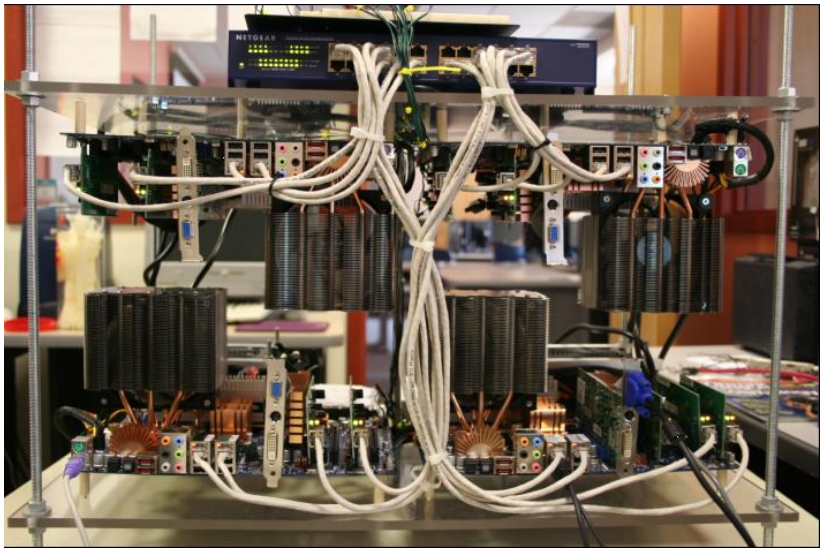


Quadrowulf Software and Network Configuration Notes

By Justin Moore
August 8, 2008

Based on:
Microwulf Software and Network Configuration Notes
By Tim Brom – 5.16.2008

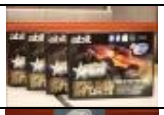
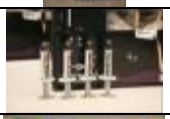
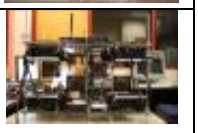
On this page I will be recreating what was done by Mr. Tim Brom at Calvin College with Professor Joel Adams at this web page. I am going to try to be a bit more verbose in my description of the creation of Quadrowulf. That is not to say that Mr. Brom's paper was not helpful, because it was indeed helpful in the inspiration, creation, and implementation of this project. This project was not really class work related – more, I was interested in implementing it and my professor, Dr. Hayden S. Porter, of Furman University, gave me a chance to build it for him. It should be noted that this project uses **quad-core** processors versus the dual-core processors used in Microwulf.

Hardware Components

First, my hardware:

I have created a similar cluster to Tim's, except that I have 2 layers instead of four, and I'm using full ATX boards instead of m-ATX. So, this isn't really “Microwulf,” but it's still very manageable to move around with just two people. Right now it resides on a cart, so it’s even easier to move. See the Table 1.1 below for the part list that accompanies this picture.

Table 1.1

Part	***Brand / Model***	***Picture***
Motherboard	Abit IP35PRO Off Limits	
Processor	Intel Core2Quad Q6600 Overclocked to 3.2 Ghz	
Heatsink	Tuniq Tower	
Thermal Paste	Artic Silver 5	
Video Card	Biostar GeForce Series 6200LE [V6202EL – 63 VER:7.0]	
Network Switch	Netgear ProSafe 24-port Gigabit Switch	
Memory (RAM)	Corsair XMS2 DDR2-800 4GB [2x2GB]	
Power Supply	Corsair VX550W Power Supply	
Network Cards	Linksys 10/100/1000 Gigabit EG1032 PCI Cards	
Hard Drive	Seagate 500GB Barracuda 7200.11	
CD/DVD Drive	Sony NEC Optiarc	
Network Cable	I made them custom lengths	
Mounting hardware	Rods/Nuts/Bolts/Washers	
Plexiglas	Two sheets 3/8” thick	

I mention in the parts list that I've overclocked the quad cores to run at 3.2 GHz in part because some of the additional cost accrued is because we needed additional materials to compensate the heat generated by the overclocked processors, e.g., the snazzy Tuniq Tower heatsink and the Artic Silver thermal paste. I am going to use pictures to walk you through the installation process of getting a new node to work properly, after all the steps have been performed on the head node, e.g. setting up the DHCP server, NFS server, and the TFTP server.

Tim's instructions for installing the head node are great; my only modification will be that I'm now using 5 Ethernet cards on the head node instead of 3, and 4 on each diskless node instead of 2. I have also partitioned 2x as many partitions on my drive because I wanted to make sure that I can get 32bit version of everything

working before I installed the 64bit versions. (: Professor Porter uses mpiJava, which somewhat died as a project, so we have to make sure that still works.

It appears now that I will have to leave it at the 32-bit version installed and let the next contender work with the 64-bit OS.

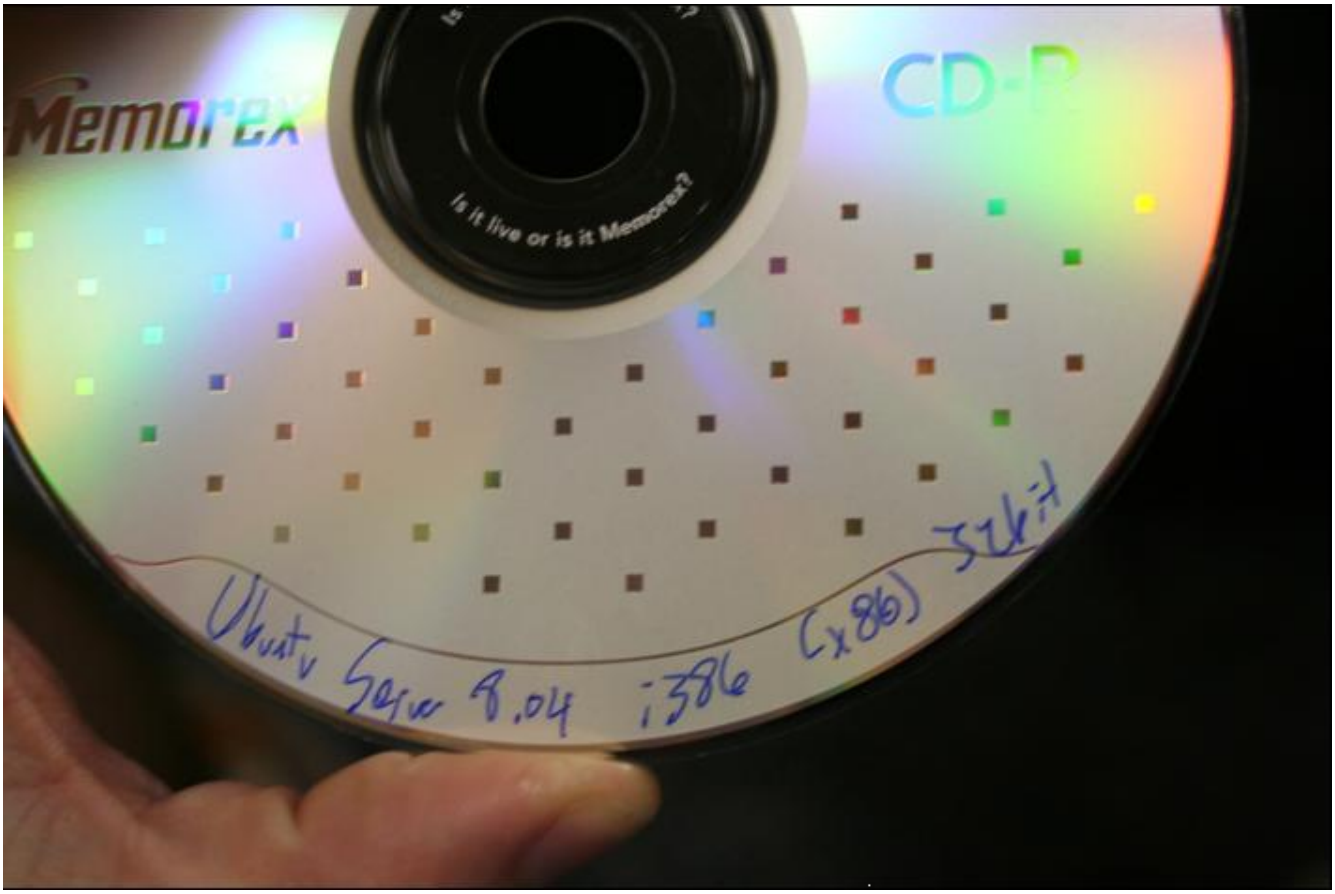
Note: *I changed the BIOS from hard drive recognition from IDE->AHCI because of a hang problem with the Ubuntu installation disks. If when you put the disk in, it just randomly drops to BusyBox after a long time (failsafe) then you should either change the BIOS like I did, or pass a boot parameter to the Linux install disk that is 'noapic' or something along those lines. Then it will work.*

Diskless node installation

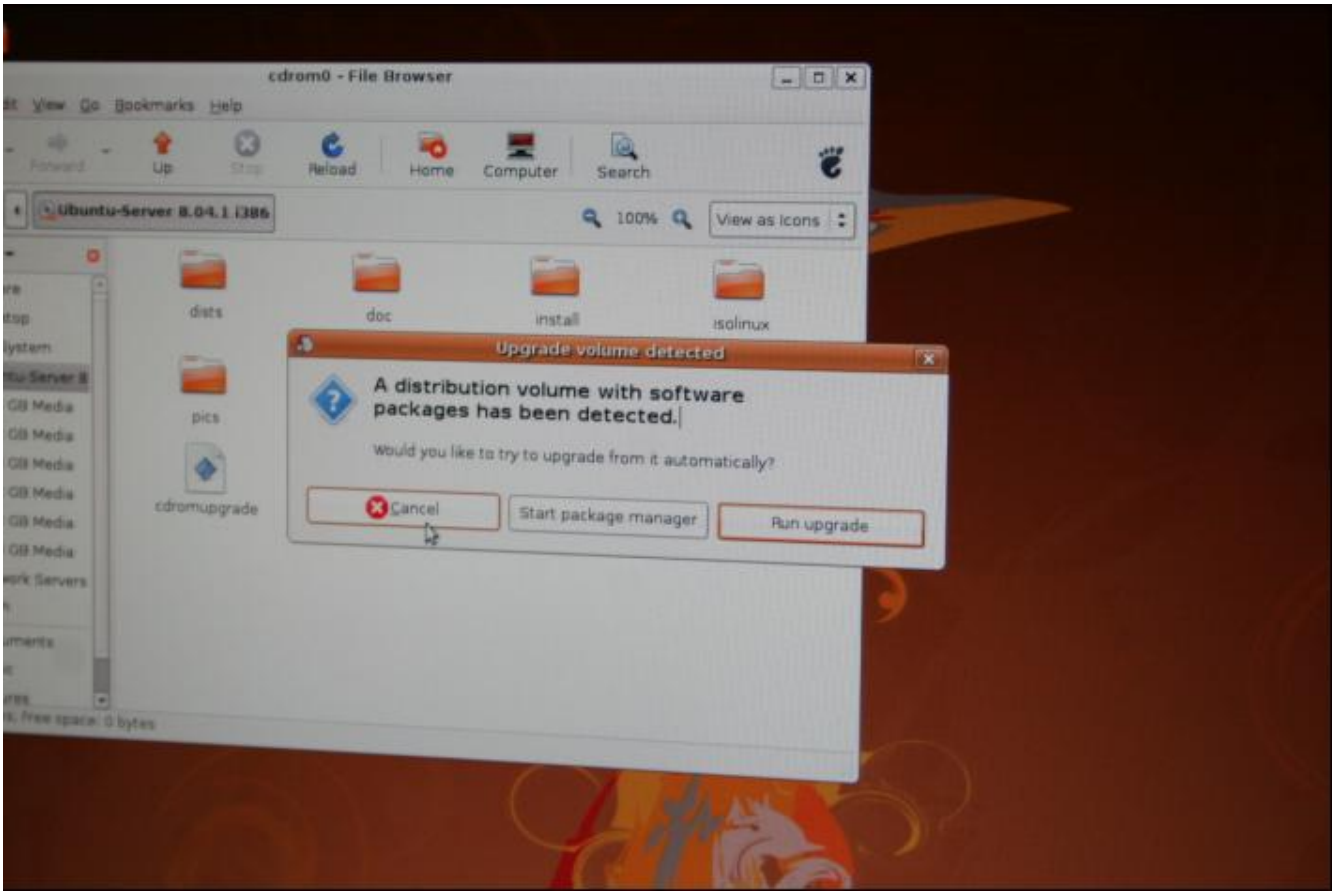
OK – now for the installation of a headless node. Again, I am assuming that you've already assembled the cluster and installed the OS [I'm using Ubuntu 8.04 Desktop] on the head node (with suitable partition tables). Here we go.

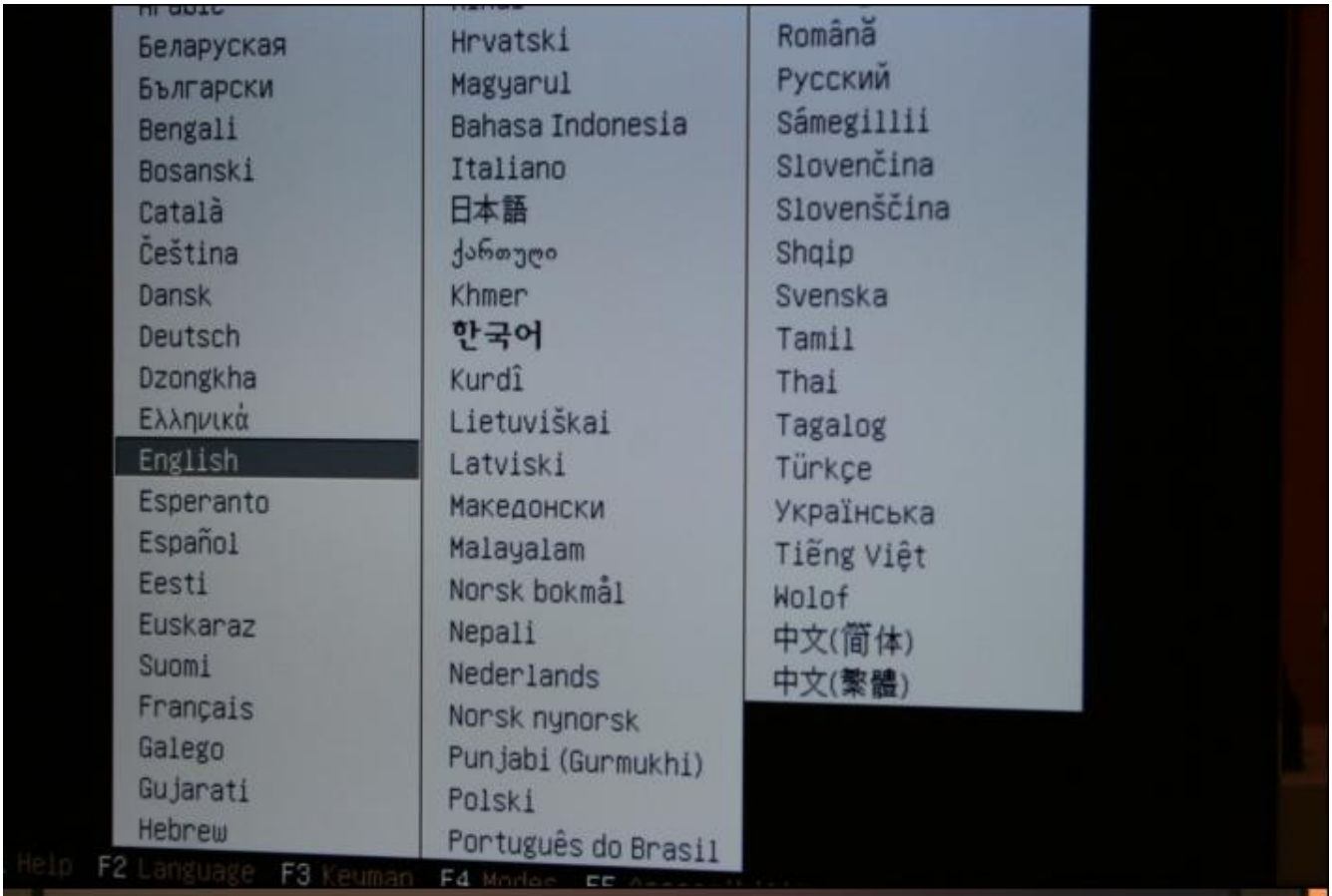
Put the Ubuntu Server 8.04 disk in the drive and reboot the head node.

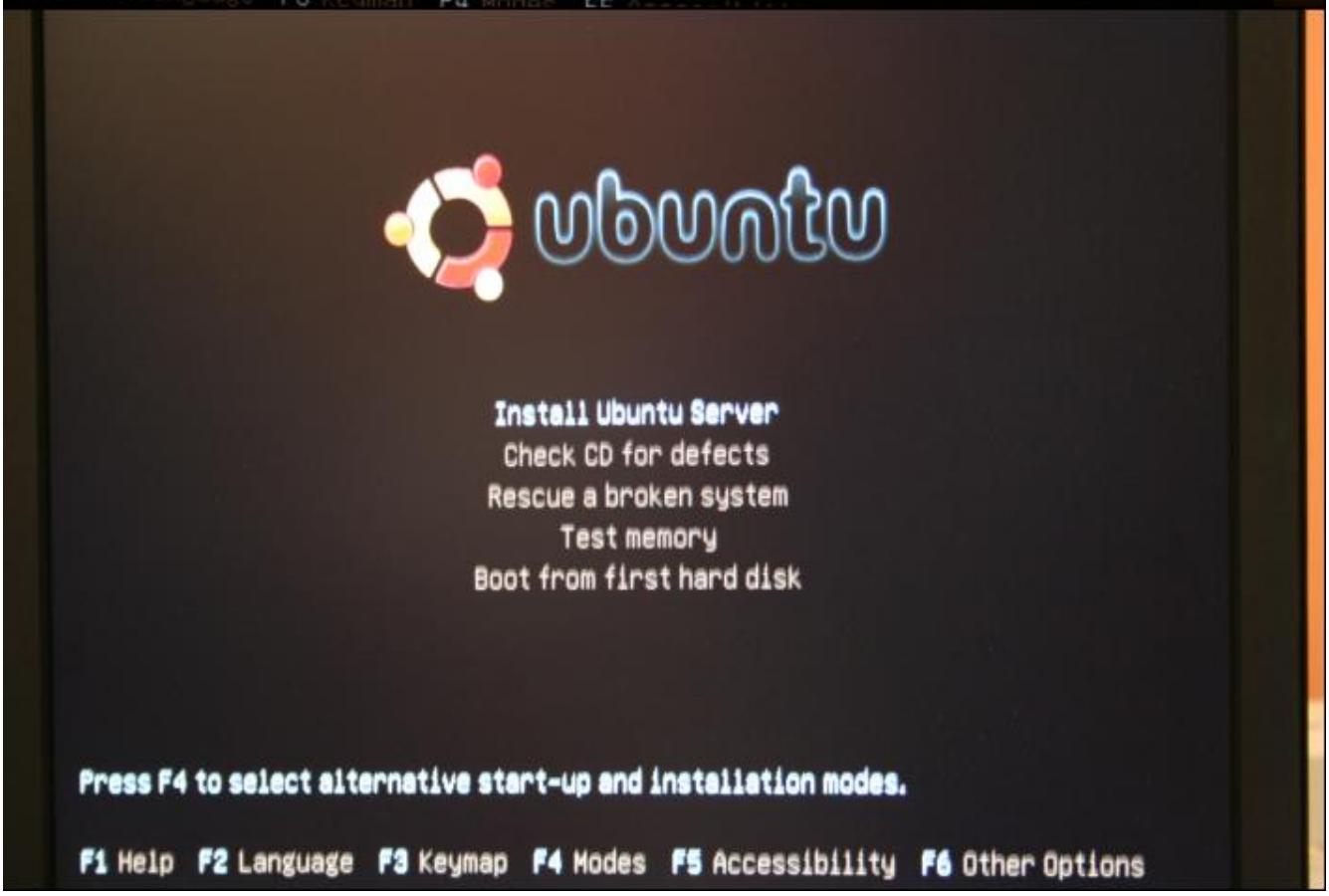
:pix:

