CSCI558L Lab8

IMPROVING TCP PERFORMANCE OVER LOSSY LINKS

10/29/2011

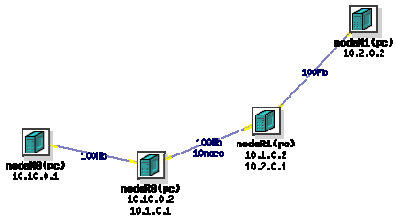
ESHA DESAI
SUELA BUZI
LEON ABURIME
HARSH GUPTA

Contents

Experiment Setup

NS File

```
set ns [new Simulator]
source tb_compat.tcl

# Nodes

set nodeN0 [$ns node]
set nodeR0 [$ns node]
set nodeR1 [$ns node]
set nodeN1 [$ns node]

tb-set-node-os $nodeN0 FBSD8-STD
tb-set-node-os $nodeR0 FBSD8-STD
tb-set-node-os $nodeR1 FBSD8-STD
tb-set-node-os $nodeN1 FBSD8-STD

# Links
set link0 [$ns duplex-link $nodeN0 $nodeR0 100Mb 0.0ms DropTail]
set link1 [$ns duplex-link $nodeR0 $nodeR1 100Mb 10ms DropTail]
#tb-set-link-loss $link1 0.10
set link2 [$ns duplex-link $nodeR1 $nodeN1 100Mb 0.0ms DropTail]

$ns rtproto Manual
$ns run
```



Experiment Details

Virtual Node Info:

ID	Type	OS	Qualified Name
nodeN0	pc	FBSD8-STD	nodeN0.esl.USC558L.isi.deterlab.net
nodeN1	pc	FBSD8-STD	nodeN1.esl.USC558L.isi.deterlab.net
nodeR0	pc	FBSD8-STD	nodeR0.esl.USC558L.isi.deterlab.net
nodeR1	pc	FBSD8-STD	nodeR1.esl.USC558L.isi.deterlab.net

Physical Node Mapping:

ID	Type	OS	Physical
nodeN0	pc3060	FBSD8-STD	pc151
nodeN1	pc3060	FBSD8-STD	pc194
nodeR0	pc3060	FBSD8-STD	pc147
nodeR1	pc3060	FBSD8-STD	pc190
tbdelay0	pc3060	FBSD62-STD	pc157

