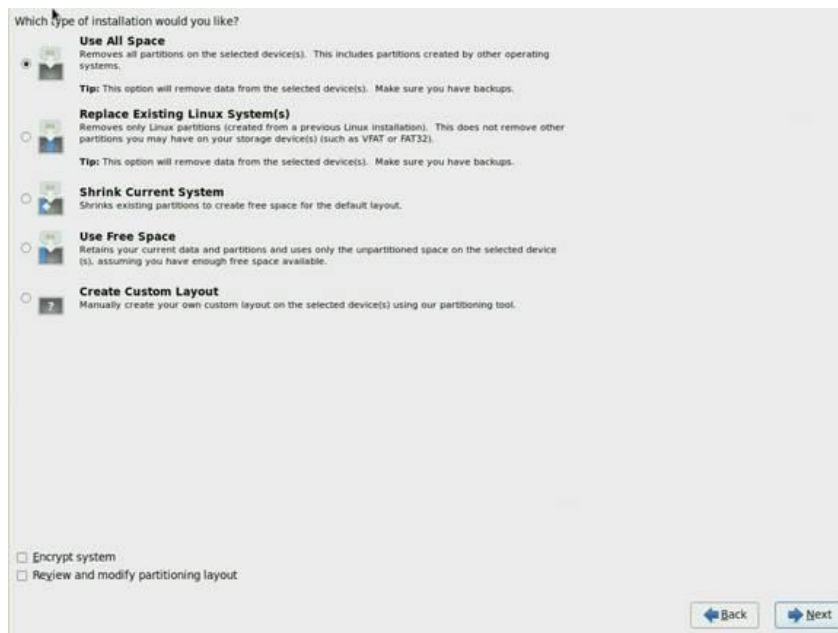
11. Click **Next**.



12. Select **Fresh Installation** and proceed with the Installation



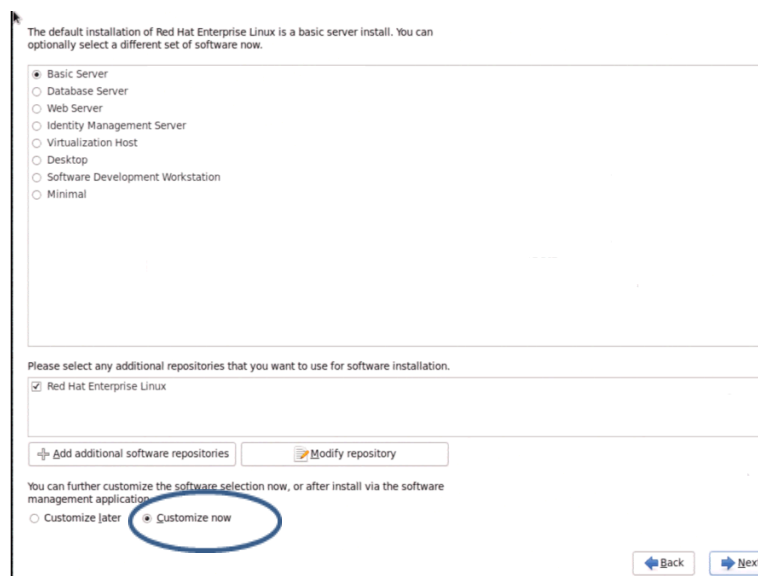
13. Select the **Use All Space** option.