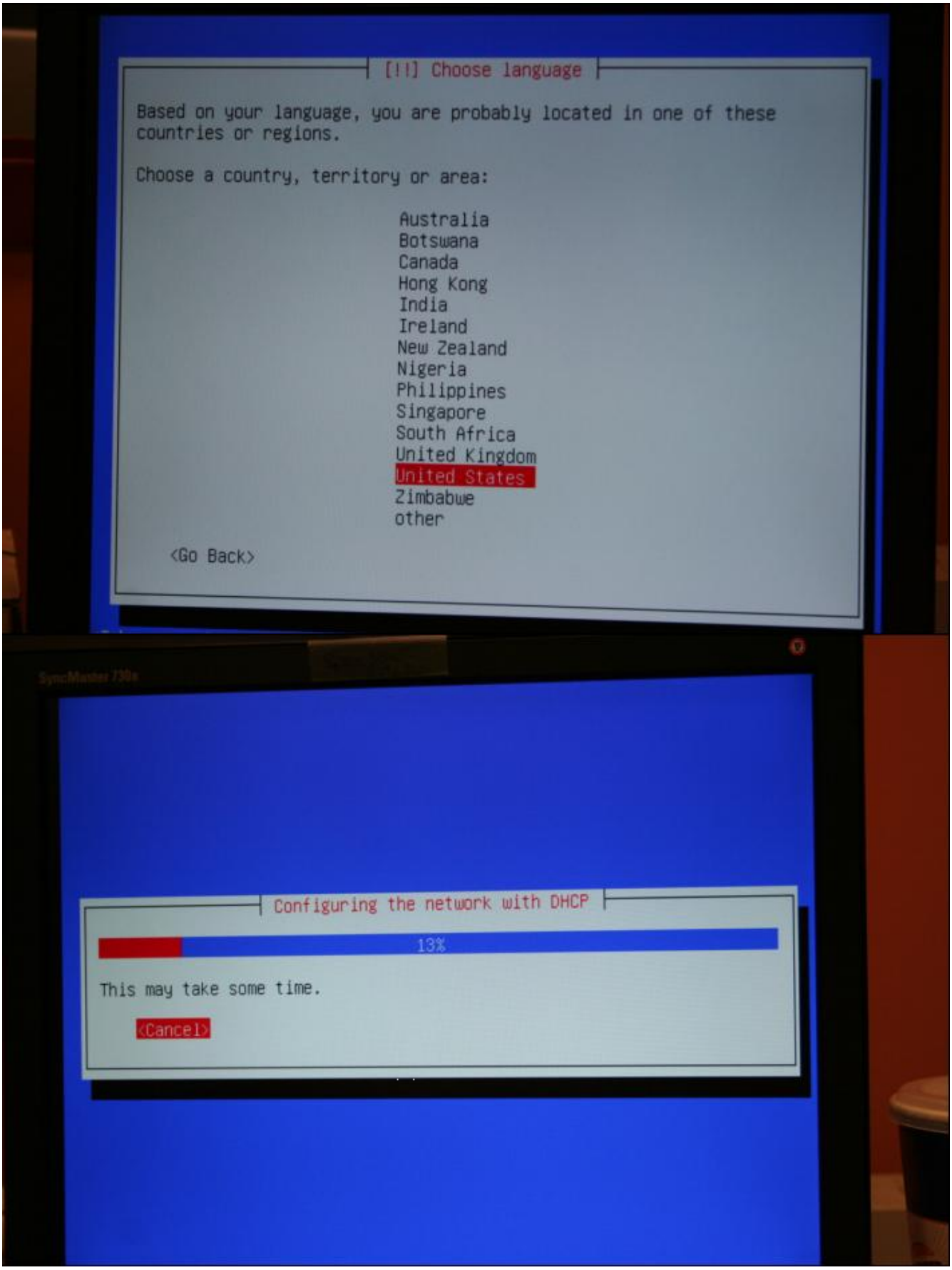


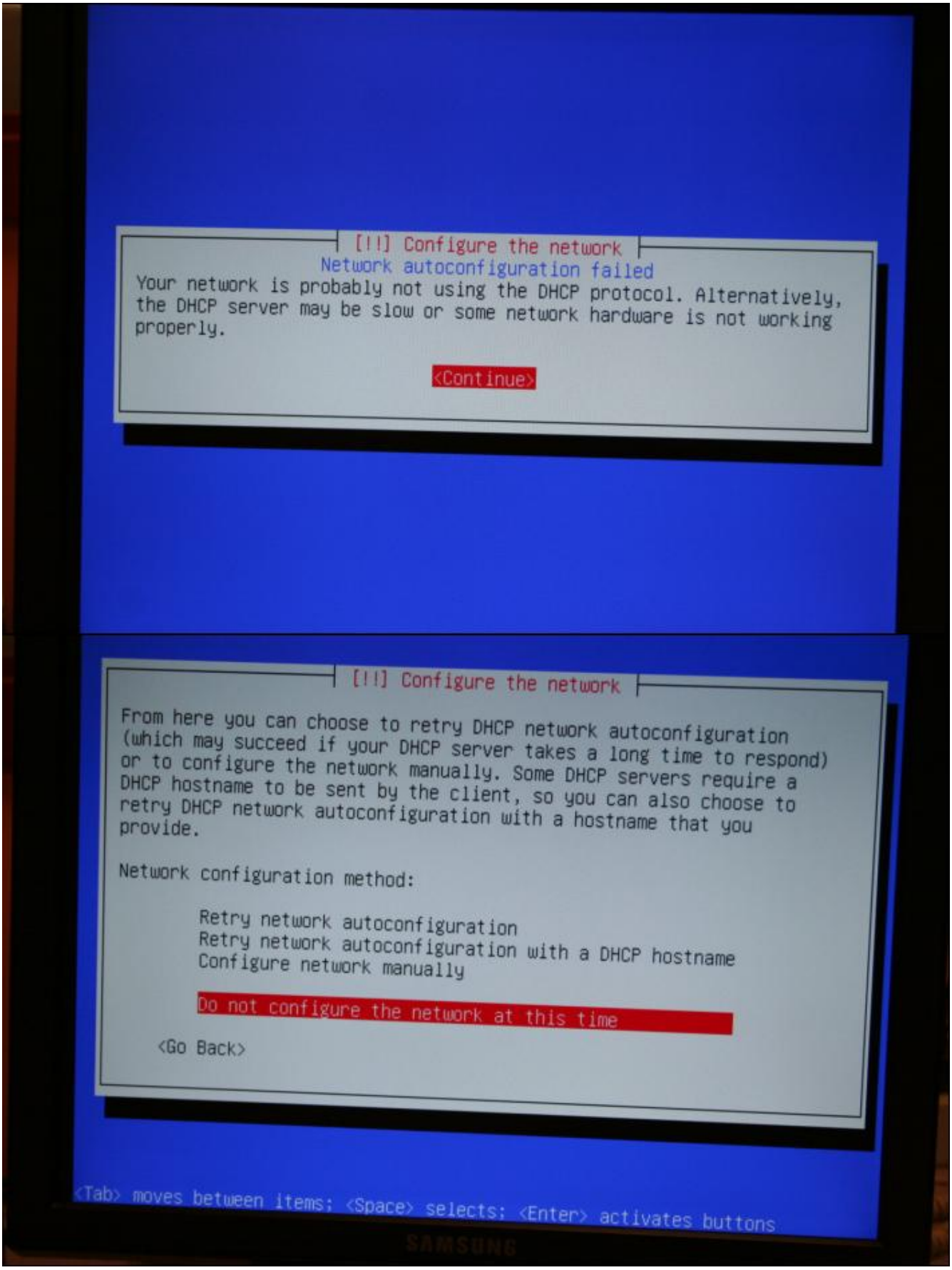
Memorex
CD-R

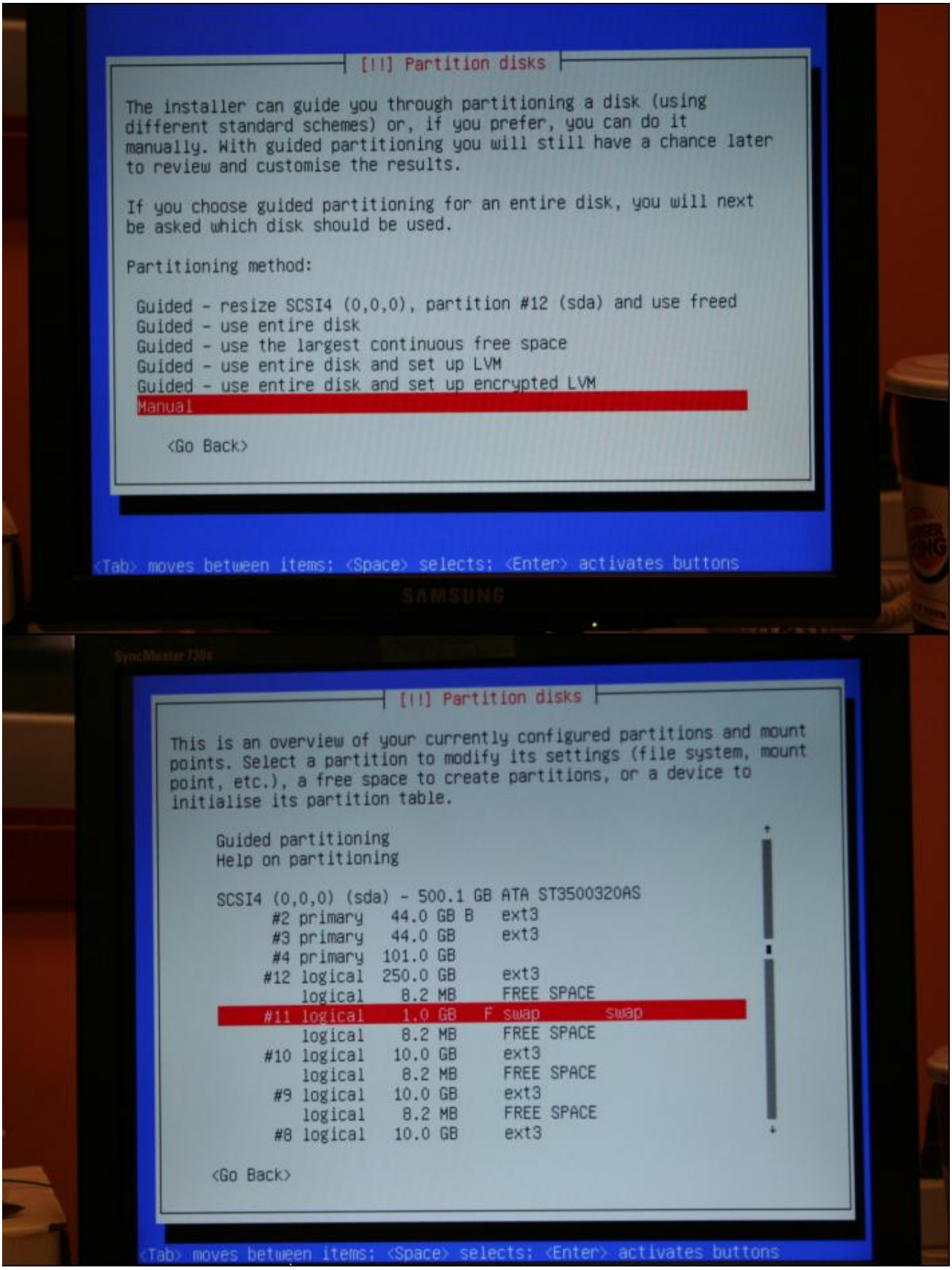

cdrom0 - File Browser
View Go Bookmarks Help
Up
Reload
Home
Computer
Search
Ubuntu-Server 8.04.1 i386
100%
View as Icons
dists
doc
install
isolinux
pics
cdromupgrade
Upgrade volume detected
A distribution volume with software packages has been detected.
Would you like to try to upgrade from it automatically?
Cancel
Start package manager
Run upgrade

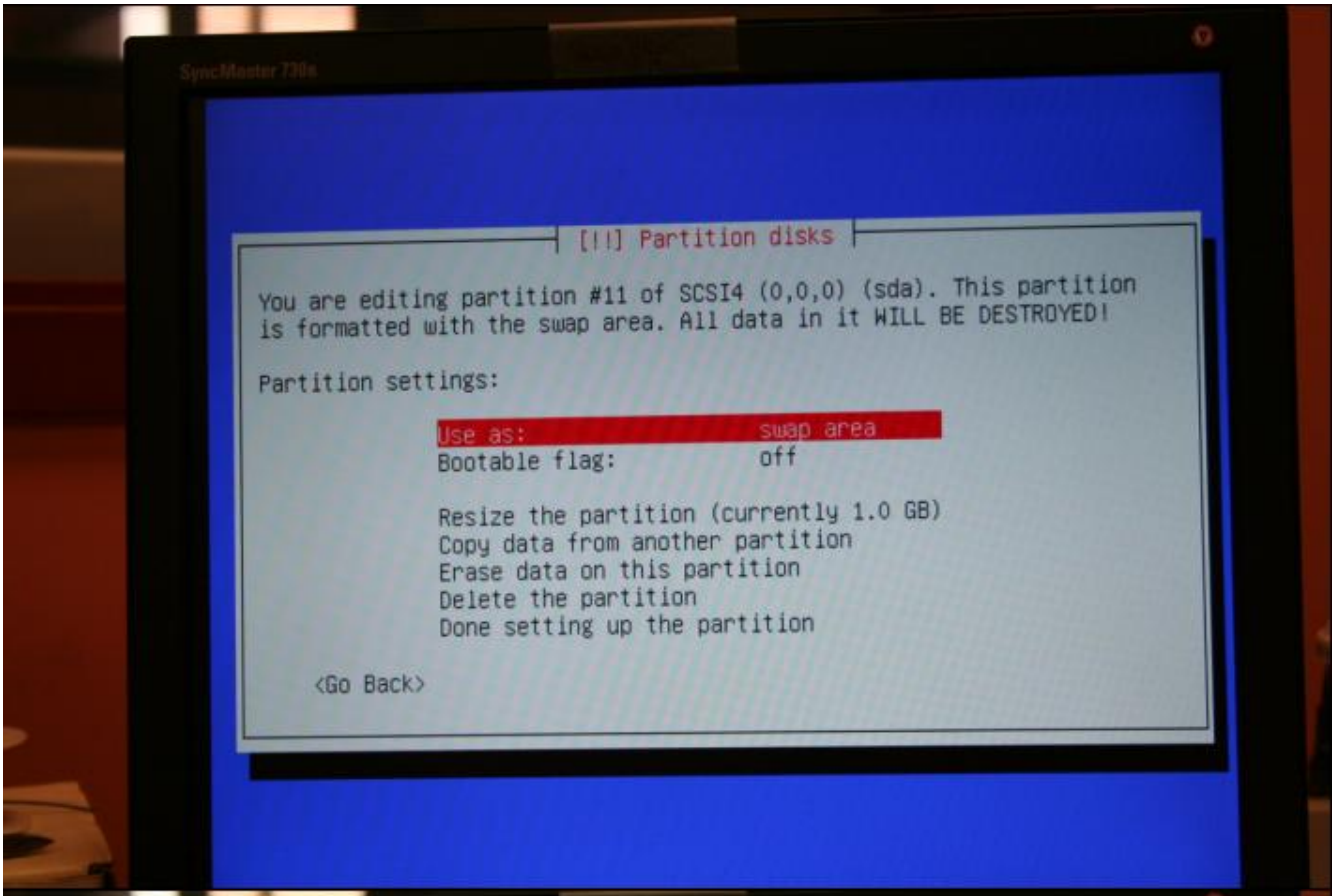

Беларуская
Български
Bengali
Bosanski
Català
Čeština
Dansk
Deutsch
Dzongkha
Ελληνικά
English
Esperanto
Español
Eesti
Euskaraz
Suomi
Français
Galego
Gujarati
Hebrew
Hrvatski
Magyarul
Bahasa Indonesia
Italiano
日本語
ქართული
Khmer
한국어
Kurdî
Lietuviškai
Latviski
Македонски
Malayalam
Norsk bokmål
Nepali
Nederlands
Norsk nynorsk
Punjabi (Gurmukhi)
Polski
Português do Brasil
Română
Русский
Sámegillii
Slovenčina
Slovenščina
Shqip
Svenska
Tamil
Thai
Tagalog
Türkçe
Українська
Tiếng Việt
Wolof
中文(简体)
中文(繁體)
Help F2 Language F3 Keyman F4 Modes

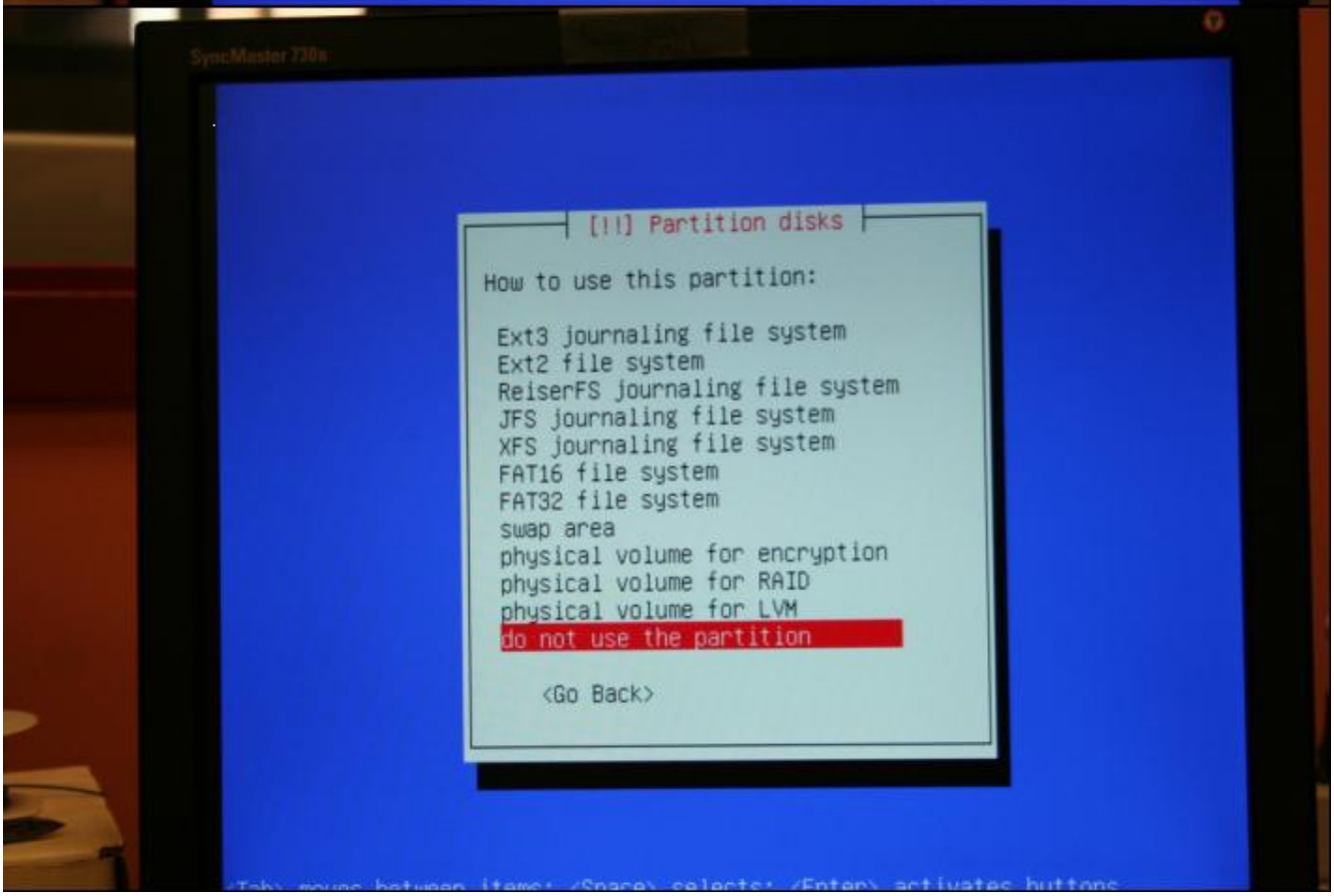

ubuntu
Install Ubuntu Server
Check CD for defects
Rescue a broken system
Test memory
Boot from first hard disk
Press F4 to select alternative start-up and installation modes.
F1 Help F2 Language F3 Keymap F4 Modes F5 Accessibility F6 Other Options


[!!] Choose language
Based on your language, you are probably located in one of these countries or regions.
Choose a country, territory or area:
Australia
Botswana
Canada
Hong Kong
India
Ireland
New Zealand
Nigeria
Philippines
Singapore
South Africa
United Kingdom
United States
Zimbabwe
other
<Go Back>
Configuring the network with DHCP
13%
This may take some time.
<Cancel>

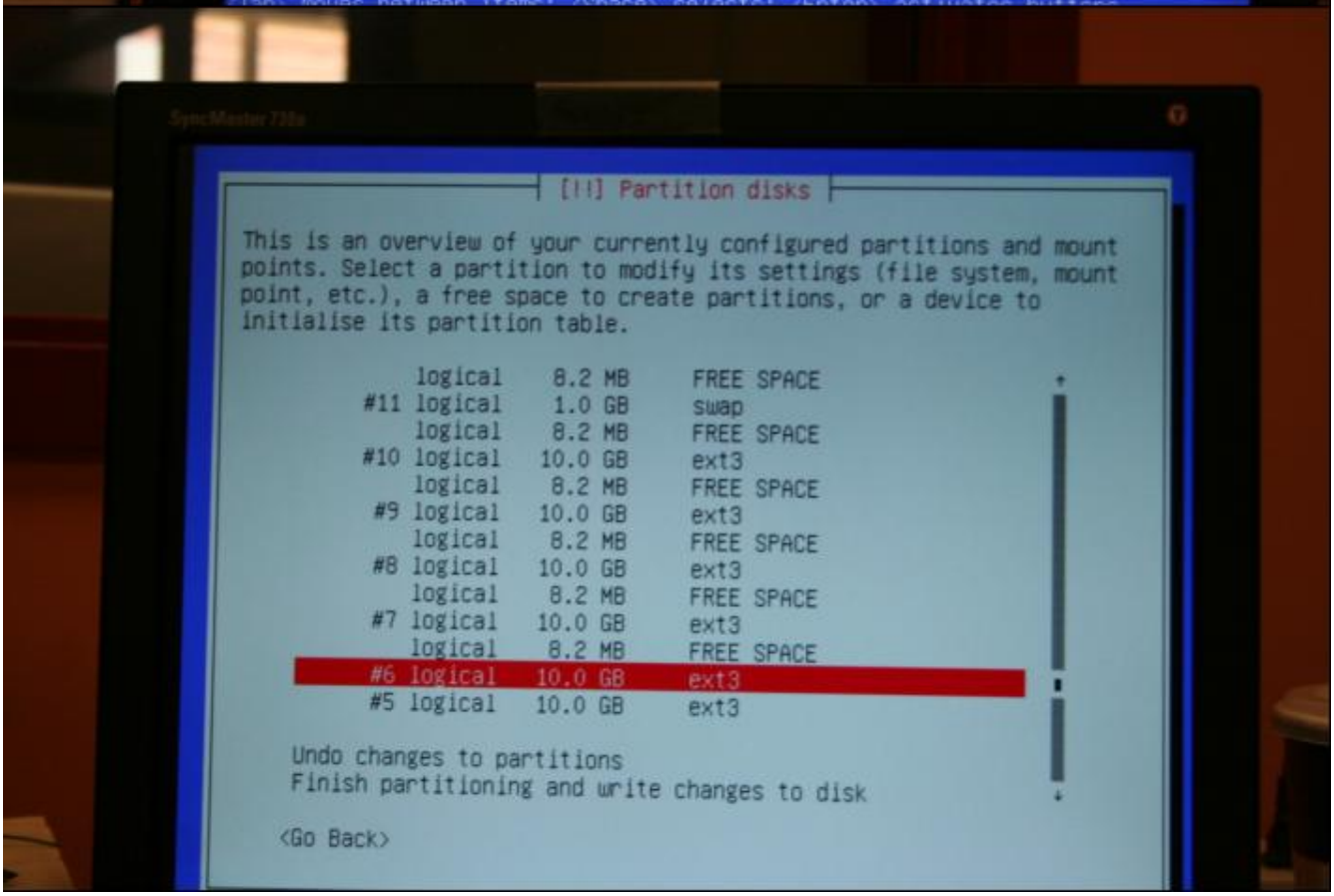

[!!] Configure the network
Network autoconfiguration failed
Your network is probably not using the DHCP protocol. Alternatively, the DHCP server may be slow or some network hardware is not working properly.
<Continue>
[!!] Configure the network
From here you can choose to retry DHCP network autoconfiguration (which may succeed if your DHCP server takes a long time to respond) or to configure the network manually. Some DHCP servers require a DHCP hostname to be sent by the client, so you can also choose to retry DHCP network autoconfiguration with a hostname that you provide.
Network configuration method:
Retry network autoconfiguration
Retry network autoconfiguration with a DHCP hostname
Configure network manually
Do not configure the network at this time
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons
SAMSUNG