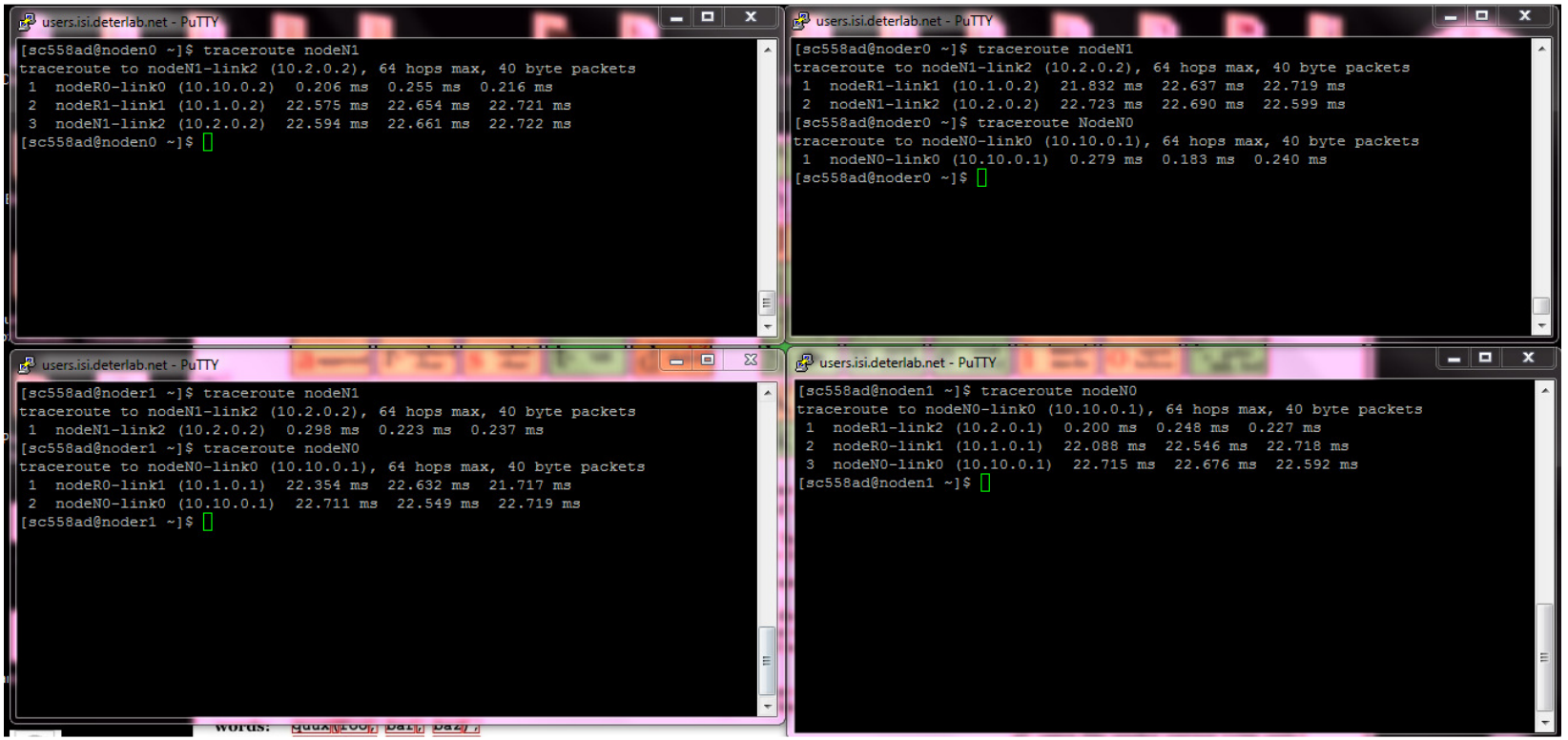
Virtual Lan/Link Info:

ID	Member/Proto	IP/Mask	Delay	BW (Kbs)	Loss Rate
link0	nodeN0:0	10.1.2.2	0.00	100000	0.00000000
ethernet	255.255.255.0	0.00	100000	0.00000000	
link0	nodeR0:0	10.1.2.3	0.00	100000	0.00000000
ethernet	255.255.255.0	0.00	100000	0.00000000	
link1	nodeR0:1	10.1.3.2	5.00	100000	0.00000000
ethernet	255.255.255.0	5.00	100000	0.00000000	
link1	nodeR1:0	10.1.3.3	5.00	100000	0.00000000
ethernet	255.255.255.0	5.00	100000	0.00000000	
link2	nodeN1:0	10.1.1.3	0.00	100000	0.00000000
ethernet	255.255.255.0	0.00	100000	0.00000000	
link2	nodeR1:1	10.1.1.2	0.00	100000	0.00000000
ethernet	255.255.255.0	0.00	100000	0.00000000	

Route configuration

nodeN0	sudo route add 192.168.253.1 192.168.1.254
	sudo route del default 192.168.1.254
	sudo route add default 10.10.0.2
nodeN1	sudo route add 192.168.253.1 192.168.1.254
	sudo route del default 192.168.1.254
	sudo route add default 10.2.0.1
nodeR0	sudo route add 192.168.253.1 192.168.1.254
	sudo route add -net 10.2.0.0/16 10.1.0.2
nodeR1	sudo route add 192.168.253.1 192.168.1.254
	sudo route add -net 10.10.0.0/16 10.1.0.1

Traceroute



Changes up to date made only in the configuration files

sudo vim /etc/inetd.conf
#ftp stream tcp no wait root …….

sudo vim /etc/rc.conf
ftpd_enable="YES"

start ftp daemon
/etc/rc.d/ftp start

ftp nodeN0
ftp> send file

```
220 noden0.esl.usc558l.isi.deterlab.net FTP server (Version 6.00LS) ready.
Name (noden0:sc558ad): sc558ad
331 Password required for sc558ad.
Password:
230 User sc558ad logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> get /mnt/1g
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||55580|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  2% |                                   | 20981 KB    1.20 MB/s    13:32 ETA^
~
```

```
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||57121|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  3% |*                                  | 37560 KB    1.11 MB/s    14:26 ETA^
Z
[12]+  Stopped                 sudo ftp noden0
[sc558ad@noden1 /mnt]$
```