14. Click **Next** and proceed with the Installation.

15. Select the **Basic Server** option.

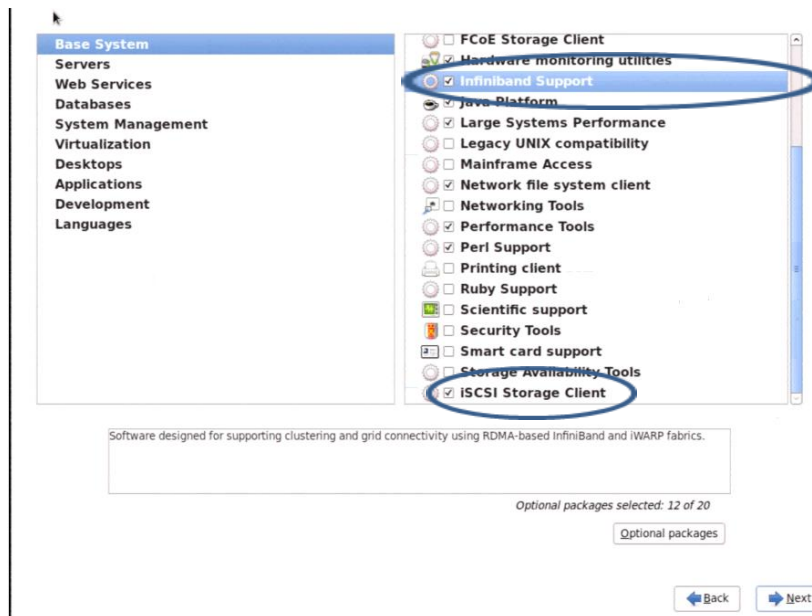
This is only one of the options that can be chosen, not the mandatory one.



16. Check the **Customize Now** checkbox.

17. Click **Next**.

18. Select **Infiniband Support** and **iSCSI Storage Client**.



19. Click **Next**. Allow the installation to reach completion.

1.3 SAN-Booting the Diskless Client with FlexBoot

When the installation process is completed, the client will ask to reboot. At that point, the DHCP server configuration for that client needs to be changed so that when it PXE boots again, it will get the root-path IQN and LUN information from the DHCP server.

For further information, please refer to section [DHCP Configuration for iSCSI Boot with FlexBoot \(PXE SAN Boot\)](#).



NOTE: Restart your DHCP service after changing the dhcp configuration file.

- Reboot the system



The expected result now is that the diskless PXE client will boot the newly installed RHEL6.4 from the iSCSI storage, and become an operational environment, accessible from any remote PC via ssh over 10GbE IP network

Appendix A: Sanity Checks

A.1 Sanity Check: iSCSI Login

From a remote PC (called *sqa070* below) with 10GE network connection to the iSCSI target, configure an iSCSI Initiator service and verify the correct target configuration by logging into the target. For CHAP configuration, refer to [SAN-Booting with FlexBoot in CHAP Environment](#).

1. Install the initiator.

```
[root@sqa070 ~]# yum install -y iscsi-initiator-utils
```

2. Configure the Initiator.

```
[root@sqa070 ~]# vim /etc/iscsi/iscsid.conf
node.startup = automatic
## Optional: for CHAP authentication, uncomment the following #lines
#discovery.sendtargets.auth.authmethod = CHAP
#discovery.sendtargets.auth.username = joe
#discovery.sendtargets.auth.password = secret

#node.session.auth.authmethod = CHAP
#node.session.auth.username = jack
#node.session.auth.password = 12charsecret

node.session.timeo.replacement_timeout = 120
node.conn[0].timeo.login_timeout = 15
node.conn[0].timeo.logout_timeout = 15
node.conn[0].timeo.noop_out_interval = 10
node.conn[0].timeo.noop_out_timeout = 15
node.session.initial_login_retry_max = 10
node.session.cmds_max = 128
node.session.queue_depth = 32
node.session.iscsi.InitialR2T = No
node.session.iscsi.ImmediateData = Yes
node.session.iscsi.FirstBurstLength = 262144
node.session.iscsi.MaxBurstLength = 16776192
node.conn[0].iscsi.MaxRecvDataSegmentLength = 131072
discovery.sendtargets.iscsi.MaxRecvDataSegmentLength = 32768
node.session.iscsi.FastAbort = No
```

3. Start the iSCSI Initiator service.

```
[root@sqa070 ~]# service iscsi start
```

4. Discover the iSCSI target host.

In the example below, the IP address 12.7.6.30 is the iSCSI target.

```
[root@sqa070 ~]# iscsid start
[root@sqa070 ~]# iscsiadm -m discovery -t st -p 12.7.6.30
Starting iscsid:
12.7.6.30:3260,1 iqn.2013-10.qalab.com:sqa030.prt9 [ OK ]
```

Achieving a successful target discovery at this stage is mandatory for proceeding with the process of iSCSI boot. A failure at this stage is probably a result of an erroneous target or network configuration and troubleshooting that is out of the scope of this document.