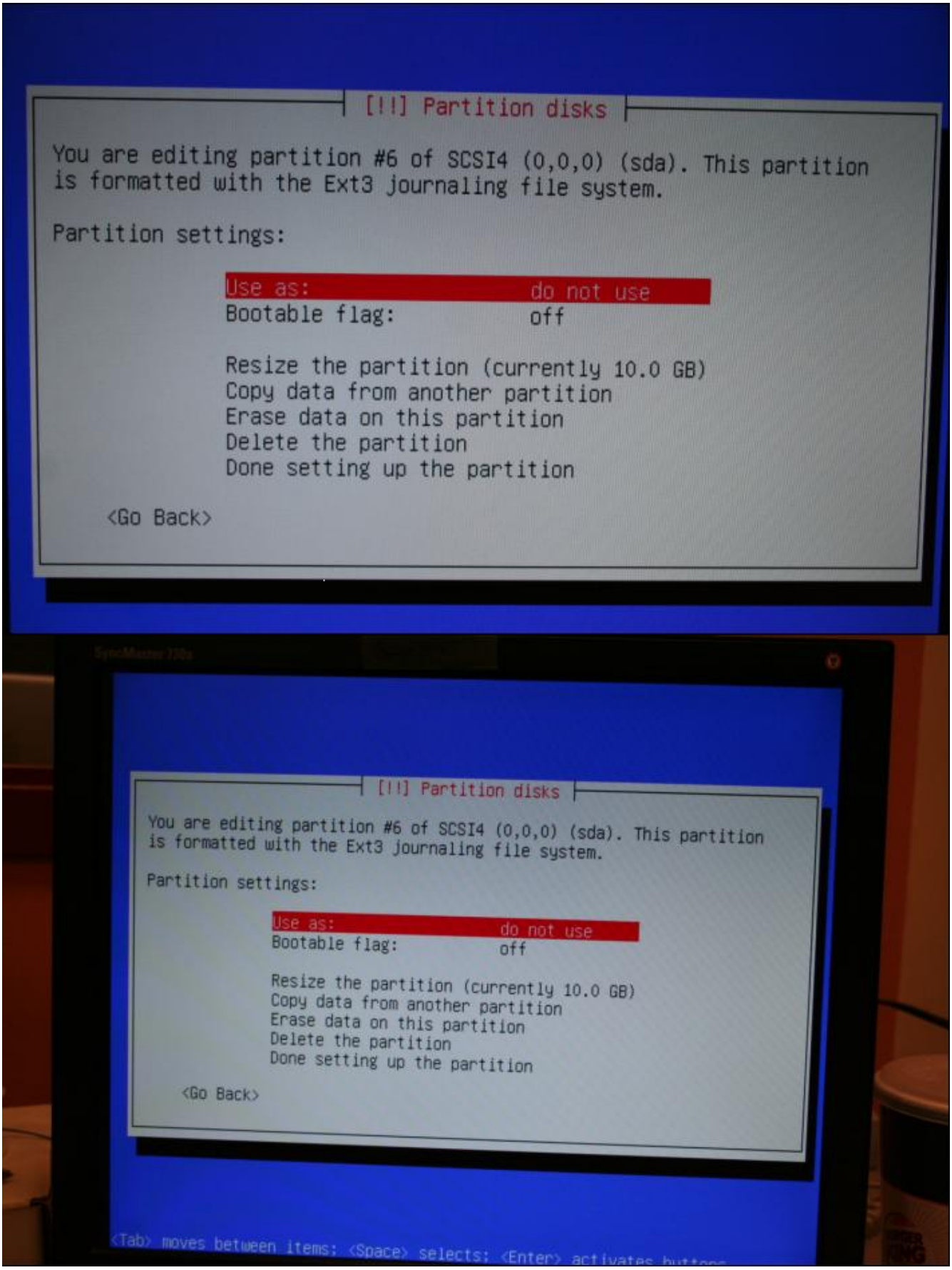

[!!] Partition disks
The installer can guide you through partitioning a disk (using different standard schemes) or, if you prefer, you can do it manually. With guided partitioning you will still have a chance later to review and customise the results.
If you choose guided partitioning for an entire disk, you will next be asked which disk should be used.
Partitioning method:
Guided - resize SCSI4 (0,0,0), partition #12 (sda) and use freed
Guided - use entire disk
Guided - use the largest continuous free space
Guided - use entire disk and set up LVM
Guided - use entire disk and set up encrypted LVM
Manual
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons
SAMSUNG
[!!] Partition disks
This is an overview of your currently configured partitions and mount points. Select a partition to modify its settings (file system, mount point, etc.), a free space to create partitions, or a device to initialise its partition table.
Guided partitioning
Help on partitioning
SCSI4 (0,0,0) (sda) - 500.1 GB ATA ST3500320AS
#2 primary 44.0 GB B ext3
#3 primary 44.0 GB ext3
#4 primary 101.0 GB
#12 logical 250.0 GB ext3
logical 8.2 MB FREE SPACE
#11 logical 1.0 GB F swap swap
logical 8.2 MB FREE SPACE
#10 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#9 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#8 logical 10.0 GB ext3
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons

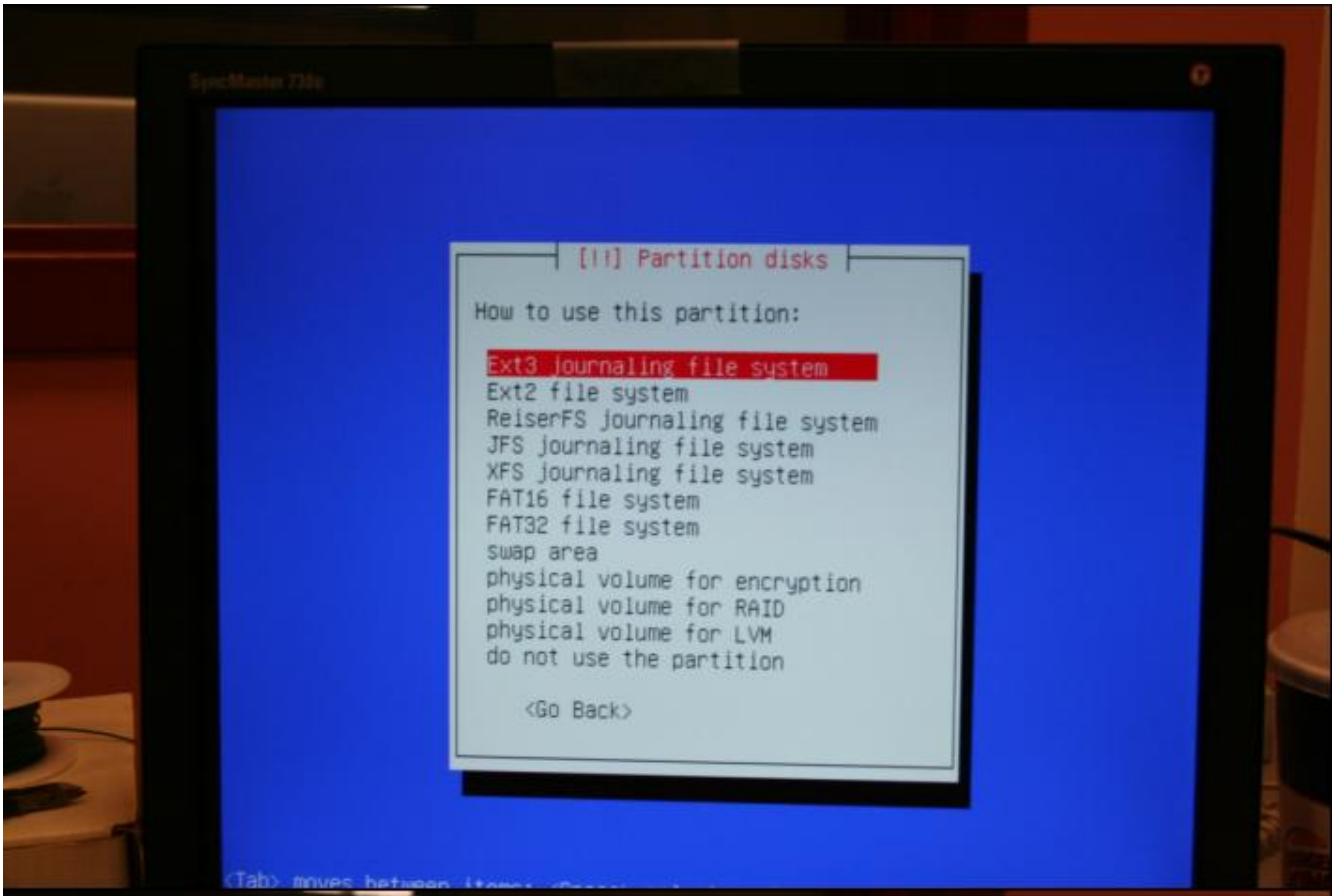

[!!] Partition disks
You are editing partition #11 of SCSI4 (0,0,0) (sda). This partition is formatted with the swap area. All data in it WILL BE DESTROYED!
Partition settings:
Use as: swap area
Bootable flag: off
Resize the partition (currently 1.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>

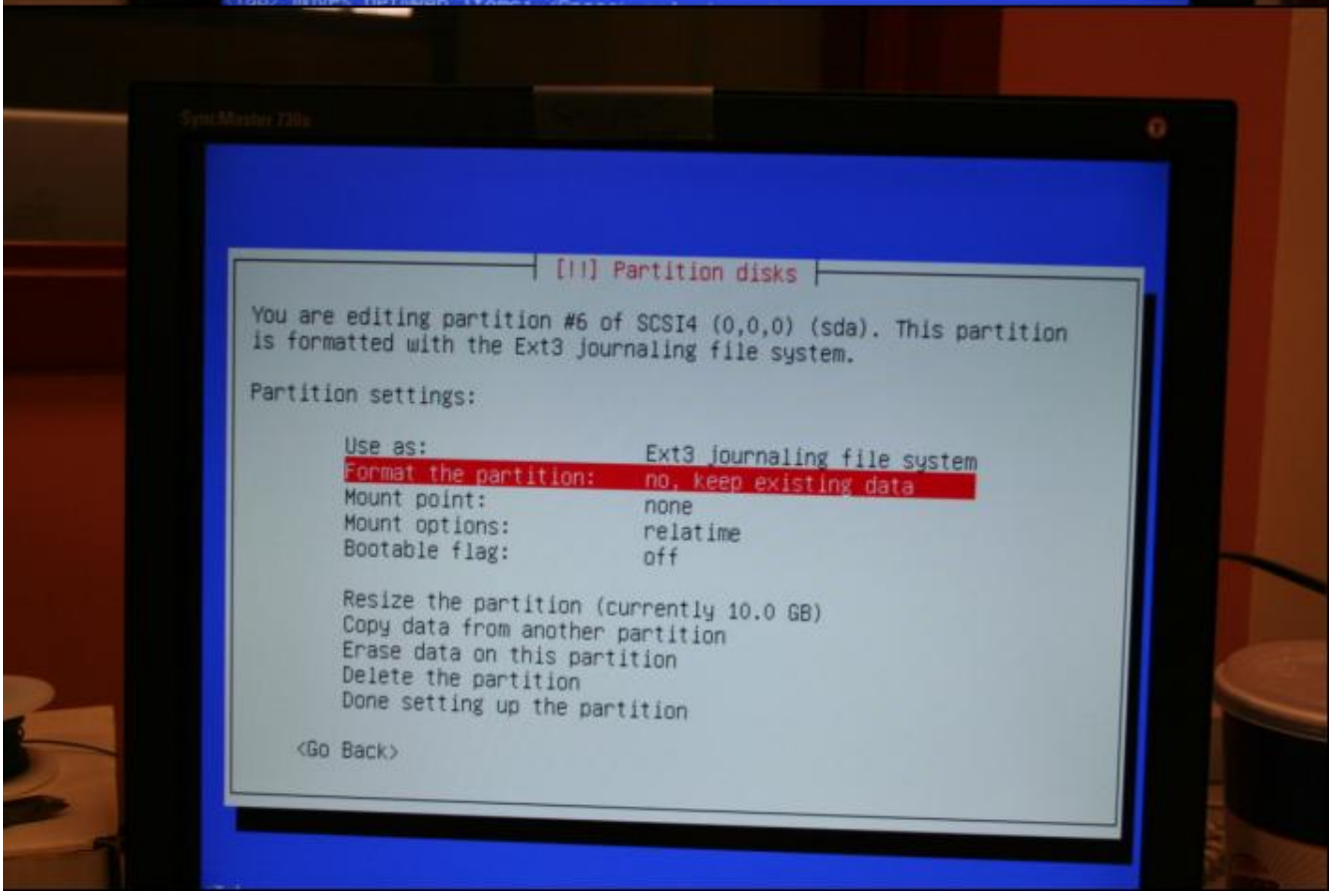

[!!] Partition disks
How to use this partition:
Ext3 journaling file system
Ext2 file system
ReiserFS journaling file system
JFS journaling file system
XFS journaling file system
FAT16 file system
FAT32 file system
swap area
physical volume for encryption
physical volume for RAID
physical volume for LVM
do not use the partition
<Go Back>

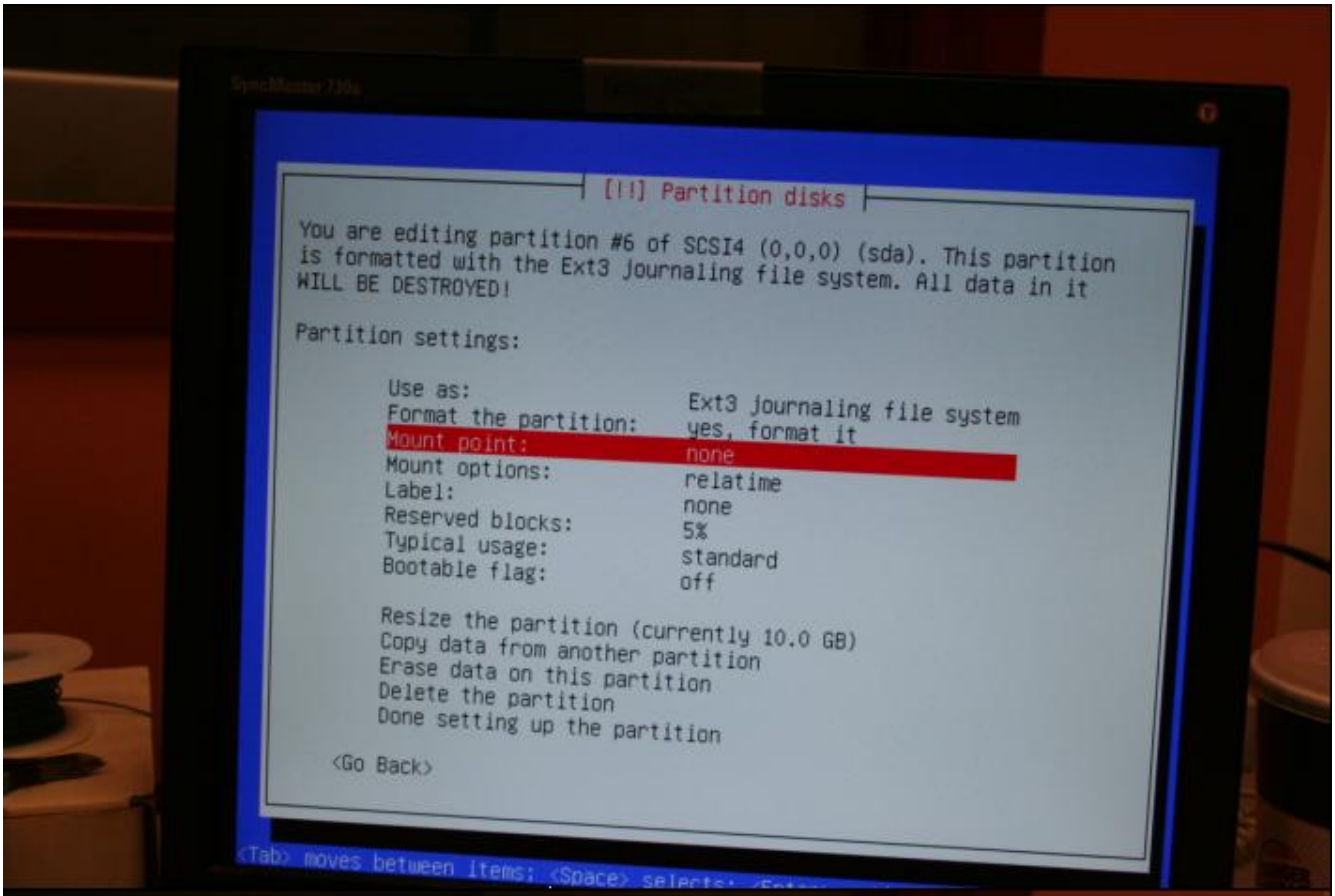
[!!] Partition disks
You are editing partition #11 of SCSI4 (0,0,0) (sda). This partition is formatted with the swap area.
Partition settings:
Use as: do not use
Bootable flag: off
Resize the partition (currently 1.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>

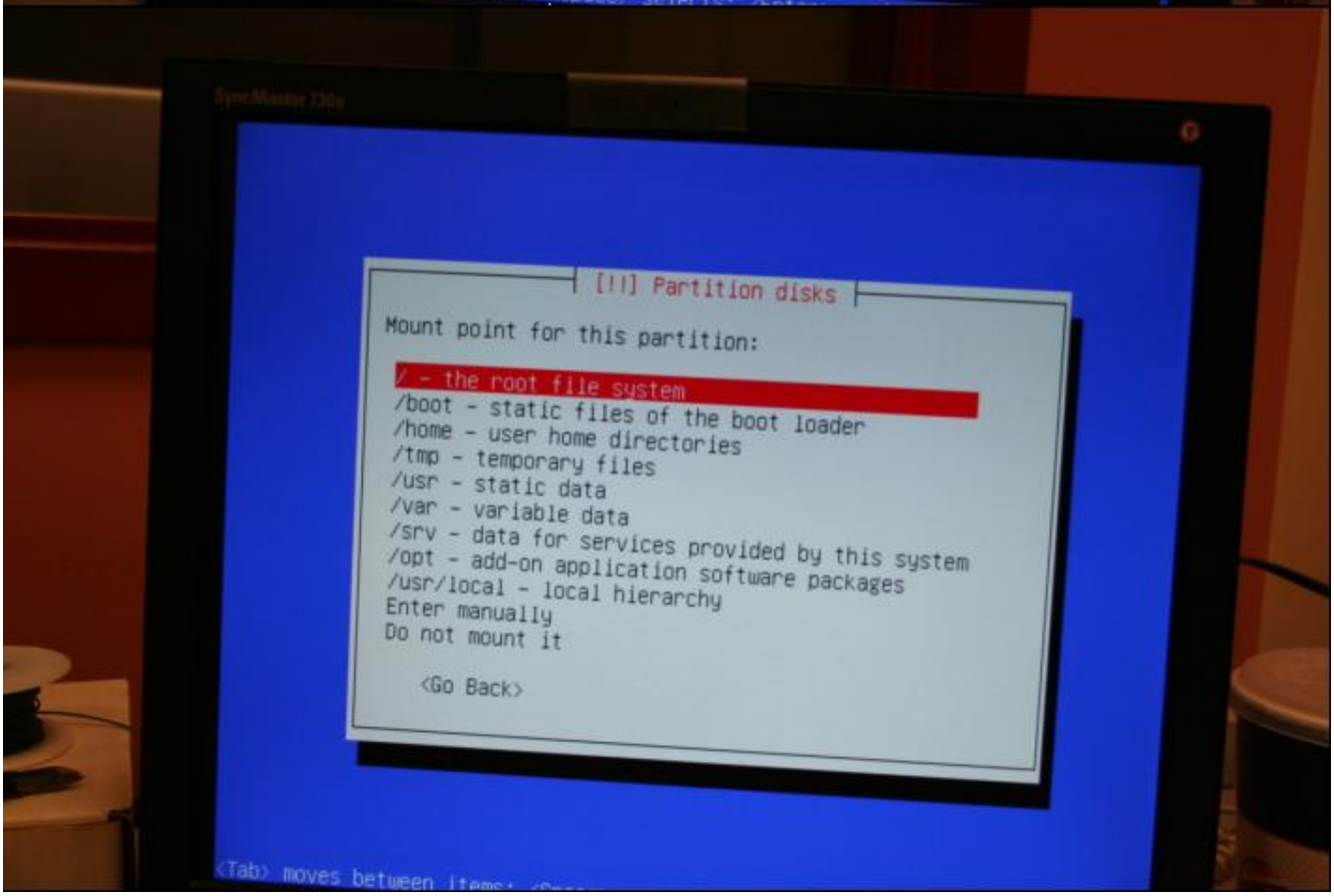

[!!] Partition disks
This is an overview of your currently configured partitions and mount points. Select a partition to modify its settings (file system, mount point, etc.), a free space to create partitions, or a device to initialise its partition table.
logical 8.2 MB FREE SPACE
#11 logical 1.0 GB swap
logical 8.2 MB FREE SPACE
#10 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#9 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#8 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#7 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#6 logical 10.0 GB ext3
#5 logical 10.0 GB ext3
Undo changes to partitions
Finish partitioning and write changes to disk
<Go Back>