25% link loss at 100ms delay

Experiment USC558L/esl

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you like
you to set the parameters for the *entire* link or lan. If you want to change the values for individual nodes, then enter new values on t
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapou

Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED)			
						q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	50.00	100000	0.12500000	100s	n/a	n/a	n/a	n/a
	nodeR1	50.00	100000	0.12500000	100s	n/a	n/a	n/a	n/a

☑ **Save?** (Check this box if you want these settings to be used next time the experiment is swapped in)

Execute

```
[sc558ad@noden1 ~]$ sudo ftp noden0
Connected to nodeN0-link0.
220 noden0.esl.usc558l.isi.deterlab.net FTP server (Version 6.00LS) ready.
Name (noden0:sc558ad): sc558ad
331 Password required for sc558ad.
Password:
230 User sc558ad logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> get /mnt/1g
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||62670|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  0% |                                     | 24616       1.20 KB/s  - stalled -

  0% |                                     | 24616       1.14 KB/s  - stalled -


  0% |                                     | 24616       1.09 KB/s  - stalled -
  0% |                                     | 24616       0.96 KB/s  - stalled -
```

10ms delay 50% loss

Experiment **USC558L/esl**

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you like at a time.
you to set the parameters for the *entire* link or lan. If you want to change the values for indvidual nodes, then enter new values on the proper
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapout, then che

Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED)			
						q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	10.00	100000	0.25000000	100s	n/a	n/a	n/a	n/a
	nodeR1	10.00	100000	0.25000000	100s	n/a	n/a	n/a	n/a

☑ **Save?** (Check this box if you want these settings to be used next time the experiment is swapped in)

Execute

```
[sc558ad@noden1 ~]$ sudo ftp noden0
Connected to nodeN0-link0.
220 noden0.esl.usc558l.isi.deterlab.net FTP server (Version 6.00LS) ready.
Name (noden0:sc558ad): sc558ad
331 Password required for sc558ad.
Password:
230 User sc558ad logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> get /mnt/1g
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||57761|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  0% |                                     | 14480       1.76 KB/s  - stalled -
  0% |                                     | 14480       1.57 KB/s  - stalled -
  0% |                                     | 14480       1.28 KB/s  - stalled -
  0% |                                     | 14480       1.17 KB/s  - stalled -
  0% |                                     | 14480       1.00 KB/s  - stalled -
  0% |                                     | 31856       2.07 KB/s    --:-- ETA
  0% |                                     | 31856       1.72 KB/s    --:-- ETA
  0% |                                     | 31856       1.63 KB/s    --:-- ETA
  0% |                                     | 31856       1.29 KB/s  - stalled -
  0% |                                     | 36200       1.26 KB/s    --:-- ETA
  0% |                                     | 36200       1.00 KB/s  - stalled -
```