5. Log into the target.

```
[root@sqa070 ~]# iscsiadm -m node -p 12.7.6.30 -T
iqn.2013-10.qalab.com:sqa030.prt9 --login
Logging in to [iface: default, target: iqn.2013-10.qalab.com:sqa030.prt9,
portal: 12.7.6.30,3260] (multiple)
```

```
Login to [iface: default, target: iqn.2013-10.qalab.com:sqa030.prt9, portal:
12.7.6.30,3260] successful.
```

A successful LUN login at this stage is mandatory for proceeding with the process of iSCSI boot. A failure at this stage is probably a result of an erroneous target or network configuration and troubleshooting that is out of the scope of this document.

6. Verify the remote partition appears to the initiator as a local HDD

```
[root@sqa070 ~]# fdisk -l

Disk /dev/sda: 500.1 GB, 500107862016 bytes
255 heads, 63 sectors/track, 60801 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x000518f2

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1  *           1          131     1048576   83   Linux
Partition 1 does not end on cylinder boundary.
/dev/sda2             131        2681     2048000    83   Linux
/dev/sda3             2681        2936      204800    82   Linux swap /
Solaris

Disk /dev/sdb: 21.5 GB, 21478670336 bytes
64 heads, 32 sectors/track, 20483 cylinders
Units = cylinders of 2048 * 512 = 1048576 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00000000
```

A.2 DHCP Configuration for PXELINUX with FlexBoot

The following DHCP configuration is presented as is, and may not work in all environments.

```
authoritative;
ddns-update-style none ;
ddns-updates off ;
allow bootp;
always-broadcast off ;
always-reply-rfc1048 off;
boot-unknown-clients on ;

option client-system-architecture          code 93 = unsigned integer 16;
option vendor-encapsulated-options         code 43 = string;
option vendor-class-identifier             code 60 = string;

class "PXECient" {
    match if substring (option vendor-class-identifier, 0, 9) = "PXECient";
    option vendor-class-identifier "PXECient";
    option vendor-encapsulated-options 06:01:08 ;
    option dhcp-parameter-request-list = concat(option
dhcp-parameter-request-list,43);
}

subnet 12.7.0.0 netmask 255.255.0.0 {
    option dhcp-server-identifier 12.7.6.30 ;
    option domain-name "pxe030.mtl.com" ;
    option domain-name-servers 12.7.6.30 ;
    default-lease-time 86400 ; # 1 day
    max-lease-time 86400 ;
    option ntp-servers 12.7.6.30;
}
```

```
host sqa070 {
    fixed-address 12.7.6.70 ;
    hardware ethernet 00:02:c9:32:e8:80 ;
    next-server 12.7.6.30;
    if option client-system-architecture = 00:00 { filename "pxelinux.0" ;
    }
}
```

A.3 pxelinux.cfg/default

```
[root@sqa030 ~]# cat /var/lib/tftpboot/pxelinux.cfg/default
LABEL rh6.4x64_inst1_manual
MENU LABEL Manual Installation RHEL6.4
    KERNEL RHEL6.4-x86_64-DVD1/images/pxeboot/vmlinuz
    APPEND initrd=RHEL6.4-x86_64-DVD1/images/pxeboot/initrd.img
```

A.4 DHCP Configuration for iSCSI Boot with FlexBoot (PXE SAN Boot)

Modify the following host declaration in your DHCP configuration file to allow PXE SAN Boot

```
host sqa070 {
    filename "";
    option root-path "iscsi:12.7.6.30::::iqn.2013-10.qalab.com:sqa030.prt9";
    fixed-address 12.7.60.70 ;
    hardware ethernet 00:02:c9:32:e8:80 ;
}
```

Appendix B: SAN-Booting with FlexBoot in CHAP Environment

For Successful SAN-Booting with FlexBoot in CHAP environment, the FlexBoot in use must support passing CHAP credentials. Please note that FlexBoot v3.4.149 does not support such operation. Hence, use FlexBoot to chain-load an UNDI software which supports passing CHAP credentials. This can be achieved by using the UNDI software module from iPXE.org, called *undionly.kpxe*. A Linux host is required to create *undionly.kpxe*.