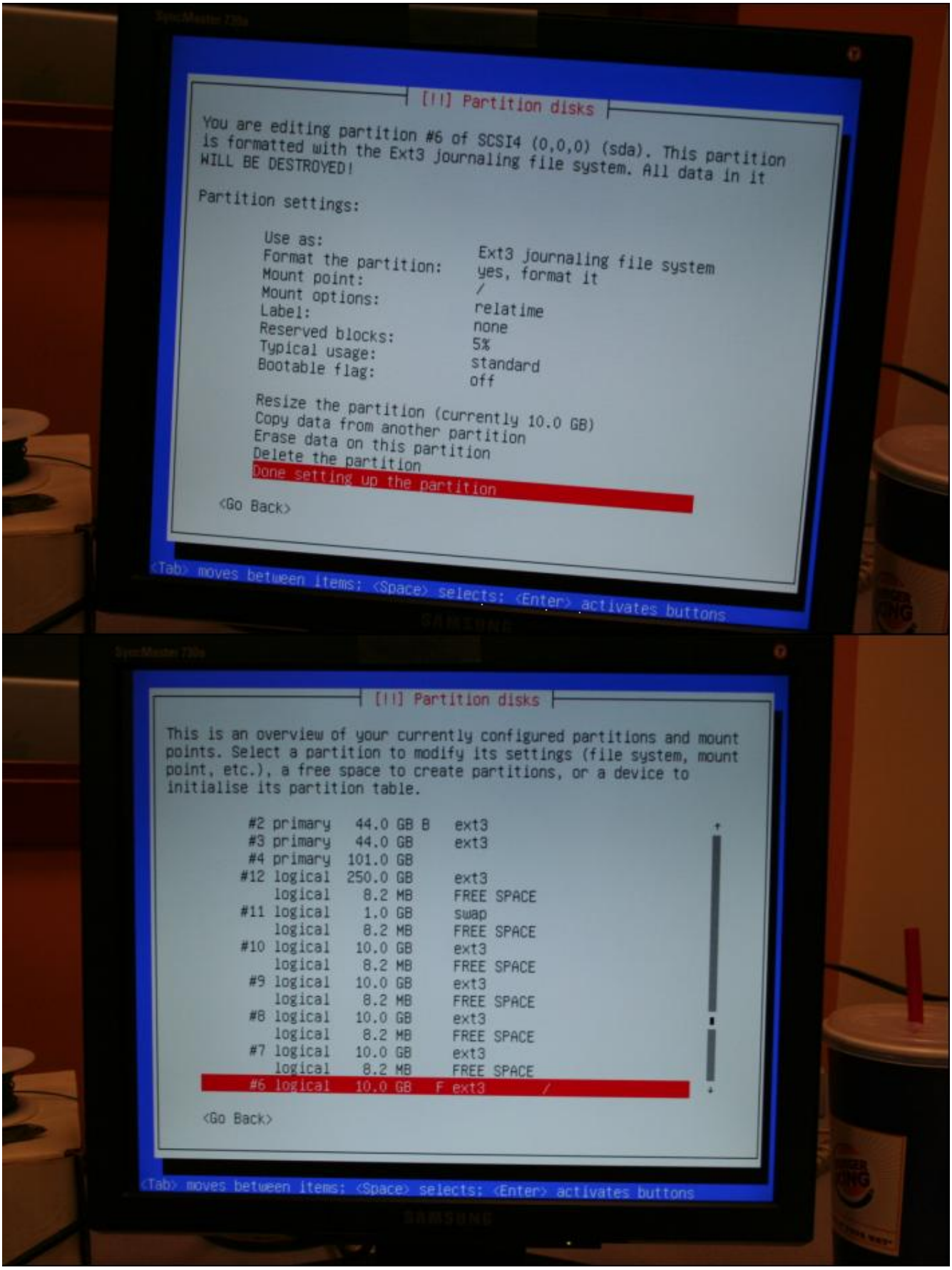

[!!] Partition disks
You are editing partition #6 of SCSI4 (0,0,0) (sda). This partition is formatted with the Ext3 journaling file system.
Partition settings:
Use as: do not use
Bootable flag: off
Resize the partition (currently 10.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>
[!!] Partition disks
You are editing partition #6 of SCSI4 (0,0,0) (sda). This partition is formatted with the Ext3 journaling file system.
Partition settings:
Use as: do not use
Bootable flag: off
Resize the partition (currently 10.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates

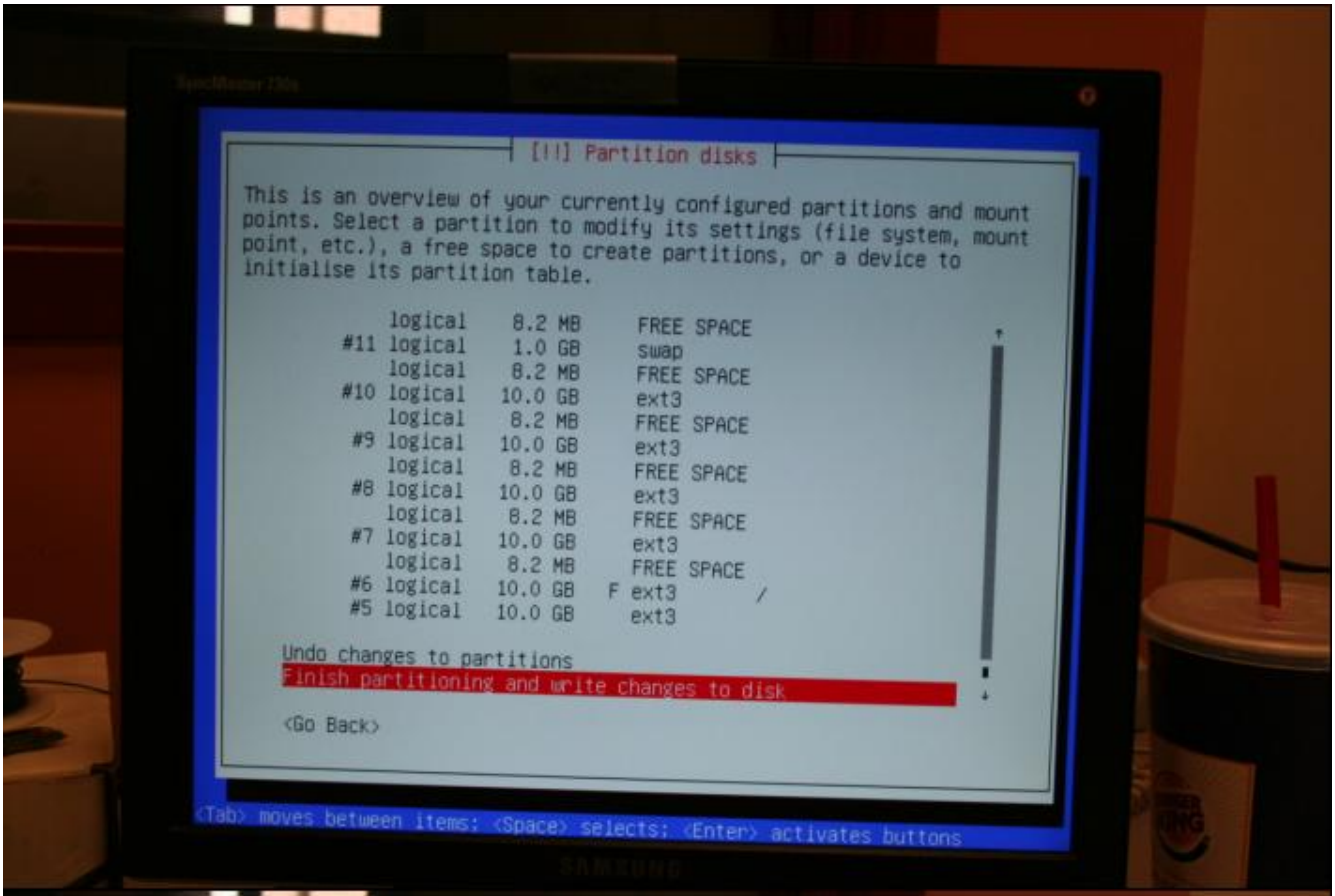

[!!] Partition disks
How to use this partition:
Ext3 journaling file system
Ext2 file system
ReiserFS journaling file system
JFS journaling file system
XFS journaling file system
FAT16 file system
FAT32 file system
swap area
physical volume for encryption
physical volume for RAID
physical volume for LVM
do not use the partition
<Go Back>

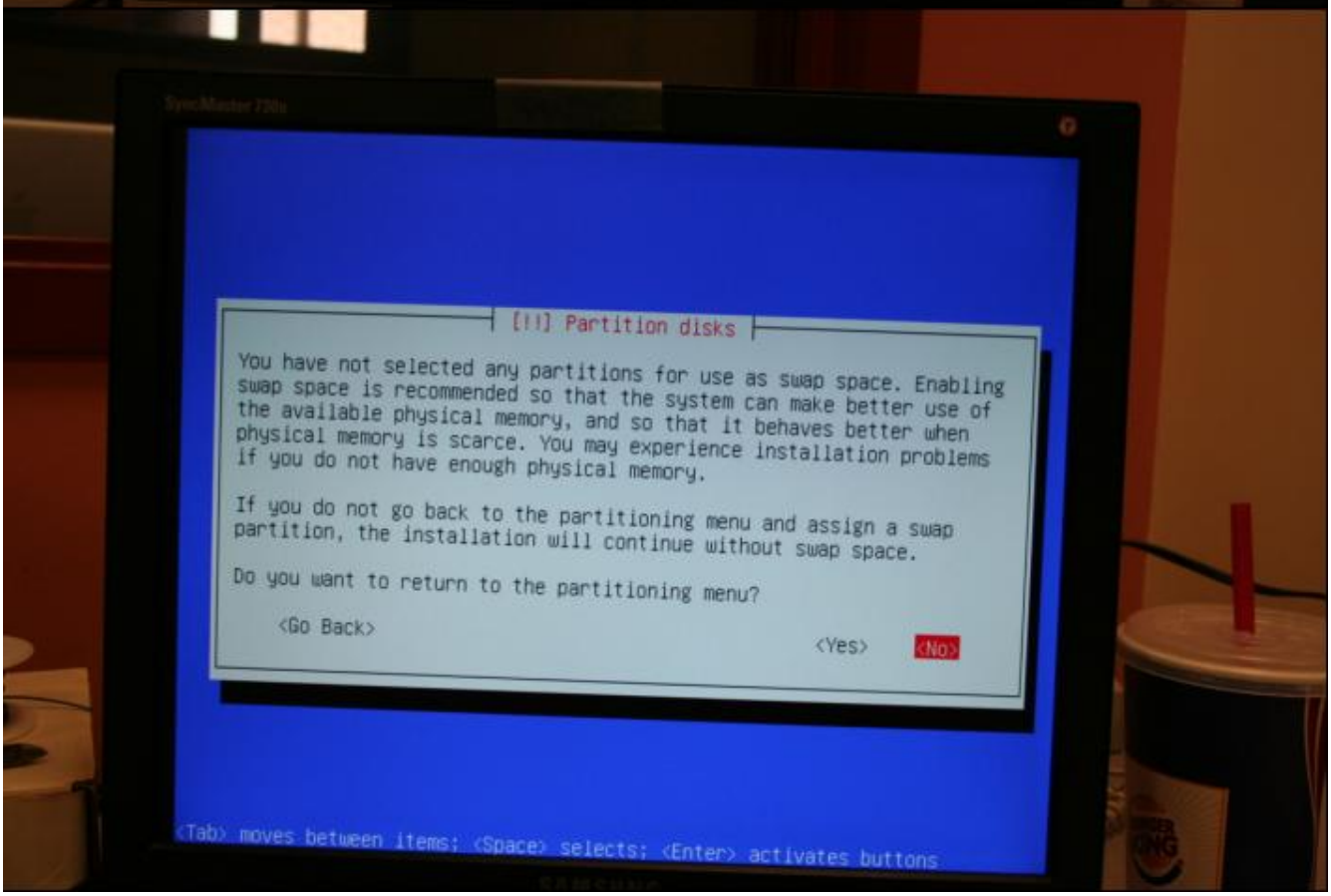

[!!] Partition disks
You are editing partition #6 of SCSI4 (0,0,0) (sda). This partition is formatted with the Ext3 journaling file system.
Partition settings:
Use as: Ext3 journaling file system
Format the partition: no, keep existing data
Mount point: none
Mount options: relatime
Bootable flag: off
Resize the partition (currently 10.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>


[!!] Partition disks
You are editing partition #6 of SCSI4 (0,0,0) (sda). This partition is formatted with the Ext3 journaling file system.
WILL BE DESTROYED!
Partition settings:
Use as: Ext3 journaling file system
Format the partition: yes, format it
Mount point: none
Mount options: relatime
Label: none
Reserved blocks: 5%
Typical usage: standard
Bootable flag: off
Resize the partition (currently 10.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>
<Tab> moves between items; <Space> selects;

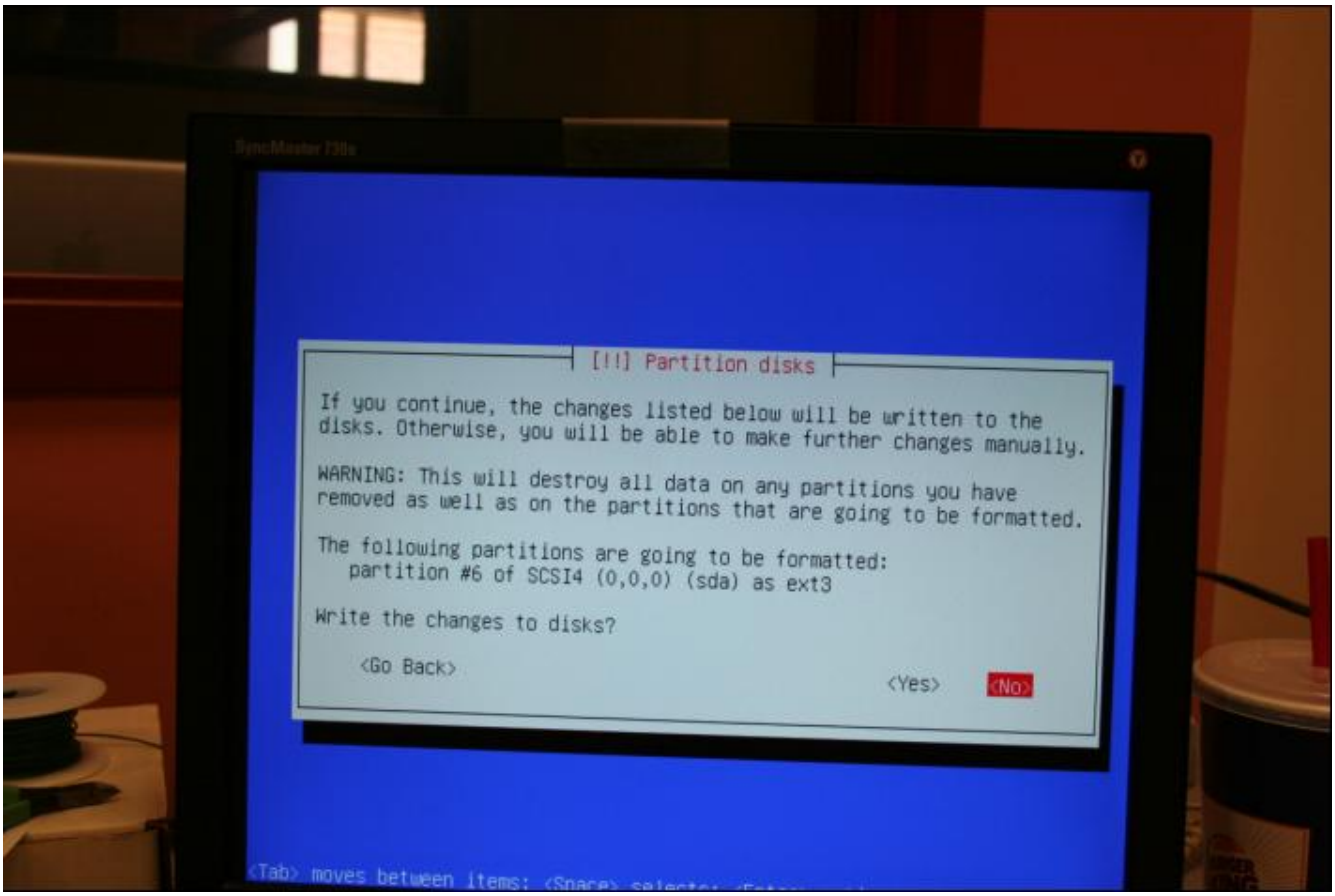

[!!] Partition disks
Mount point for this partition:
/ - the root file system
/boot - static files of the boot loader
/home - user home directories
/tmp - temporary files
/usr - static data
/var - variable data
/srv - data for services provided by this system
/opt - add-on application software packages
/usr/local - local hierarchy
Enter manually
Do not mount it
<Go Back>
<Tab> moves between items;

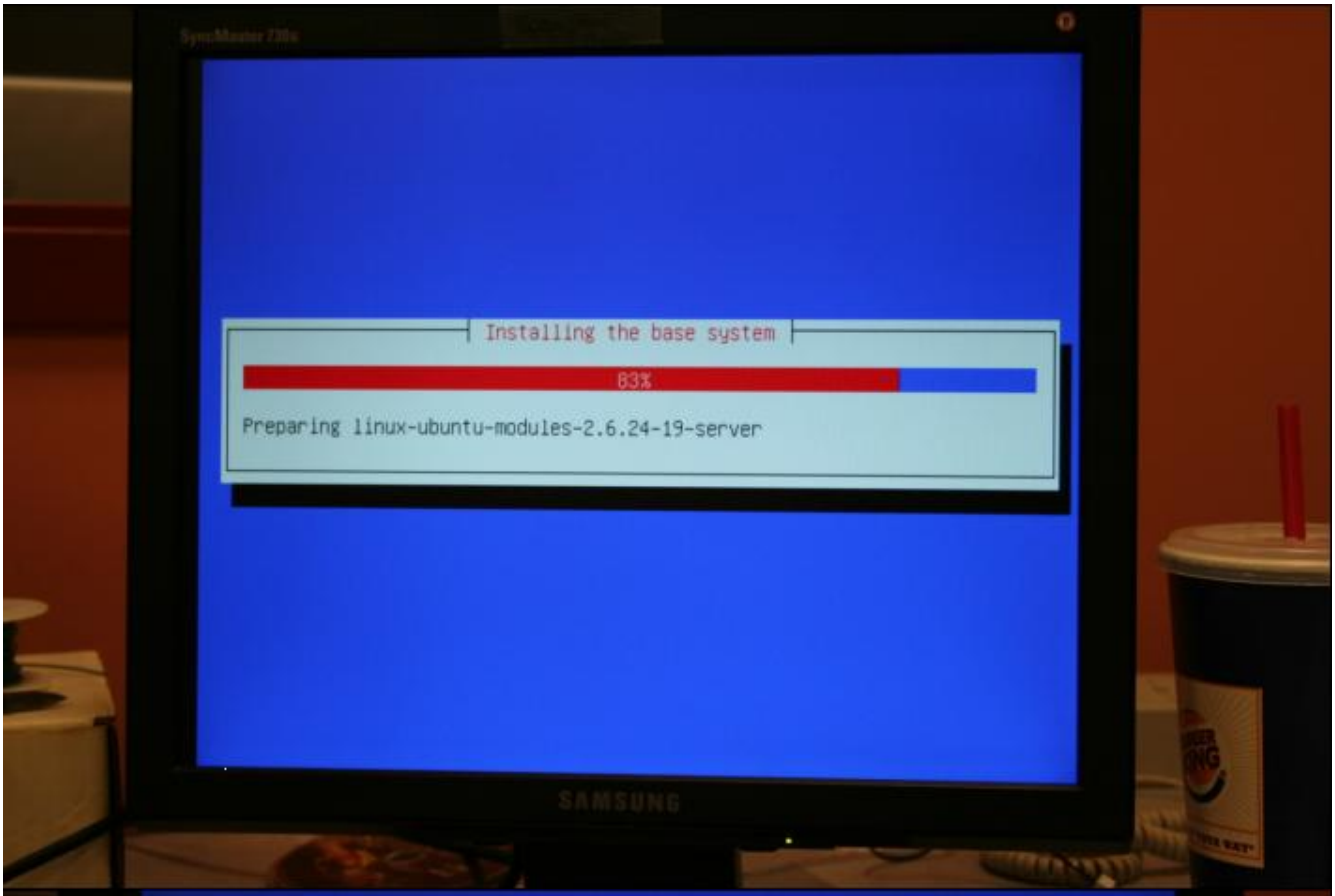

[!!] Partition disks
You are editing partition #6 of SCSI4 (0,0,0) (sda). This partition is formatted with the Ext3 journaling file system. All data in it WILL BE DESTROYED!
Partition settings:
Use as: Ext3 journaling file system
Format the partition: yes, format it
Mount point: /
Mount options: relatime
Label: none
Reserved blocks: 5%
Typical usage: standard
Bootable flag: off
Resize the partition (currently 10.0 GB)
Copy data from another partition
Erase data on this partition
Delete the partition
Done setting up the partition
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons
[!!] Partition disks
This is an overview of your currently configured partitions and mount points. Select a partition to modify its settings (file system, mount point, etc.), a free space to create partitions, or a device to initialise its partition table.
#2 primary 44.0 GB B ext3
#3 primary 44.0 GB ext3
#4 primary 101.0 GB
#12 logical 250.0 GB ext3
logical 8.2 MB FREE SPACE
#11 logical 1.0 GB swap
logical 8.2 MB FREE SPACE
#10 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#9 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#8 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#7 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#6 logical 10.0 GB F ext3 /
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons