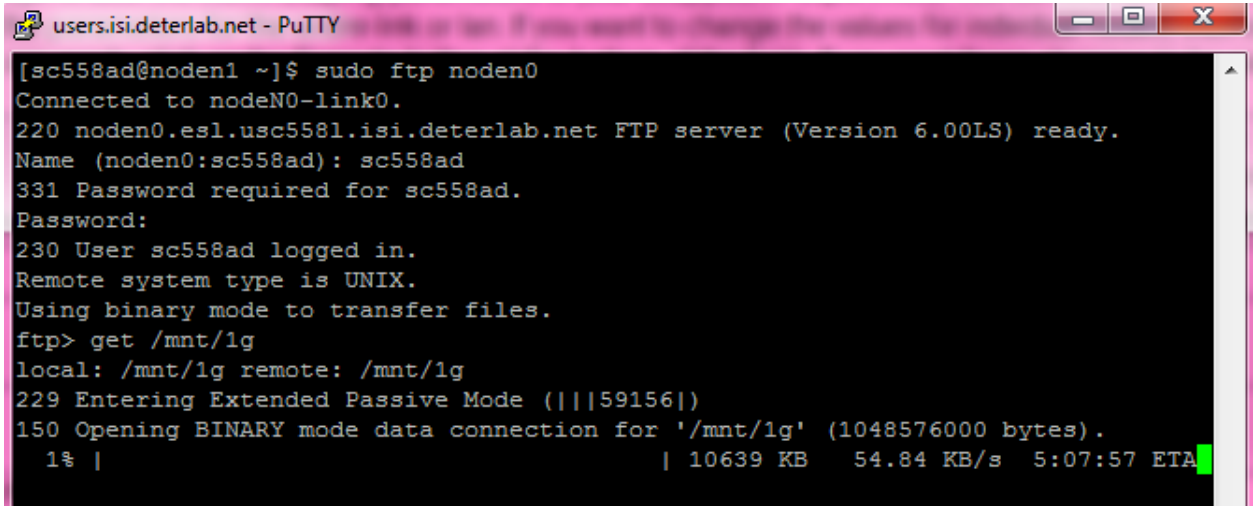
10ms delay and 25% round trip loss

Experiment USC558L/esl

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you like at a
you to set the parameters for the *entire* link or lan. If you want to change the values for indvidual nodes, then enter new values on the
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapout, th

Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED)			
						q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	5.00	100000	0.12500000	100s	n/a	n/a	n/a	n/a
	nodeR1	5.00	100000	0.12500000	100s	n/a	n/a	n/a	n/a

☑ **Save?** (Check this box if you want these settings to be used next time the experiment is swapped in)



10ms delay and 10% round trip loss

Experiment USC558L/esl

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you like
you to set the parameters for the *entire* link or lan. If you want to change the values for indvidual nodes, then enter new values on
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapou

Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED)			
						q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	5.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a
	nodeR1	5.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a

☑ **Save?** (Check this box if you want these settings to be used next time the experiment is swapped in)

```
[sc558ad@noden1 ~]$ sudo ftp noden0
Connected to nodeN0-link0.
220 noden0.esl.usc558l.isi.deterlab.net FTP server (Version 6.00LS) ready.
Name (noden0:sc558ad): sc558ad
331 Password required for sc558ad.
Password:
230 User sc558ad logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> get /mnt/1g
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||58120|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  3% |*                                  | 32144 KB  365.27 KB/s    45:15 ETA
```

0ms delay and 10% round trip loss

Experiment USC558L/esl

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you like at
you to set the parameters for the *entire* link or lan. If you want to change the values for indvidual nodes, then enter new values on the
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapout, th

Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED) q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	0.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a
	nodeR1	0.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a

☑ **Save?** (Check this box if you want these settings to be used next time the experiment is swapped in)

Execute

```
[sc558ad@noden1 ~]$ sudo ftp noden0
Connected to nodeN0-link0.
220 noden0.esl.usc558l.isi.deterlab.net FTP server (Version 6.00LS) ready.
Name (noden0:sc558ad): sc558ad
331 Password required for sc558ad.
Password:
230 User sc558ad logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> get /mnt/1g
local: /mnt/1g remote: /mnt/1g
229 Entering Extended Passive Mode (|||63337|)
150 Opening BINARY mode data connection for '/mnt/1g' (1048576000 bytes).
  4% |*                                  | 48223 KB   698.89 KB/s    23:16 ETA
```