1. Install the below prerequisite software to support necessary UNDI's compilation.

```
[root@sqa030 ~]# yum install -y gcc binutils make perl
```

2. Download the UNDI sources from iPXE.org.

(For More info, visit <http://www.ipxe.org/download>)

```
[root@sqa030 ~]# git clone git://git.ipxe.org/ipxe.git
```

3. Edit a command file named *sanbootnchap.ipxe* (the name is given as an example whereas the .ipxe file extension is mandatory) with the following lines.

Make sure to enter your own values for username and password per your CHAP configuration. For reasons of simplicity, and coherence with this document examples, we gave our CHAP the username *joe*, and the password *secret*. Note that these CHAP settings are going to be hardcoded into this module and any modification done to them will require you to recompile the module after changing *sanbootnchap.ipxe*.

```
[root@sqa030 ~]# cd ipxe/src
[root@sqa030 src]# vim sanbootnchap.ipxe
#!ipxe
dhcp || dhcp || exit
set username joe
set password secret
echo ${root-path}
isset ${root-path} || exit
echo Booting from iSCSI tgt
sanboot --no-describe ${root-path}
```

4. Compile the *undionly.kpxe* module.

Refer to <http://www.ipxe.org/howto/chainloading>

```
[root@sqa030 src]# make bin/undionly.kpxe EMBED=sanbootnchap.ipxe
```

5. Copy the bin/*undionly.kpxe* to your TFTP root directory (E.G.: /var/lib/tftpboot).
6. Edit the client's host declaration in the DHCP configuration file for chain-loading *undionly.kpxe*.

The outcome of this procedure is to have FlexBoot download *undionly.kpxe* to the client's RAM, and then have *undionly.kpxe* authenticate iSCSI and login with the iSCSI target.

```
host sqa070 {
    next-server 12.7.6.30;
    if option client-system-architecture = 00:00 { filename "undionly.kpxe" ;
    }
    fixed-address 12.7.60.70 ;
    hardware ethernet 00:02:c9:32:e8:80 ;
    if exists user-class and option user-class = "iPXE" {
        option root-path
        "iscsi:12.7.6.30:::iqn.2013-10.qalab.com:sqa030.prt9";
        filename "";
    }
}
```

}

For CHAP users: All the CHAP authentication lines mentioned as comments in the iSCSI target and initiator configuration examples in sections [Configuring the iSCSI Target Machine](#), and [Sanity Check: iSCSI Login](#) must be uncommented beforehand, and during RHEL installation, follow these steps:



iSCSI Discovery Details

To use iSCSI disks, you must provide the address of your iSCSI target and the iSCSI initiator name you've configured for your host.

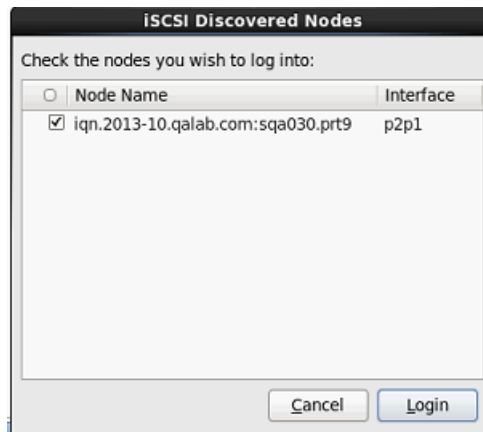
Target IP Address:

iSCSI Initiator Name:

What kind of iSCSI **discovery authentication** do you wish to perform:

CHAP Username:

CHAP Password:



iSCSI Discovered Nodes

Check the nodes you wish to log into:

☐	Node Name	Interface
☒	iqn.2013-10.qalab.com:sqa030.prt9	p2p1



iSCSI Nodes Login

What kind of iSCSI **login authentication** do you wish to perform:

CHAP Username:

CHAP Password:



iSCSI Login Results

Successfully logged in and attached the following nodes:

iqn.2013-10.qalab.com:sqa030.prt9 via p2p1