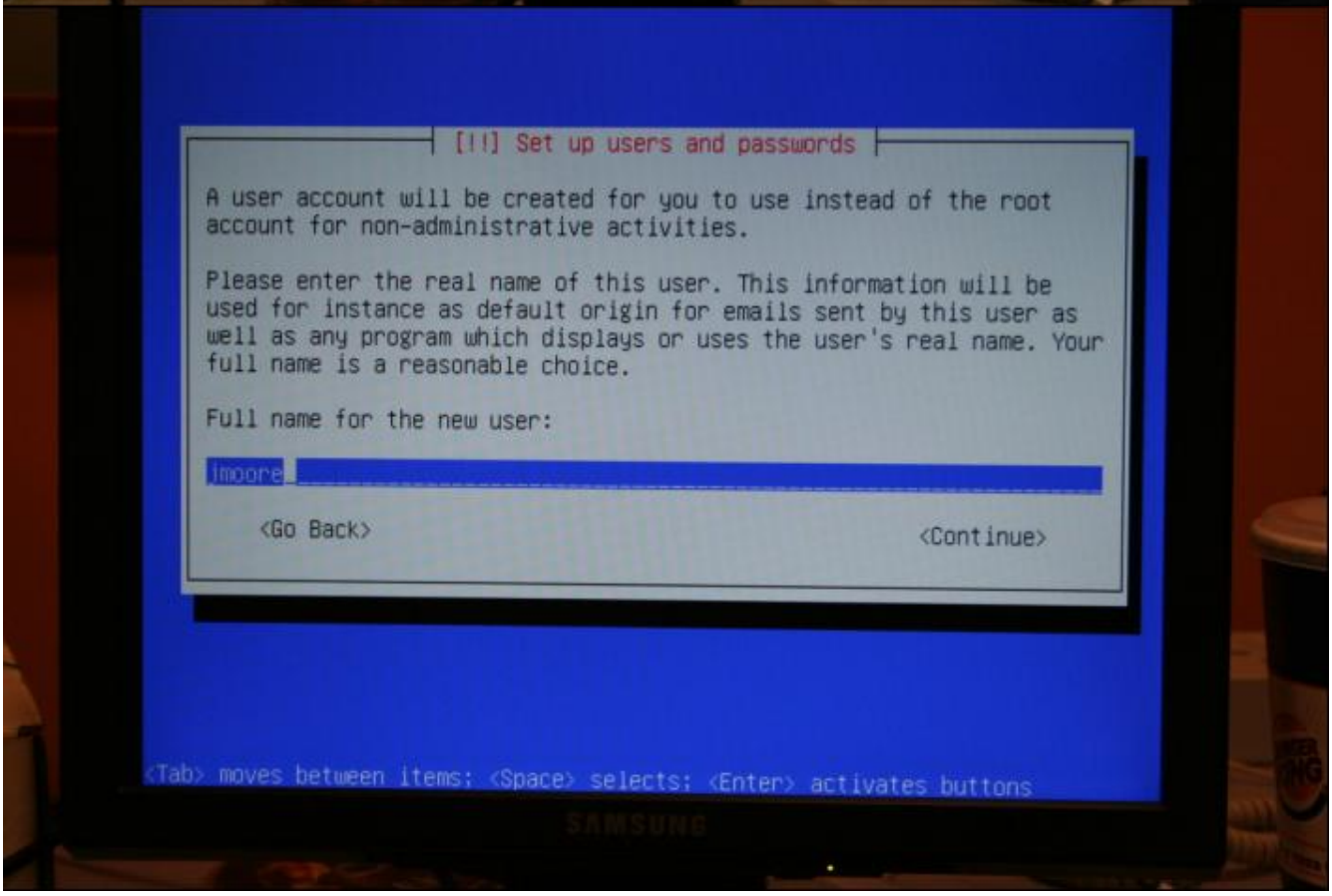

[!!] Partition disks
This is an overview of your currently configured partitions and mount points. Select a partition to modify its settings (file system, mount point, etc.), a free space to create partitions, or a device to initialise its partition table.
logical 8.2 MB FREE SPACE
#11 logical 1.0 GB swap
logical 8.2 MB FREE SPACE
#10 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#9 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#8 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#7 logical 10.0 GB ext3
logical 8.2 MB FREE SPACE
#6 logical 10.0 GB F ext3 /
#5 logical 10.0 GB ext3
Undo changes to partitions
Finish partitioning and write changes to disk
<Go Back>
<Tab> moves between items; <Space> selects; <Enter> activates buttons

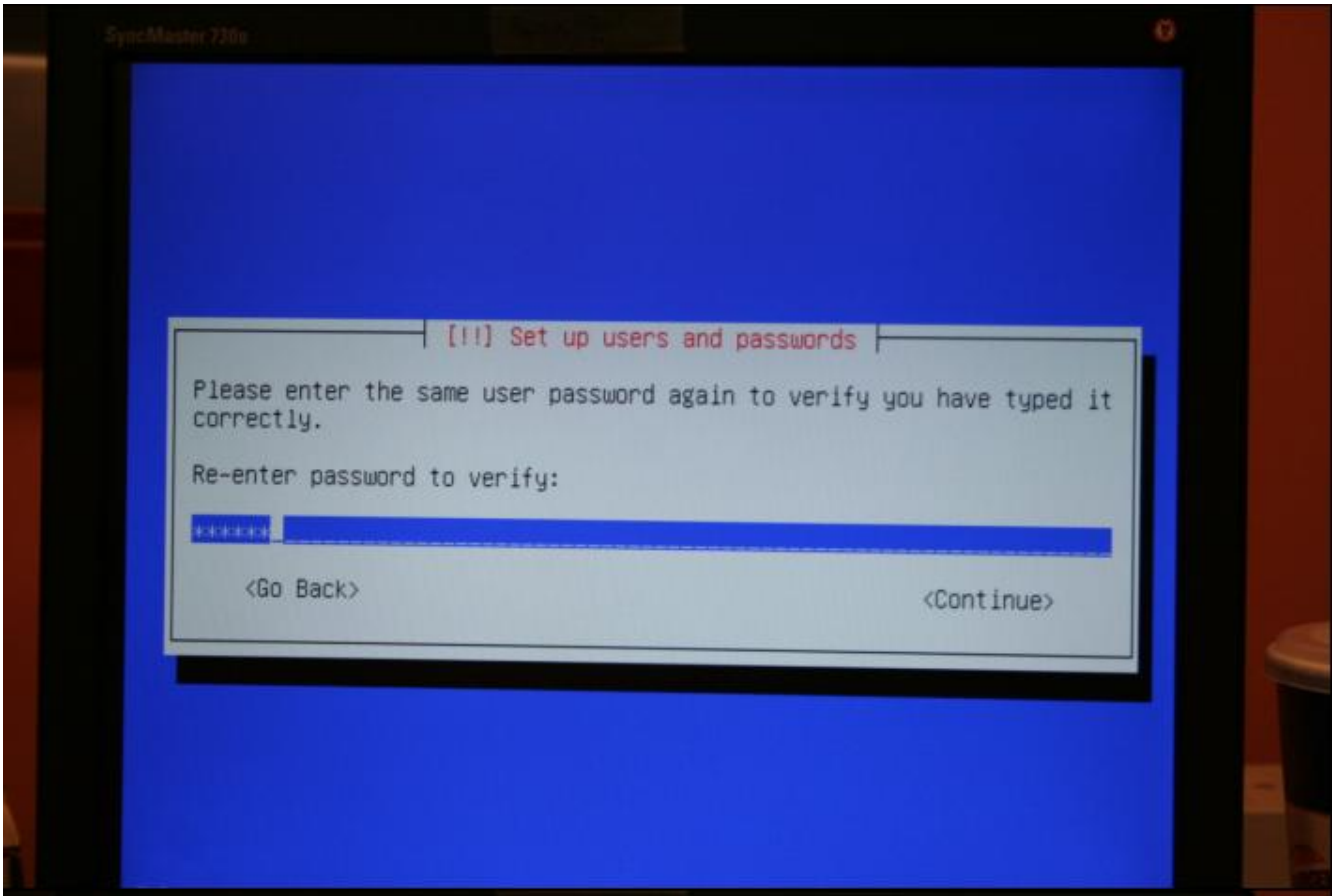

[!!] Partition disks
You have not selected any partitions for use as swap space. Enabling swap space is recommended so that the system can make better use of the available physical memory, and so that it behaves better when physical memory is scarce. You may experience installation problems if you do not have enough physical memory.
If you do not go back to the partitioning menu and assign a swap partition, the installation will continue without swap space.
Do you want to return to the partitioning menu?
<Go Back>
<Yes>
<No>
<Tab> moves between items; <Space> selects; <Enter> activates buttons

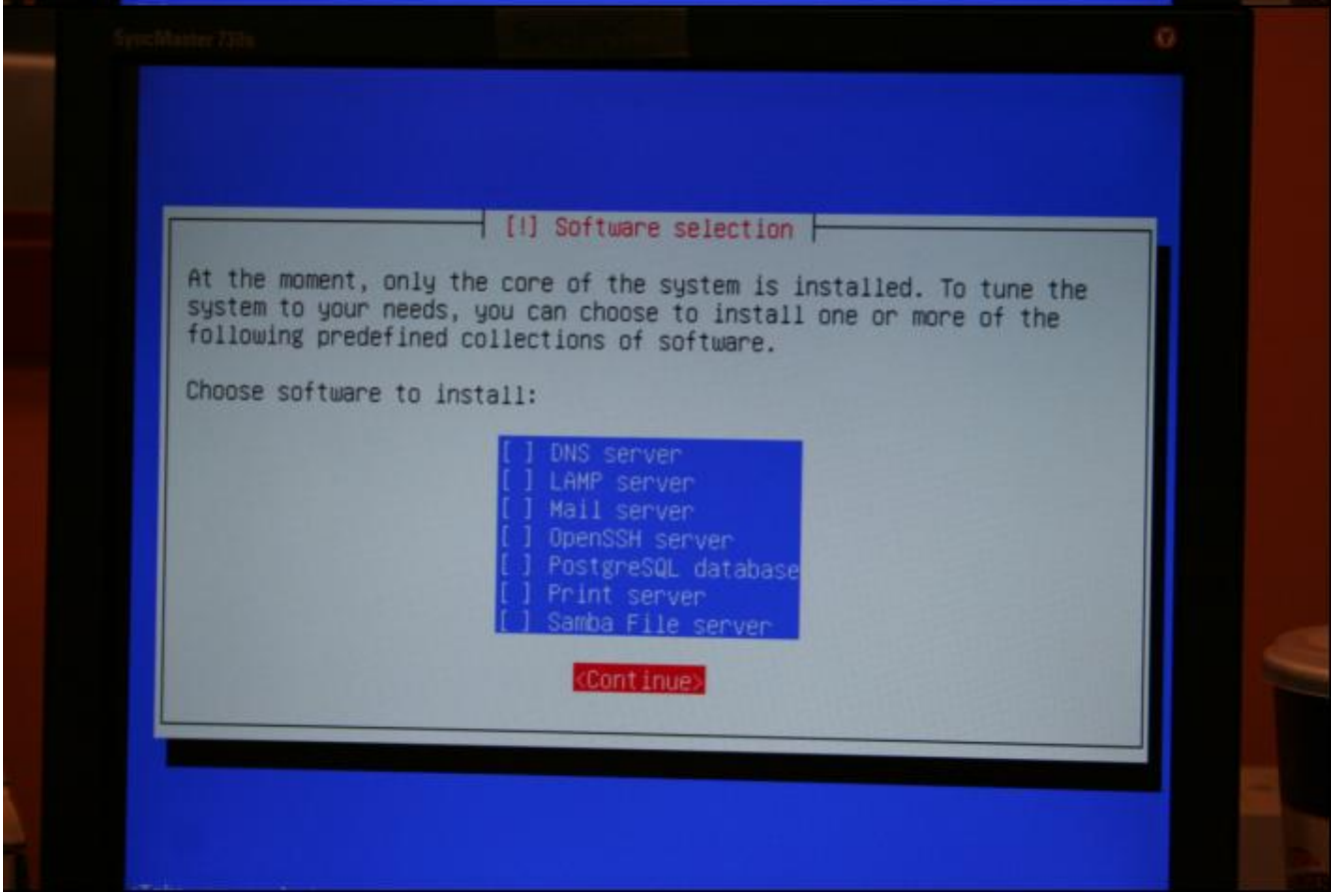
The next image should have the “YES” highlighted, but I snapped the picture and hit enter before I realized it.



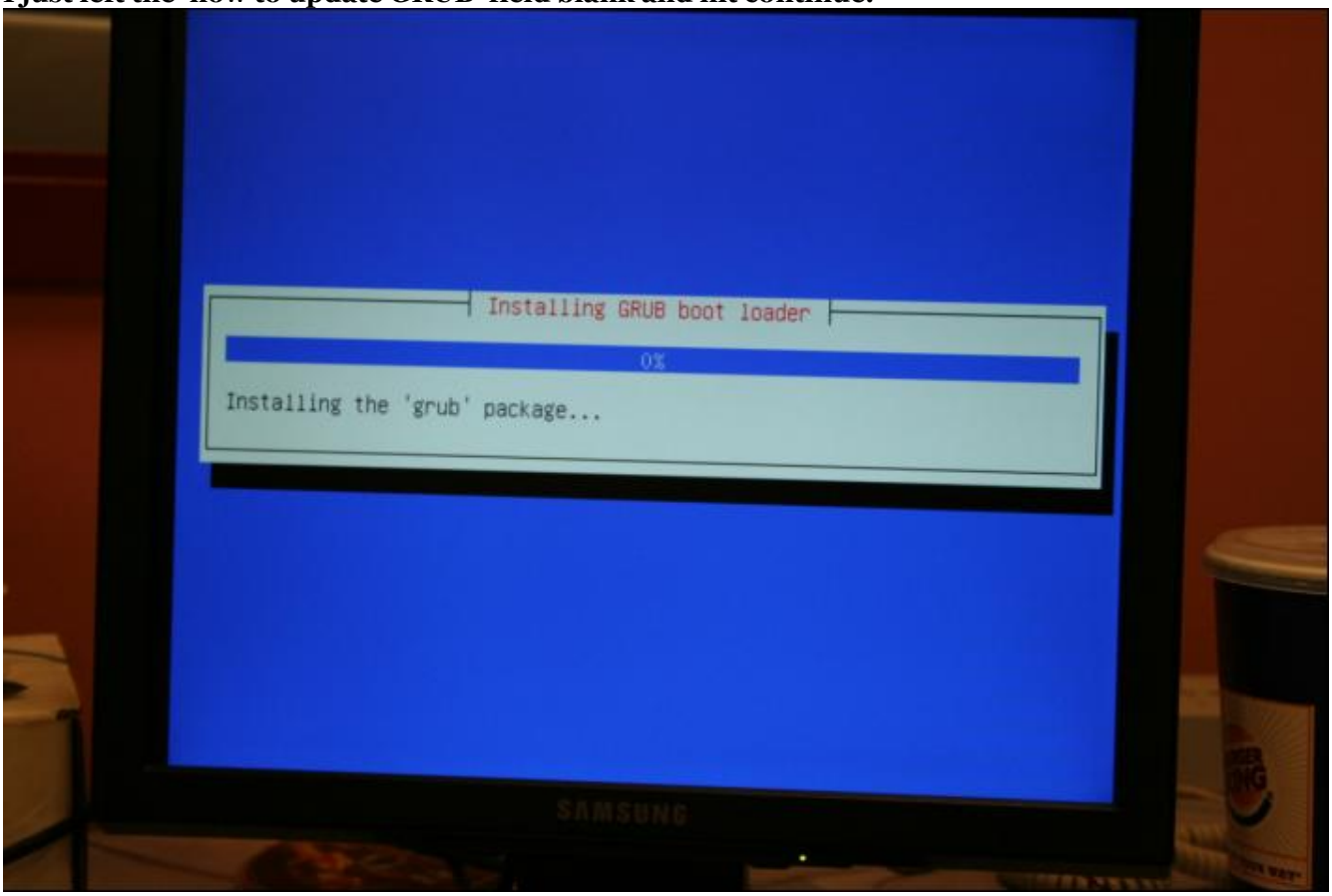

Installing the base system
83%
Preparing linux-ubuntu-modules-2.6.24-19-server
SAMSUNG

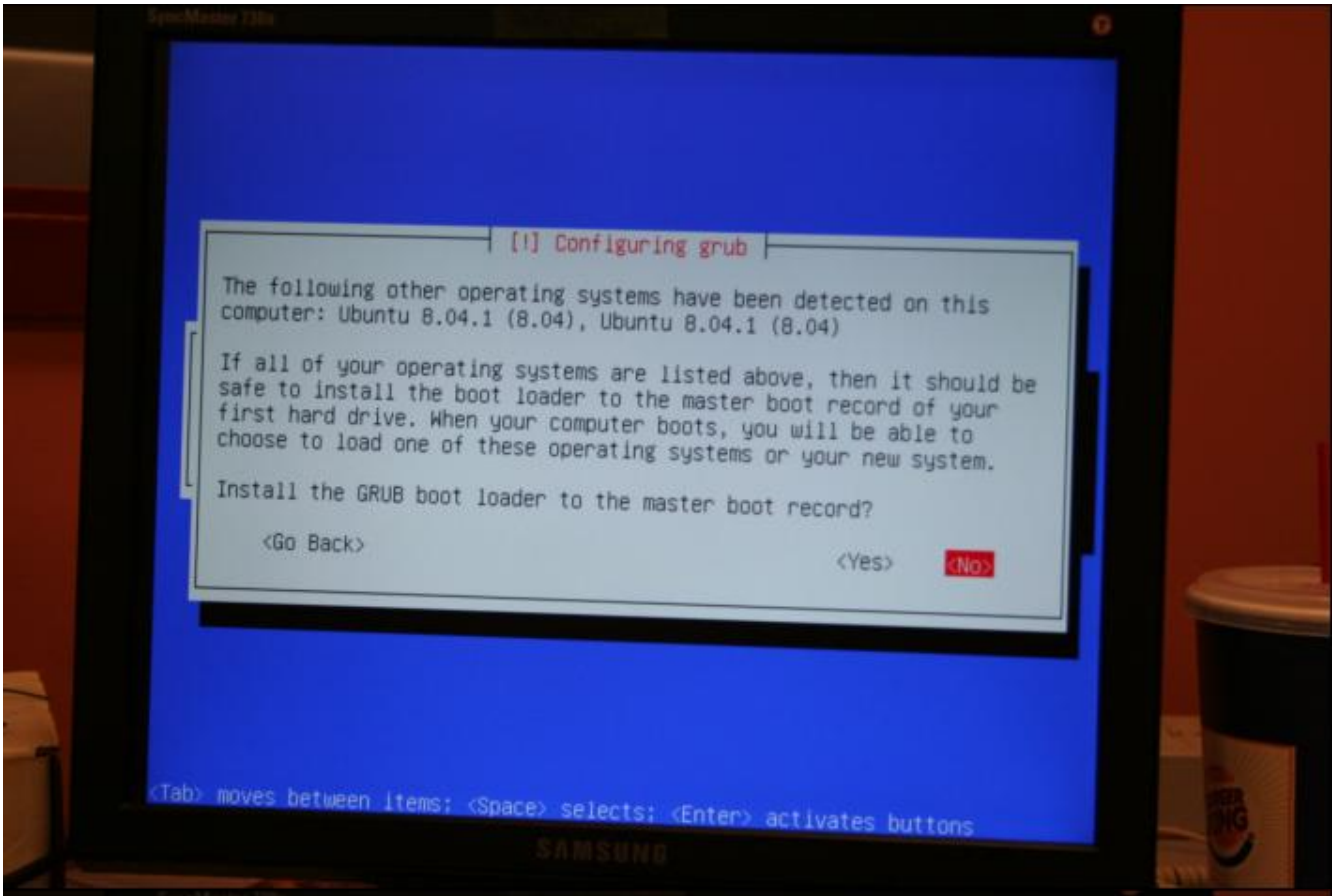

[!!] Set up users and passwords
A user account will be created for you to use instead of the root account for non-administrative activities.
Please enter the real name of this user. This information will be used for instance as default origin for emails sent by this user as well as any program which displays or uses the user's real name. Your full name is a reasonable choice.
Full name for the new user:
jmoore
<Go Back>
<Continue>
<Tab> moves between items; <Space> selects; <Enter> activates buttons
SAMSUNG

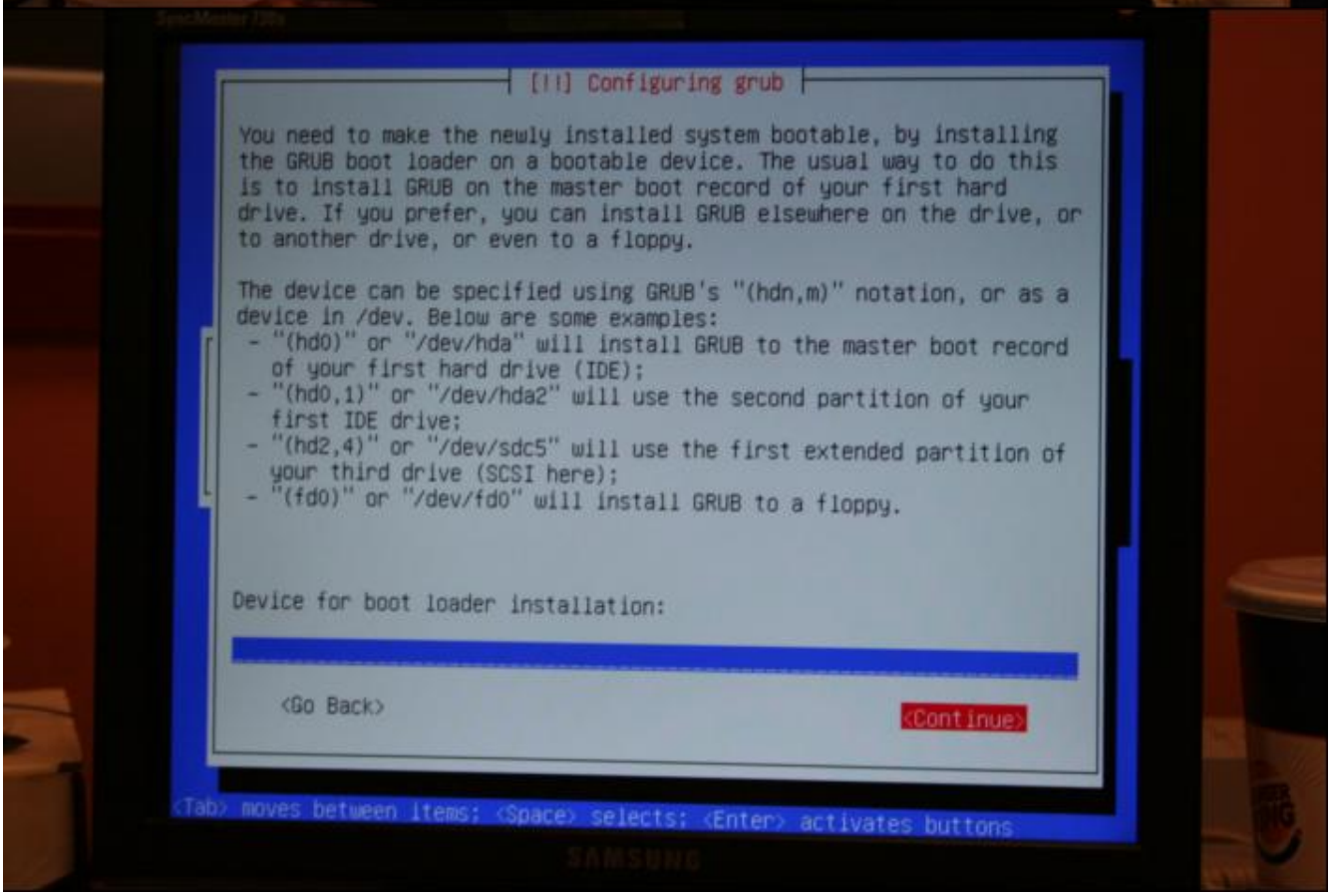

[!!] Set up users and passwords
Please enter the same user password again to verify you have typed it correctly.
Re-enter password to verify:
******
<Go Back>
<Continue>


[!] Software selection
At the moment, only the core of the system is installed. To tune the system to your needs, you can choose to install one or more of the following predefined collections of software.
Choose software to install:
[ ] DNS server
[ ] LAMP server
[ ] Mail server
[ ] OpenSSH server
[ ] PostgreSQL database
[ ] Print server
[ ] Samba File server
<Continue>

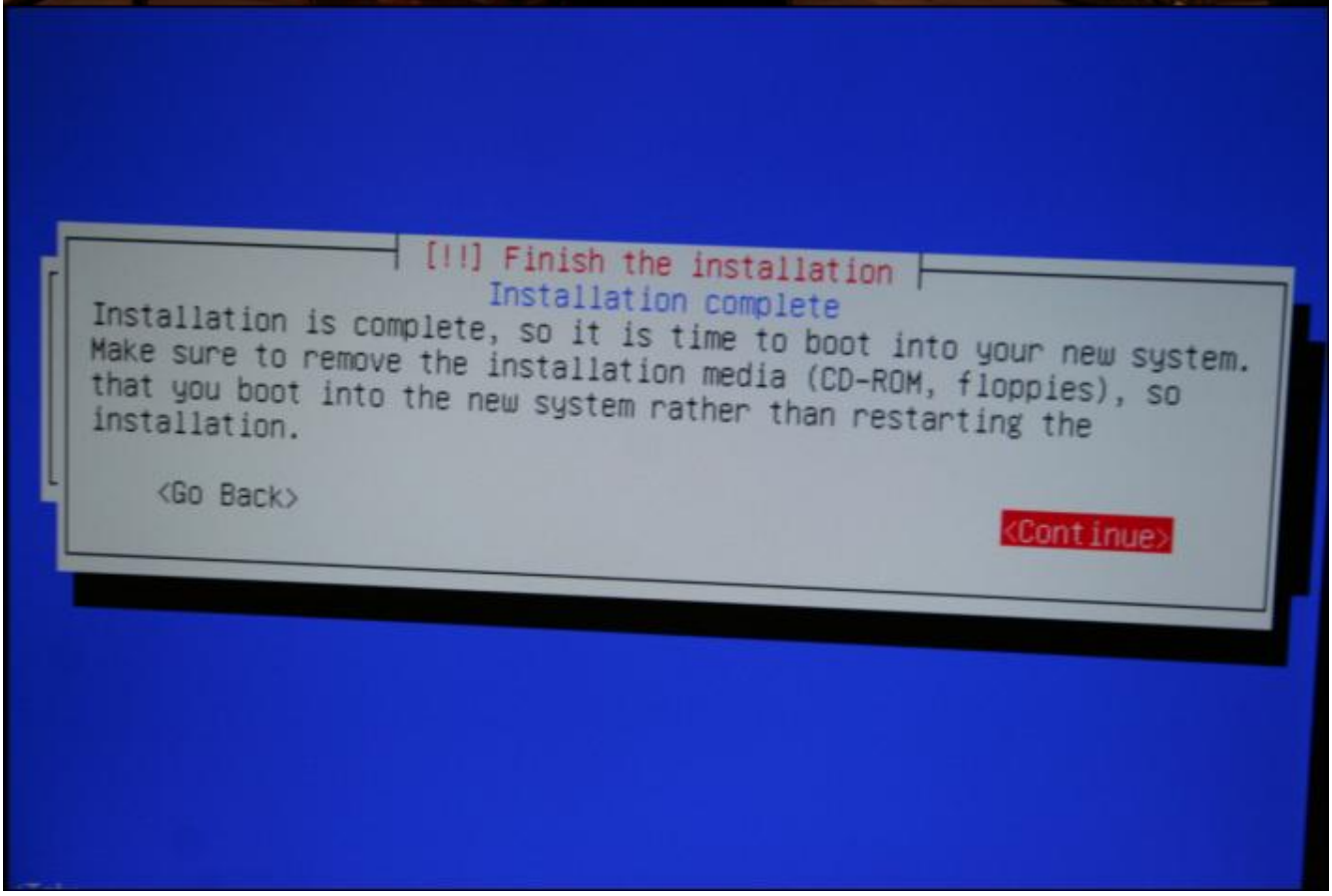
This next part is *<u>CRUCIAL</u>*
-- Do not install GRUB, it will overwrite what you have already done in installing the Head Node!
I just left the 'how to update GRUB' field blank and hit continue.




[!] Configuring grub
The following other operating systems have been detected on this computer: Ubuntu 8.04.1 (8.04), Ubuntu 8.04.1 (8.04)
If all of your operating systems are listed above, then it should be safe to install the boot loader to the master boot record of your first hard drive. When your computer boots, you will be able to choose to load one of these operating systems or your new system.
Install the GRUB boot loader to the master boot record?
<Go Back>
<Yes>
<No>
<Tab> moves between items; <Space> selects; <Enter> activates buttons


[!!] Configuring grub
You need to make the newly installed system bootable, by installing the GRUB boot loader on a bootable device. The usual way to do this is to install GRUB on the master boot record of your first hard drive. If you prefer, you can install GRUB elsewhere on the drive, or to another drive, or even to a floppy.
The device can be specified using GRUB's "(hdn,m)" notation, or as a device in /dev. Below are some examples:
- "(hd0)" or "/dev/hda" will install GRUB to the master boot record of your first hard drive (IDE);
- "(hd0,1)" or "/dev/hda2" will use the second partition of your first IDE drive;
- "(hd2,4)" or "/dev/sdc5" will use the first extended partition of your third drive (SCSI here);
- "(fd0)" or "/dev/fd0" will install GRUB to a floppy.
Device for boot loader installation:
<Go Back>
<Continue>
<Tab> moves between items; <Space> selects; <Enter> activates buttons

[!] Finish the installation
System clocks are generally set to Coordinated Universal Time (UTC).
The operating system uses your time zone to convert system time into
local time. This is recommended unless you also use another operating
system that expects the clock to be set to local time.
Is the system clock set to UTC?
<Go Back>
<Yes>
<No>
<Tab> moves between items; <Space> selects; <Enter> activates buttons
SAMSUNG


[!!] Finish the installation
Installation complete
Installation is complete, so it is time to boot into your new system.
Make sure to remove the installation media (CD-ROM, floppies), so
that you boot into the new system rather than restarting the
installation.
<Go Back>
<Continue>

After the Install

I'm going to outline the process for one node. After you do the first node, you can either repeat all the steps, including the installation, replacing every instance of "node1" with "node2".

Commands for the Head Node / Master Node

```
sudo mkdir -p /nodes/nfs/node1
```

Next we need to update the fstab.
Update the fstab to reflect your new partition.

```
/dev/sda6       /nodes/nfs/node1     ext3     noatime          0          0
```

Here is my complete fstab on the Head Node / Master Node

```
# /etc/fstab: static file system information.
#
# <file system> <mount point>   <type>  <options>       <dump>  <pass>
proc            /proc           proc    defaults        0       0
# /dev/sda2
UUID=1aa2ddf9-f7f4-4f11-935e-000ebde2c5cd /               ext3    relatime,errors=remount-ro 0       1
# /dev/sda12
UUID=a74802af-473c-46aa-bb70-90772ced1463 /home           ext3    relatime        0       2
# /dev/sda11
UUID=cdc7f7ec-aafc-4875-9e02-e7066e5e7c7e none            swap    sw              0       0
/dev/scd0       /media/cdrom0   udf,iso9660 user,noauto,exec,utf8 0       0
/dev/sda5       /nodes/nfs/node1        ext3    noatime 0       0
/dev/sda6       /nodes/nfs/node2        ext3    noatime 0       0
/dev/sda7       /nodes/nfs/node3        ext3    noatime 0       0
```

Mount everything in fstab by typing

```
sudo mount -a
```

If your /etc/exports file does not look like this:

```
# /etc/exports: the access control list for filesystems which may be exported
#               to NFS clients.  See exports(5).
#
# Example for NFSv2 and NFSv3:
# /srv/homes       hostname1(rw,sync) hostname2(ro,sync)
#
# Example for NFSv4:
# /srv/nfs4        gss/krb5i(rw,sync,fsid=0,crossmnt)
# /srv/nfs4/homes  gss/krb5i(rw,sync)
#

/usr                192.168.2.0/24(ro,no_subtree_check)
/nodes/nfs/node1        192.168.2.0/24(rw,no_root_squash,sync,no_subtree_check)
/nodes/nfs/node2        192.168.2.0/24(rw,no_root_squash,sync,no_subtree_check)
/nodes/nfs/node3        192.168.2.0/24(rw,no_root_squash,sync,no_subtree_check)
/nodes              192.168.2.0/24(rw,no_root_squash,sync,no_subtree_check)
/home               192.168.2.0/24(rw,no_root_squash,sync,no_subtree_check)
```

Fix it. (:

We now need to add updates for the DHCP server so that it will find our new node when it boots.

For example:

```
host node1 {
      hardware ethernet 00:50:8D:BC:A2:32;
      fixed-address 192.168.2.5;
      option root-path "/nodes/nfs/node1";
}
```

Here is my complete dhcpd.conf:

```
#
# Sample configuration file for ISC dhcpd for Debian
#
# Attention: If /etc/ltsp/dhcpd.conf exists, that will be used as
# configuration file instead of this file.
#
# $Id: dhcpd.conf,v 1.1.1.1 2002/05/21 00:07:44 peloy Exp $
#

# The ddns-updates-style parameter controls whether or not the server will
# attempt to do a DNS update when a lease is confirmed. We default to the
# behavior of the version 2 packages ('none', since DHCP v2 didn't
# have support for DDNS.)
ddns-update-style none;
#allow booting;
#allow bootp;
# option definitions common to all supported networks...
option domain-name "furman.edu";
option subnet-mask 255.255.255.0;
option domain-name-servers 156.143.128.18, 156.143.128.1;
next-server 192.168.2.1; #TFTP server
filename "/tftpboot/pxelinux.0";

subnet 192.168.2.0 netmask 255.255.255.0 {
      range 192.168.2.2 192.168.2.200;
      option domain-name-servers 192.168.2.1;
      option broadcast-address 192.168.2.255;
      option routers 192.168.2.1;
}

host node1 {
      hardware ethernet 00:50:8D:BC:A2:32;
      fixed-address 192.168.2.5;
      option root-path "/nodes/nfs/node1";
}

host node2 {
      hardware ethernet 00:50:8D:BC:9D:28;
      fixed-address 192.168.2.9;
      option root-path "/nodes/nfs/node2";
}

host node3 {
      hardware ethernet 00:50:8D:B7:71:22;
      fixed-address 192.168.2.13;
      option root-path "/nodes/nfs/node3";
```

```
}
```

Restart the DHCP server with:

sudo /etc/init.d/dhcp3-server restart

Now we have to make some changes to the filesystem that was created with the server installation cd, because there are certain settings that were put in that are not correct with respect to that diskless node. For example, the hard drive and the CD/DVDROM drive are set to mount in the node's fstab, but the node does not have anything connected to it, so those mounts will fail.

Commands for the Headless Node

We now change root over to that new partition (we can pretend like we're loaded on Linux on that partition and begin to make some changes to the disk, even before the diskless node gets a chance to boot)

sudo chroot /nodes/nfs/node1 /bin/bash

I get a weird situation when I try to use tab completion, so I just type out the commands in full while chrooted.

Use your favorite text editor to edit this file, I use vim – lots of people don't like vim, but whatever.

sudo vim /etc/initramfs-tools/initramfs.conf

Look for the line that says:

BOOT=local

And then change it to:

BOOT=nfs

Note: This next section only needs to be done on the first node install. The other nodes can use the same *initrd* and *vmlinuz* – e.g. filesystem and kernel, respectively

To know your kernel version run the command

uname –r

Now, just realize that the generated files of the next command will use that kernel name in their filename.

sudo update-initramfs –u

This outputs two files to the default location of /boot/initrd-img-KERNEL and /boot/vmlinuz-KERNEL. Exit this by doing

exit

Now copy over these files with:

sudo cp /nodes/nfs/node1/boot/initrd.img-KERNEL_VERSION /tftpboot/
sudo cp /nodes/nfs/node1/boot/vmlinuz-KERNEL_VERSION /tftpboot/

Finished this part on the first node – connect back to the chroot now

sudo chroot /nodes/nfs/node1 /bin/bash

Please tell me you made use of the up arrow and didn't re-type that command. (:

Next I create /etc/resolv.conf and put in the nameserver of the Internet connection on the head node (so that I can have internet access through this chroot).

sudo vim /etc/resolv.conf

Place the following in the file:

nameserver ip.address.of.your.dns.server

Next, let's make the fstab look right.

Comment out the lines that begin with UUID= and /dev/scd0 – Use the # character at the front of the line to comment it out. Add in the following:

```
192.168.2.1:/usr      /usr    nfs     defaults          0       0
192.168.2.1:/home     /home nfs     defaults      0       0
192.168.2.1:/nodes    /nodesnfs     defaults      0       0
```

Here is my complete fstab for node1:

```
# /etc/fstab: static file system information.
#
# <file system> <mount point>   <type>  <options>     <dump>  <pass>
proc          /proc         proc   defaults      0      0
# /dev/sda5
#UUID=41b9f976-fa90-448f-9177-9929ba199aac /            ext3   relatime,errors=remount-ro 0
1
#/dev/scd0      /media/cdrom0   udf,iso9660 user,noauto,exec,utf8 0      0
192.168.2.1:/usr      /usr   nfs    defaults      0      0
192.168.2.1:/home     /home  nfs    defaults      0      0
192.168.2.1:/nodes    /nodes  nfs    defaults      0      0
```

Be sure to leave the trailing white line.
You can edit your /etc/hosts file now, or wait until you get the node up and running, whatever you fancy.

sudo vim /etc/hosts

Here's my hosts file:

```
127.0.0.1      localhost
127.0.0.1      HeadOfCluster

127.0.0.1      pc0
127.0.0.1      pc1
127.0.0.1      pc2
127.0.0.1      pc3
192.168.2.5    pc4
192.168.2.6    pc5
192.168.2.7    pc6
192.168.2.8    pc7
192.168.2.9    pc8
192.168.2.10   pc9
192.168.2.11   pc10
192.168.2.12   pc11
192.168.2.13   pc12
192.168.2.14   pc13
```

```
192.168.2.15    pc14
192.168.2.16    pc15
```

We need to create the nodes directory – (otherwise there won't be a place to mount to)

```
sudo mkdir /nodes
```

Next we need to update the /etc/network/interfaces file so that it will automatically set up our additional Ethernet cards properly.
NOTE I am still experiencing a weird instance where the 3rd NIC on the system takes control on boot and tries to get the NFS mounts. This causes an error because that NIC is not supposed to be third.

Here is my /etc/network/interfaces for my node1:

```
# This file describes the network interfaces available on your system
# and how to activate them. For more information, see interfaces(5).

# The loopback network interface
auto lo
iface lo inet loopback

iface eth0 inet static
address 192.168.2.5
netmask 255.255.255.0
auto eth0

iface eth1 inet static
address 192.168.2.6
netmask 255.255.255.0

auto eth1

iface eth2 inet static
address 192.168.2.7
netmask 255.255.255.0

auto eth2

iface eth3 inet static
address 192.168.2.8
netmask 255.255.255.0

auto eth3
```

Finally, the Ubuntu Server installation does not come with the NFS-common utilities, which includes the binaries necessary to mount an nfs drive. **Imagine that! :** Anyway, now that we're using chroot on the head node and we've created the */etc/resolv.conf* with the right nameserver, we can use the following:

```
sudo apt-get update
```

This will essentially tell you whether your internet is working or not, but you should see some stuff go by about getting package lists, now run the following:

```
sudo apt-get install nfs-common
```

Choose Y to install – don't worry about the fail portion of the "portmap" starting, after all – we're only on chroot,

so the /proc directory is not mounted for this installer.

We are now finished setting up the headless node. Get out of the chroot by typing:

exit

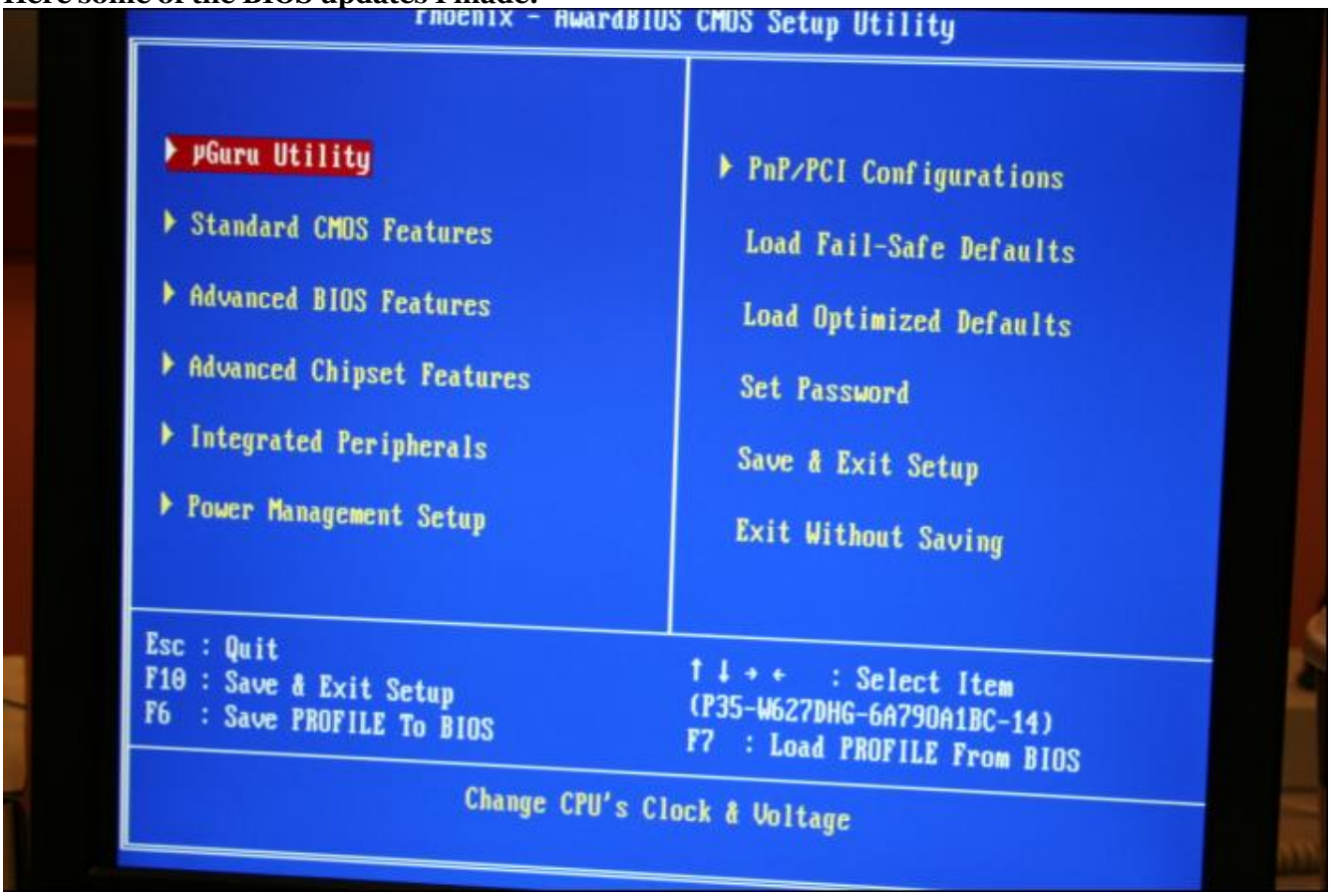
Now we just need to create the proper file in the pxelinux.cfg directory inside of /tftpboot with the MAC address of the new node and I think we're done! Here is my file for my node1:

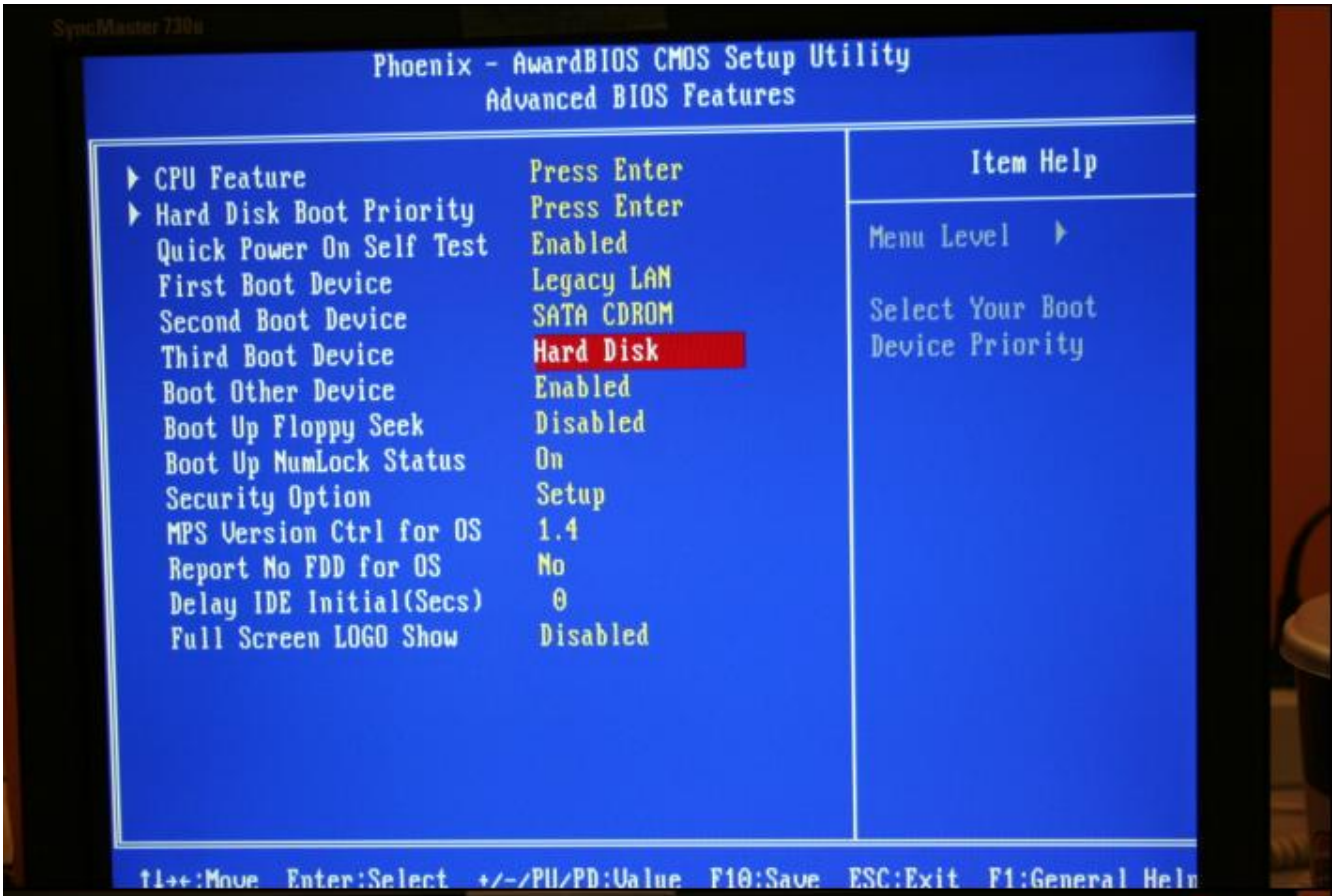
```
default linux

label linux
kernel vmlinuz-2.6.24-19-server
append initrd=initrd.img-2.6.24-19-server \\
      nfsroot=192.168.2.1:/nodes/nfs/node1
1
```

Let's try it out!

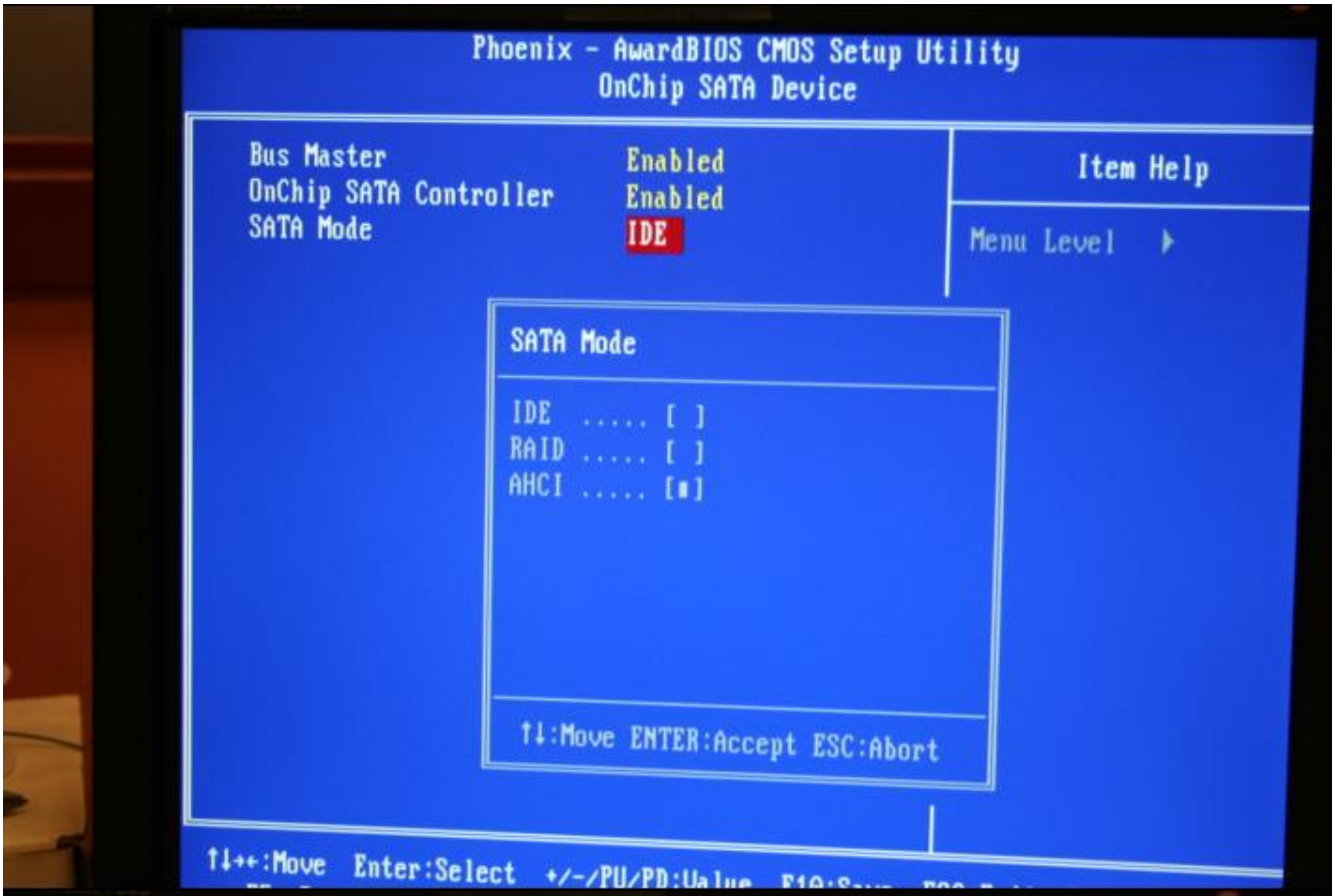
There will be a couple bios changes before this will work – I'll take some pictures.

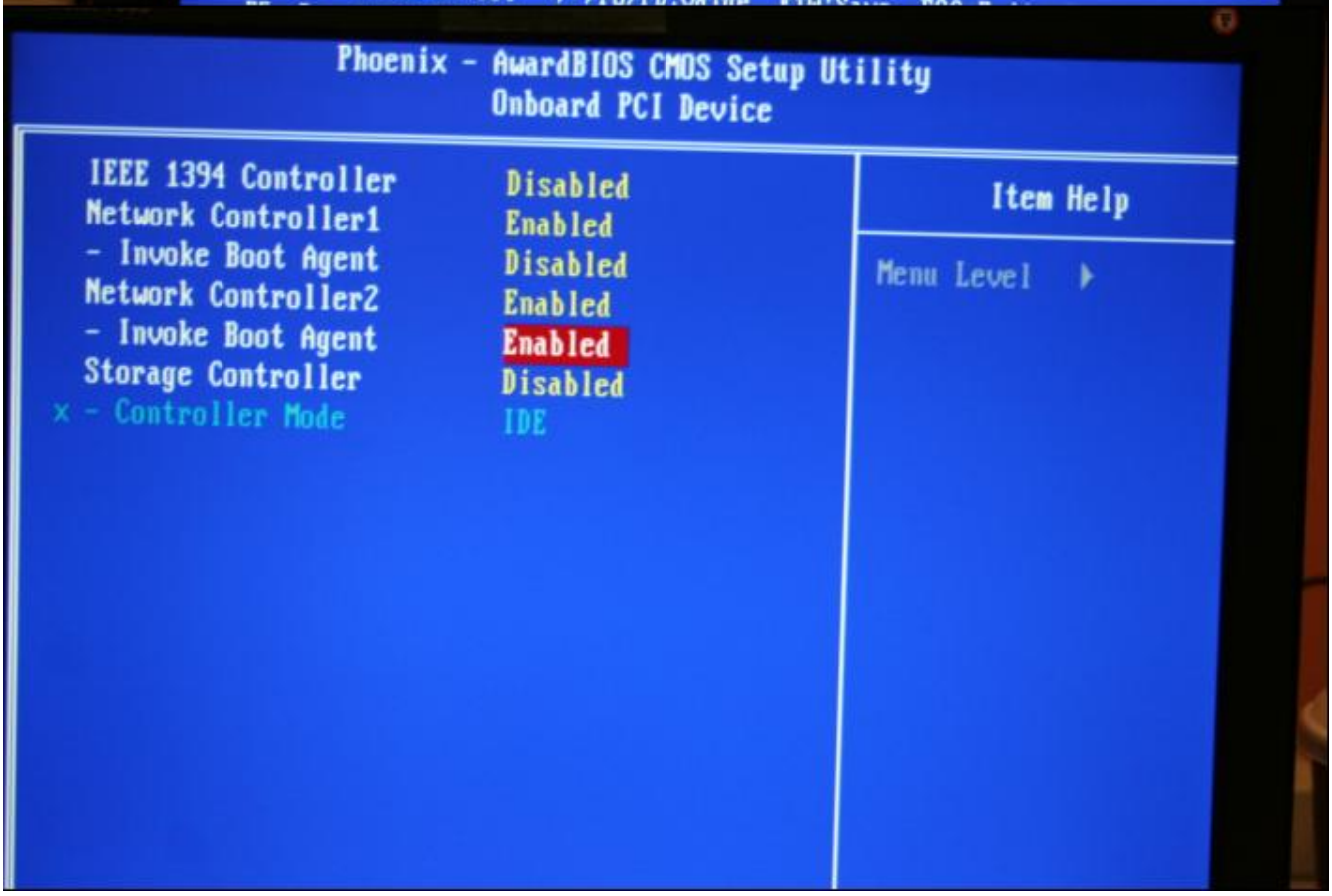
Here some of the BIOS updates I made:



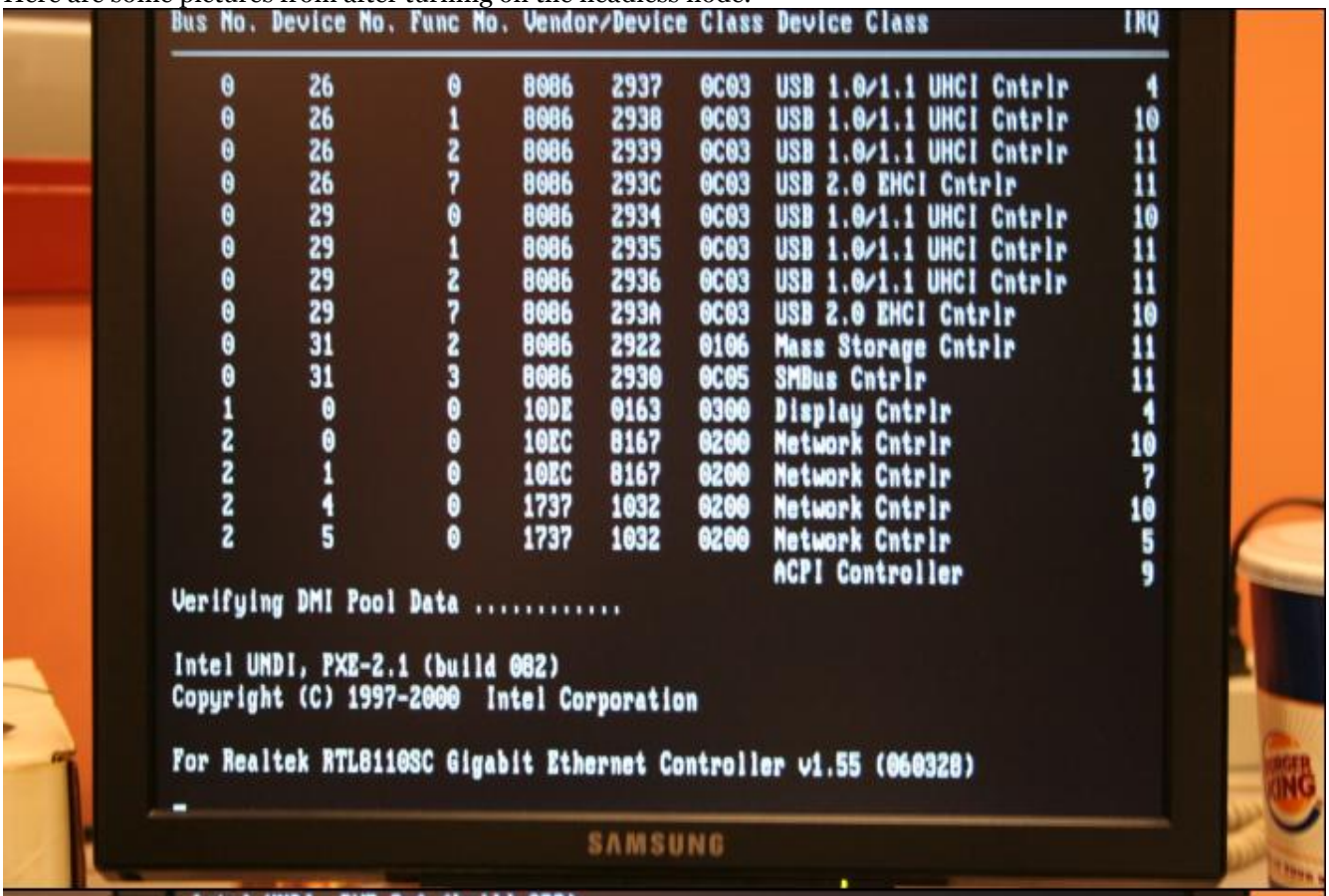

Phoenix - AwardBIOS CMOS Setup Utility
Advanced BIOS Features
CPU Feature Press Enter
Hard Disk Boot Priority Press Enter
Quick Power On Self Test Enabled
First Boot Device Legacy LAN
Second Boot Device SATA CDROM
Third Boot Device Hard Disk
Boot Other Device Enabled
Boot Up Floppy Seek Disabled
Boot Up NumLock Status On
Security Option Setup
MPS Version Ctrl for OS 1.4
Report No FDD for OS No
Delay IDE Initial(Secs) 0
Full Screen LOGO Show Disabled
Item Help
Menu Level
Select Your Boot Device Priority
↑↓→←:Move Enter:Select +/-/PU/PD:Value F10:Save ESC:Exit F1:General Help

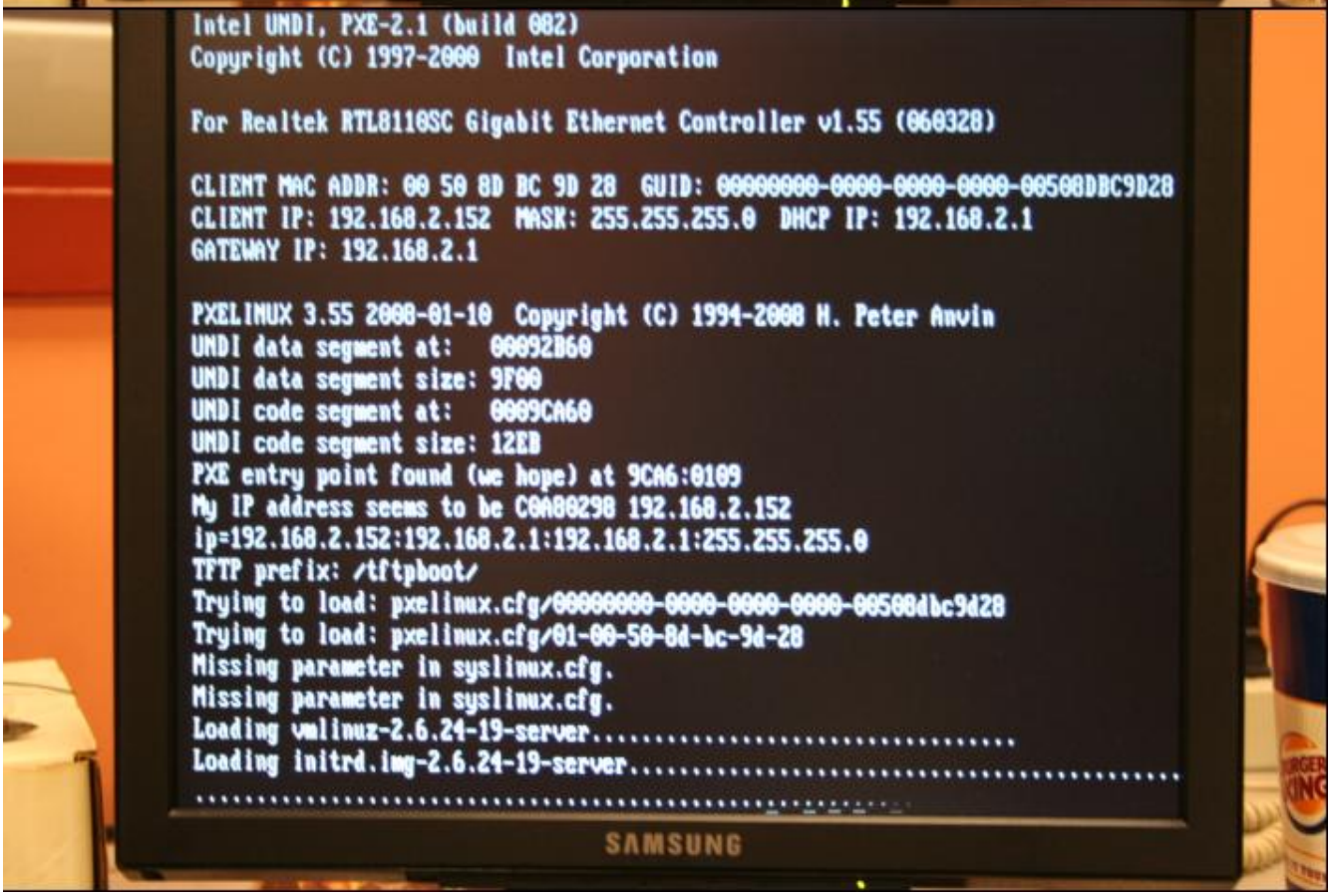
Phoenix - AwardBIOS CMOS Setup Utility
Advanced Chipset Features
DRAM Timing Selectable By SPD
x - CAS Latency Time (tCL) Auto
x - RAS# to CAS# Delay(tRCD) Auto
x - RAS# Precharge (tRP) Auto
x - Precharge Delay
x - Refresh Cycle T
x - Write Recovery
x - Write to Read D
x - Act to Act Time
x - Read to Prechar
x - Command Rate
PCIe Compliancy M
PEG Force X1
Init Display Firs
Item Help
Menu Level
Init Display First
PCI Slot ..... [ ]
PCIEx ..... [*]
↑↓:Move ENTER:Accept ESC:Abort

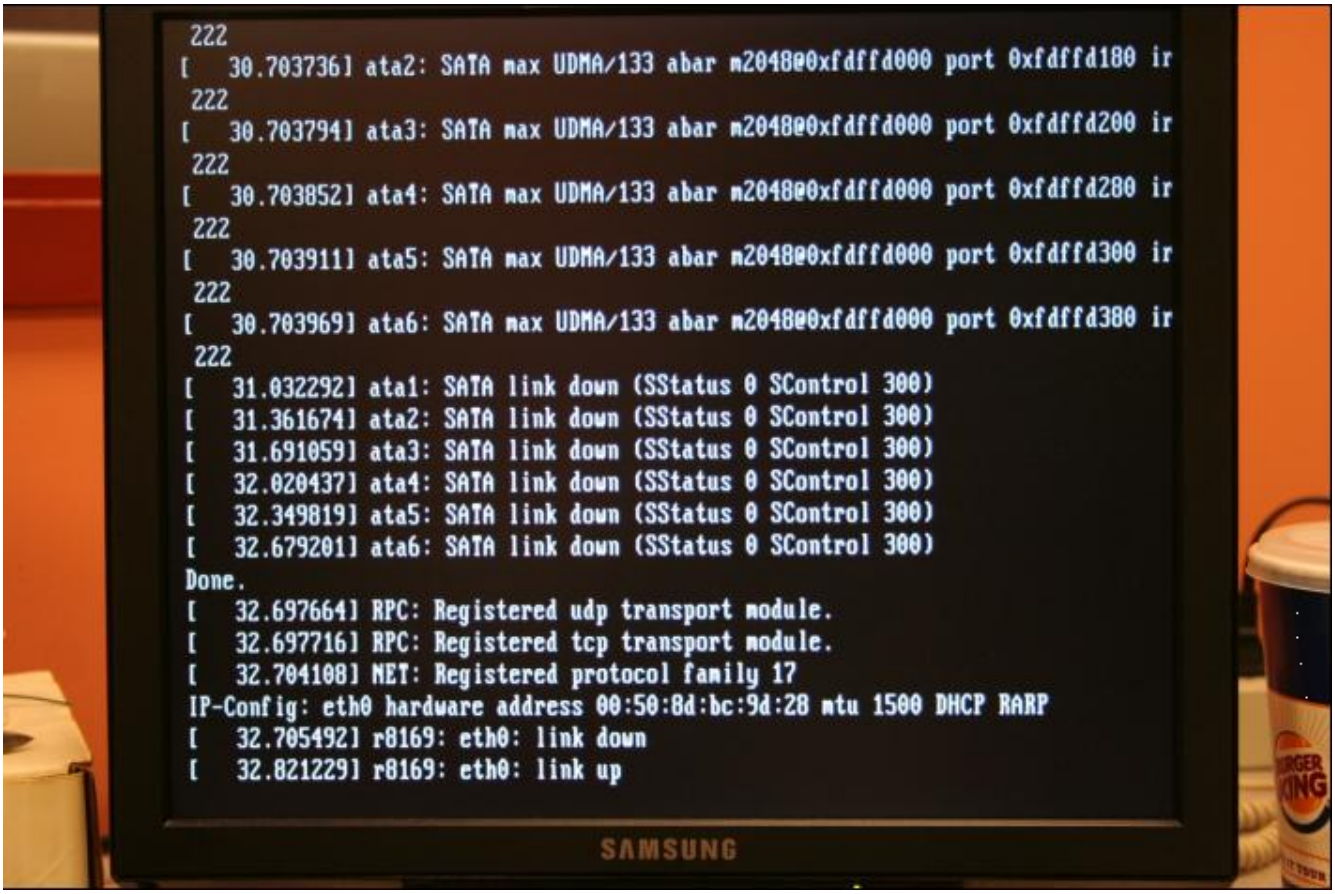

Phoenix - AwardBIOS CMOS Setup Utility
OnChip SATA Device
Bus Master
Enabled
OnChip SATA Controller
Enabled
SATA Mode
IDE
Item Help
Menu Level
SATA Mode
IDE ..... [ ]
RAID ..... [ ]
AHCI ..... [*]
↑↓:Move ENTER:Accept ESC:Abort
↑↓→←:Move Enter:Select +/-/PU/PD:Value


Phoenix - AwardBIOS CMOS Setup Utility
Onboard PCI Device
IEEE 1394 Controller
Disabled
Network Controller1
Enabled
- Invoke Boot Agent
Disabled
Network Controller2
Enabled
- Invoke Boot Agent
Enabled
Storage Controller
Disabled
x - Controller Mode
IDE
Item Help
Menu Level

Here are some pictures from after turning on the headless node.







There will probably be an error here, that's OK – we have now made it easier on ourselves to fix the problem.

After the first boot, when Ubuntu tries to set up the networking, we will easily be able to modify instead of tying it all in manually.

Now we can fix the network interface problem – Ubuntu automatically generates a list of udev rules for networking cards based on the installation and MAC/hardware addresses. The problem is that we have different MAC addresses for each card on each node!! So when the server tries to map out the ethernet cards with the MAC addresses that were stored there when we installed, those Ethernet cards are not on node1. They're residing on the head node, and so we get a weird renaming error. So, I was getting eth1 renamed to eth6, etc...

Here goes the fix:
We will have to chroot over to node1 again from the Head Node.

We now change root over to that new partition (we can pretend like we're loaded on Linux on that partition and begin to make some changes to the disk, even before the diskless node gets a chance to boot)

sudo chroot /nodes/nfs/node1 /bin/bash

I get a weird situation when I try to use tab completion, so I just type out the commands in full while chrooted.

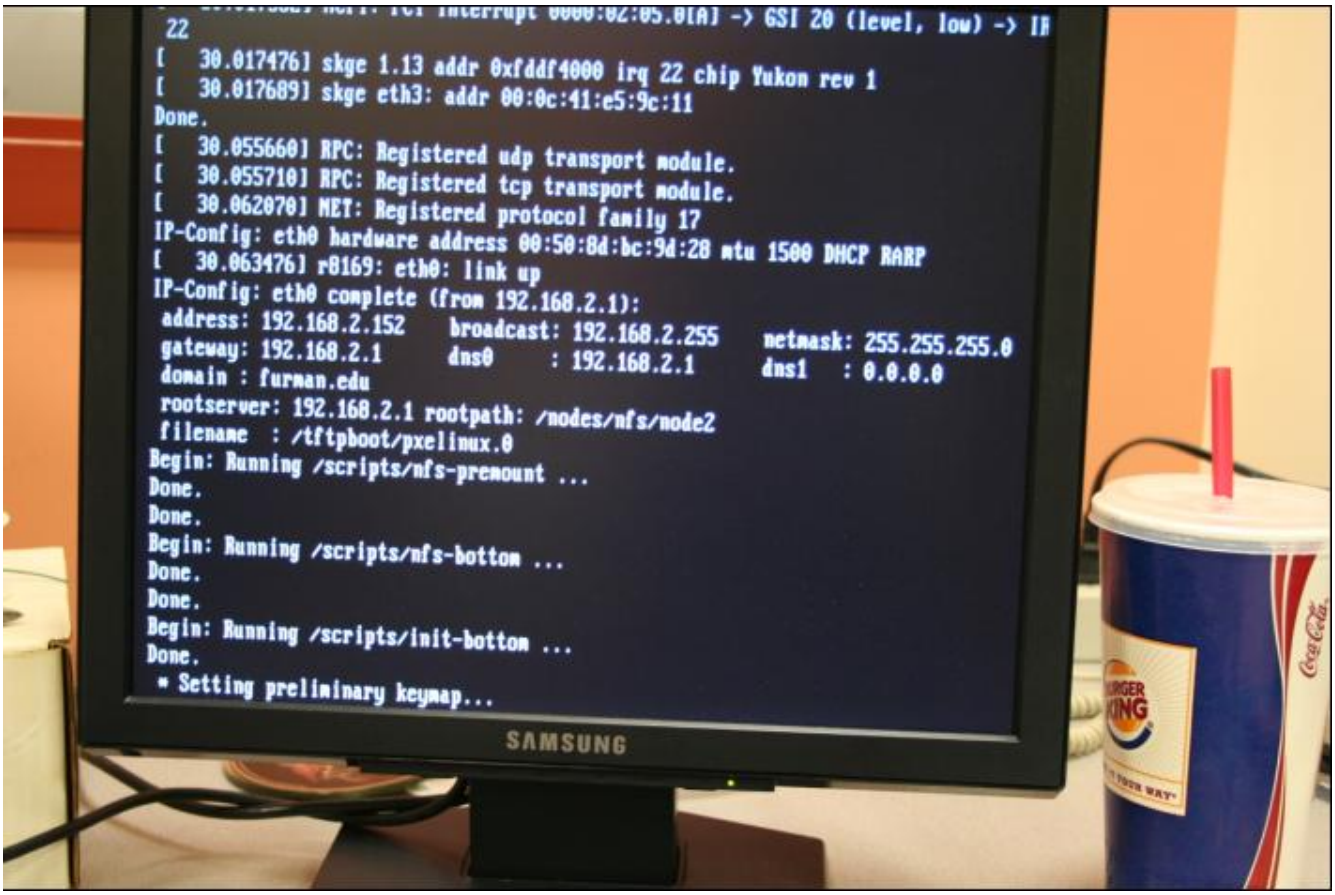
The file is:

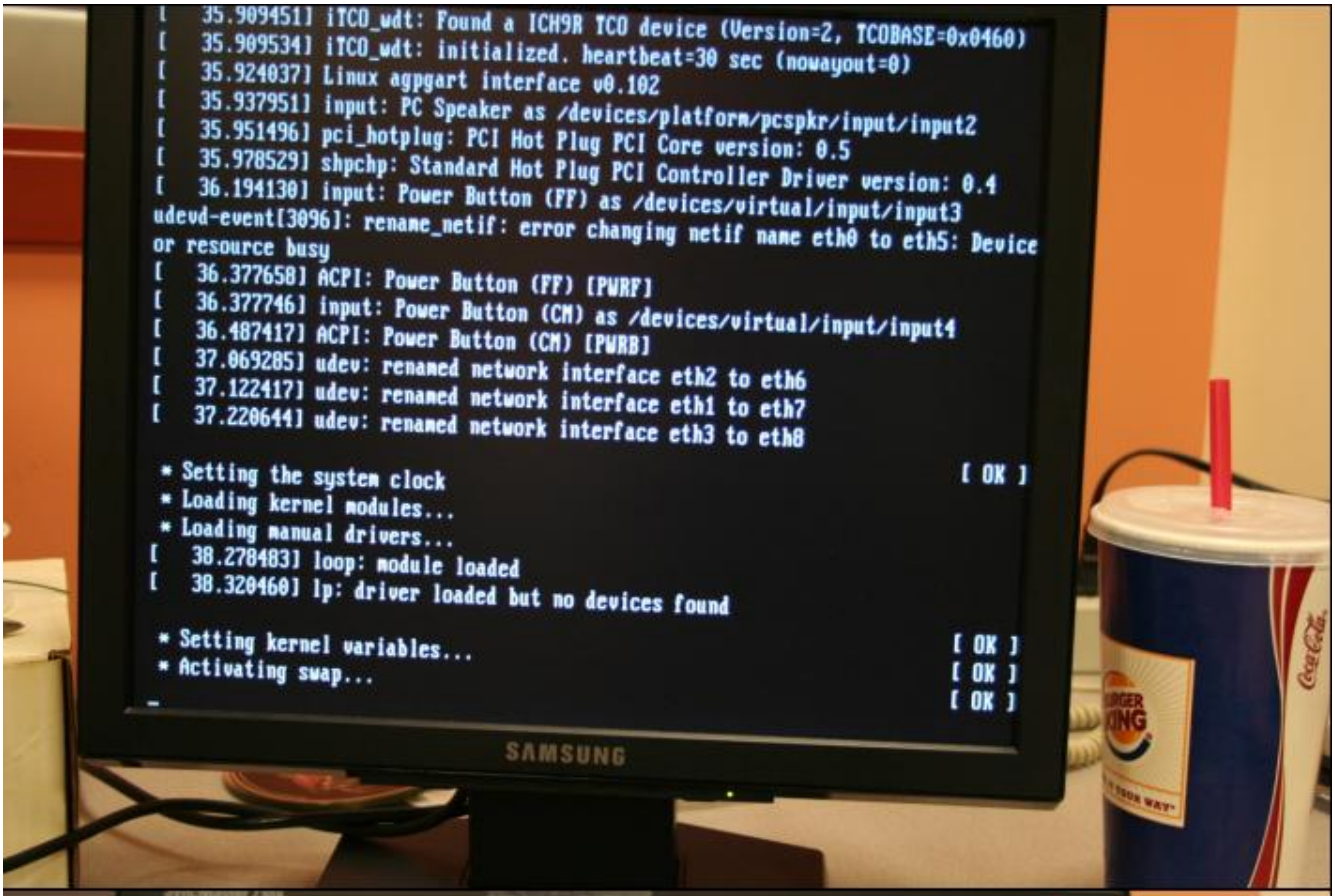
sudo vim /etc/udev/rules.d/70-persistent-net.rules

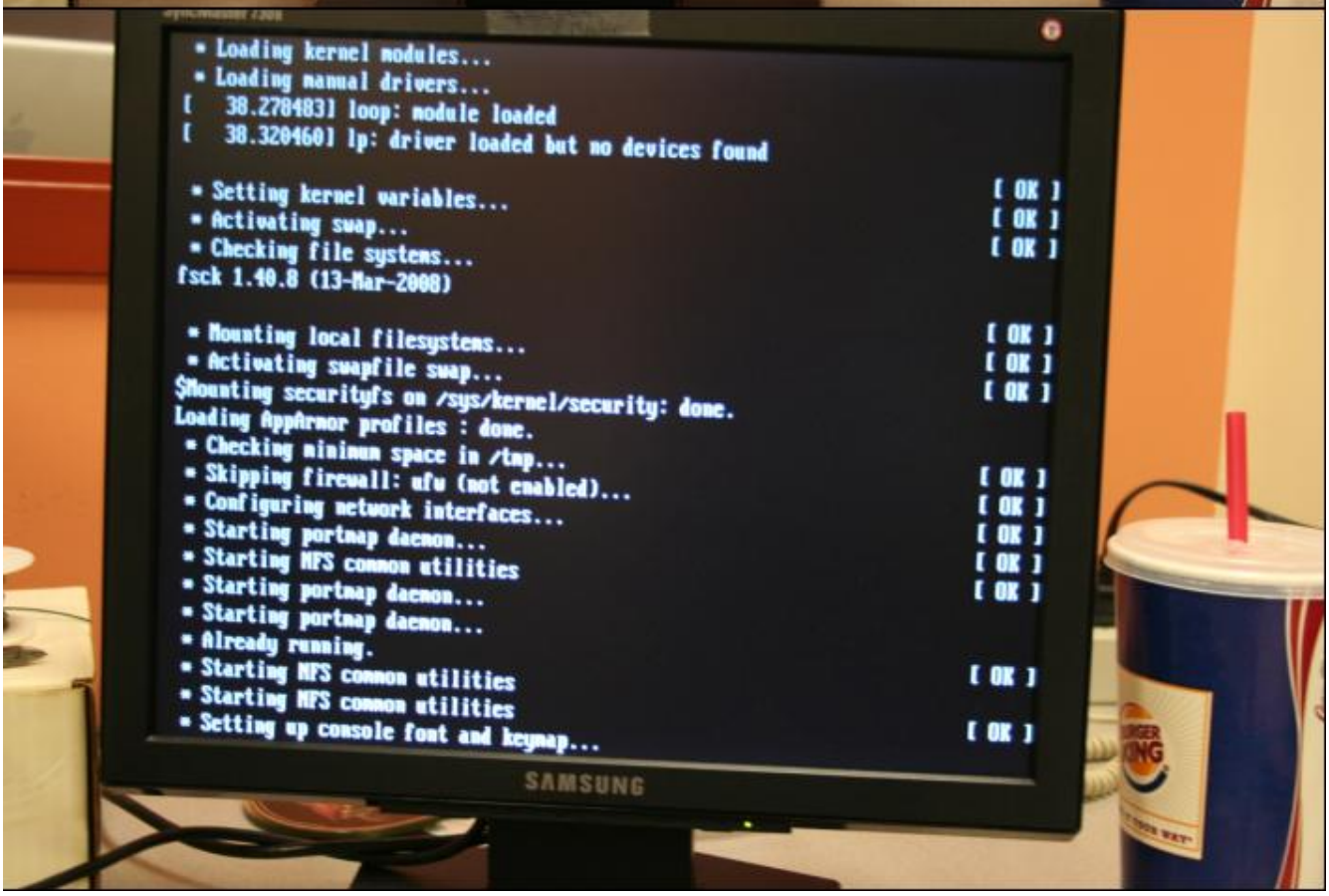
Remove the entries that were there from the head node when we first installed and replace them with the addresses of the cards that you know are in that machine.
We need to make sure our network interfaces are set up properly, with the proper MAC addresses and driver modules.

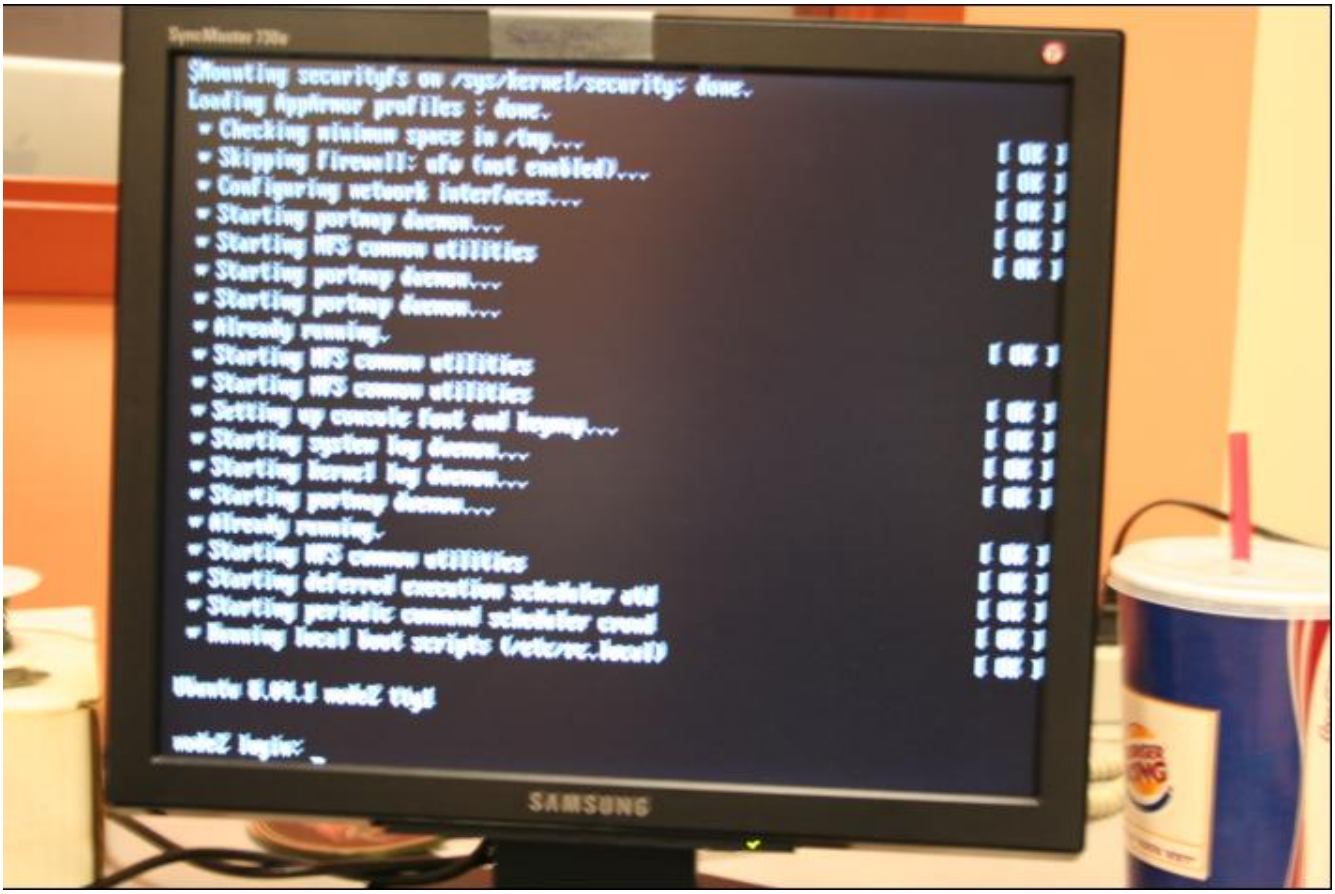
As an example, here is my node1 70-persisten-net.rules:

```
# This file was automatically generated by the /lib/udev/write_net_rules
# program run by the persistent-net-generator.rules rules file.
#
# You can modify it, as long as you keep each rule on a single line.

# PCI device 0x10ec:0x8167 (r8169)
SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="00:50:8d:bc:a2:32",
ATTR{type}=="1", KERNEL=="eth*", NAME="eth0"

#other realtek card.
SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="00:50:8d:bc:a2:33",
ATTR{type}=="1", KERNEL=="eth*", NAME="eth1"

# PCI device 0x1737:0x1032 (skge)
SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="00:0c:41:e5:9e:5e",
ATTR{type}=="1", KERNEL=="eth*", NAME="eth2"

# PCI device 0x1737:0x1032 (skge)
SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="00:0c:41:e5:9e:a8",
ATTR{type}=="1", KERNEL=="eth*", NAME="eth3"
```




[35.909451] iTCO_wdt: Found a ICH9R TCO device (Version=2, TCOBASE=0x0460)
[35.909534] iTCO_wdt: initialized. heartbeat=30 sec (nowayout=0)
[35.924037] Linux agpgart interface v0.102
[35.937951] input: PC Speaker as /devices/platform/pcspkr/input/input2
[35.951496] pci_hotplug: PCI Hot Plug PCI Core version: 0.5
[35.978529] shpchp: Standard Hot Plug PCI Controller Driver version: 0.4
[36.194130] input: Power Button (FF) as /devices/virtual/input/input3
udevd-event[3096]: rename_netif: error changing netif name eth0 to eth5: Device or resource busy
[36.377658] ACPI: Power Button (FF) [PWRF]
[36.377746] input: Power Button (CM) as /devices/virtual/input/input4
[36.487417] ACPI: Power Button (CM) [PWRB]
[37.069285] udev: renamed network interface eth2 to eth6
[37.122417] udev: renamed network interface eth1 to eth7
[37.220644] udev: renamed network interface eth3 to eth8
* Setting the system clock [OK]
* Loading kernel modules...
* Loading manual drivers...
[38.278483] loop: module loaded
[38.320460] lp: driver loaded but no devices found
* Setting kernel variables... [OK]
* Activating swap... [OK]
[OK]
SAMSUNG


* Loading kernel modules...
* Loading manual drivers...
[38.278483] loop: module loaded
[38.320460] lp: driver loaded but no devices found
* Setting kernel variables... [OK]
* Activating swap... [OK]
* Checking file systems... [OK]
fsck 1.40.8 (13-Mar-2008)
* Mounting local filesystems... [OK]
* Activating swapfile swap... [OK]
$Mounting securityfs on /sys/kernel/security: done.
Loading AppArmor profiles : done.
* Checking minimum space in /tmp...
* Skipping firewall: ufw (not enabled)... [OK]
* Configuring network interfaces... [OK]
* Starting portmap daemon... [OK]
* Starting NFS common utilities [OK]
* Starting portmap daemon... [OK]
* Starting portmap daemon...
* Already running.
* Starting NFS common utilities [OK]
* Starting NFS common utilities
* Setting up console font and keymap... [OK]
SAMSUNG


Mounting securityfs on /sys/kernel/security: done.
Loading AppArmor profiles : done.
* Checking minimum space in /tmp...
* Skipping firewall: ufw (not enabled)...
* Configuring network interfaces...
* Starting portmap daemon...
* Starting NFS common utilities
* Starting portmap daemon...
* Starting portmap daemon...
* Already running.
* Starting NFS common utilities
* Starting NFS common utilities
* Setting up console font and keymap...
* Starting system log daemon...
* Starting kernel log daemon...
* Starting portmap daemon...
* Already running.
* Starting NFS common utilities
* Starting deferred execution scheduler atd
* Starting periodic command scheduler crond
* Running local boot scripts (/etc/rc.local)
[OK]
SAMSUNG

Now to installing LAM-MPI, JAVA, and MPIJAVA

I installed both of these packages, but I don't think they're having any effect, especially not mpich - - because I explicitly compiled mpiJava with the LAM option instead of the default mpich.

JUST FOR DOCUMENTATION – NOT USEFUL
sudo apt-get install mpich-bin
sudo apt-get install lam-runtime

Now here's the rub. I've done most of the work already in zipping up all the files I used, but here is how it goes.
First, I downloaded mpiJava -- http://www.hpjava.org/mpiJava.html
Then I downloaded LAM-MPI -- http://www.lam-mpi.org/
Then I got several version of the JDK – 1.4.2 , Java5, and Java6

The configure options for lam-mpi are in Jeff Larkin's paper [page 16], though I didn't use the Fortran compiler option, so –without-fortran (or something like that) was passed as a parameter.

Unzip the java jdk and put it somewhere in /usr (I chose /usr/lib/jvm/)

Then, make sure you get these 3 things for building lam-mpi:

sudo apt-get install libc6-dev gcc g++

installing mpiJava only consists of passing in the jvm location (/usr/lib/jvm/jdk1.6.0_07 for me.)

you have to:

./configure –with-mpi=lam

or something along those lines

then once the configure is done you have to

sudo make

then the mpiJava/src/scripts/prunjava file (which actually runs the java MPI calls) has a parameter location for the jvm (same as above). Also, I added in the –xms128 –xmx1024 command to increase the heap size though it didn't seem to matter for gsky tests.

To run gsky:

Edit StarMPI.java to change any of the parameters to what you want – this is /home/jmoore/Desktop/mpiJava/examples/starArray/

There is a file in there called machines which looks like this:

```
HeadOfCluster cpu=4
Pc4 cpu=4
Pc8 cpu=4
Pc12 cpu=4
```

This has the hostname / cpu count for the entire cluster

Kick off LAM by running

lamboot –l machines

The –l parameter is crucial because otherwise it will complain about localhost resolving, and if you take localhost out it will complain that localhost is not in there. Catch 22.

The –l says “hey stop being a moron and just use the /etc/hosts file that I provided for you.”

That command should come back without any errors, and probably only spit out one line of text
You can pass –v if you want verbose or –d if you want debug

Then edit RunTests script to make sure that the number of procs param nprocs is less than or equal to the number of processors in the machines file. Ie, in this case if yours looks like mine above then it should be 16.
Then execute the runtests script by:

./RunTests

Viola.
If you have writeToFiles Boolean on then you should get some output in /tmp – if not, you’ll just get visual cues that the software is running because …… will fill the screen and you’ll see the timing analysis breakdown when the program is finished.

BOOTING NOTES

If you have problems booting the node just reboot it. Sounds cliché, sure, but whatever.
The third or fourth NIC occasionally decides that it wants to take control despite the fact that the node had to boot from the first NIC and it also received DHCP on the first NIC.
Go figure.
Anyway, if you see
"SKGE" loading as a driver right before the node tries to get the NFS scripts, then just restart the node.
I call it skeegee, and I hate skeegee.

Reboot and you should get something like r#### -- that is good. The onboard NIC cards are RealTek or something like that – so the r driver loading is what you want.

Reboot till that happens.

SSH KEY GENERATION

http://linuxproblem.org/art_9.html

use that.

If you're adding nodes or logging in for the first time, make sure that you do login via ssh once before you kick off lamboot. If any kind of message comes back from the server then it will error and you will get pissed.

There's an error like "this mac address is new, are you sure you trust this source" or some shit like that.
Just say yes and it won't happen again – and subsequently you can boot this node now with lamboot.