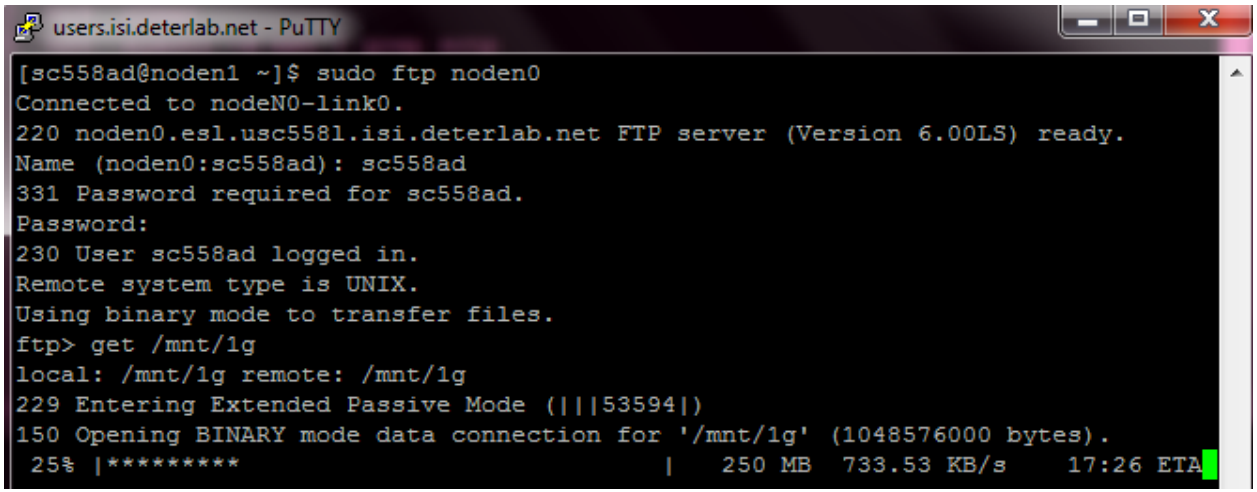
Changing protocol to veno

Experiment USC558L/esl

Use this page to alter the traffic shaping parameters of your *swapped in* experiment. You can change as many values as you lik
you to set the parameters for the *entire* link or lan. If you want to change the values for indvidual nodes, then enter new values or
When you are ready, click on the Execute button at the bottom of the form. If you want these changes to be saved across swapc

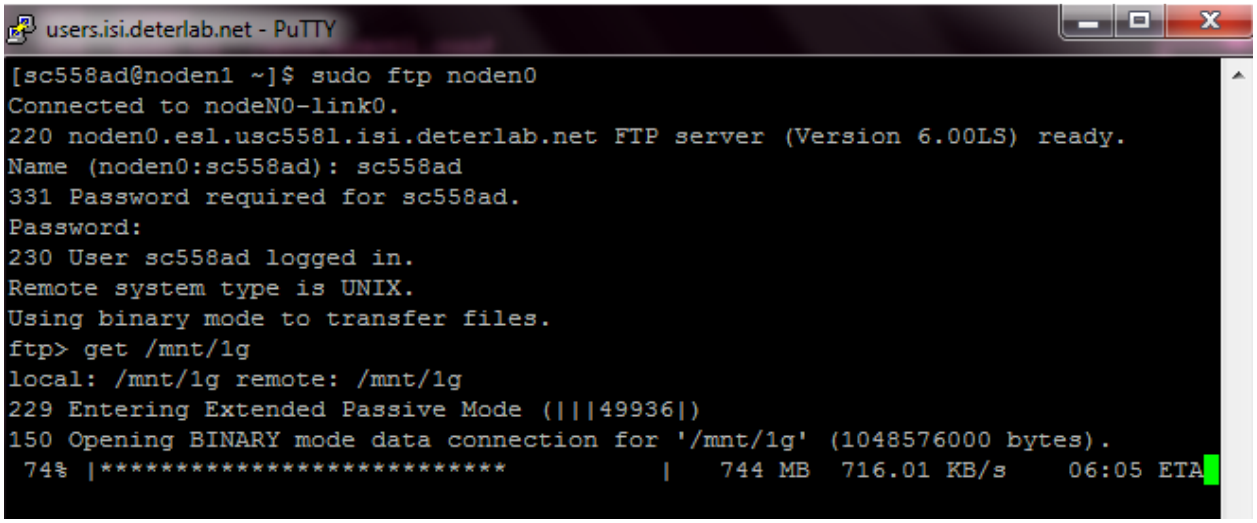
Link Name	Node	Delay (msec)	Bandwidth (kb/s)	Loss (ratio)	Queue Size	RED/GRED (only if link specified as RED)			
						q_weight	minthresh	maxthresh	linterm
link1	All Nodes					n/a	n/a	n/a	n/a
	nodeR0	0.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a
	nodeR1	0.00	100000	0.05000000	100s	n/a	n/a	n/a	n/a

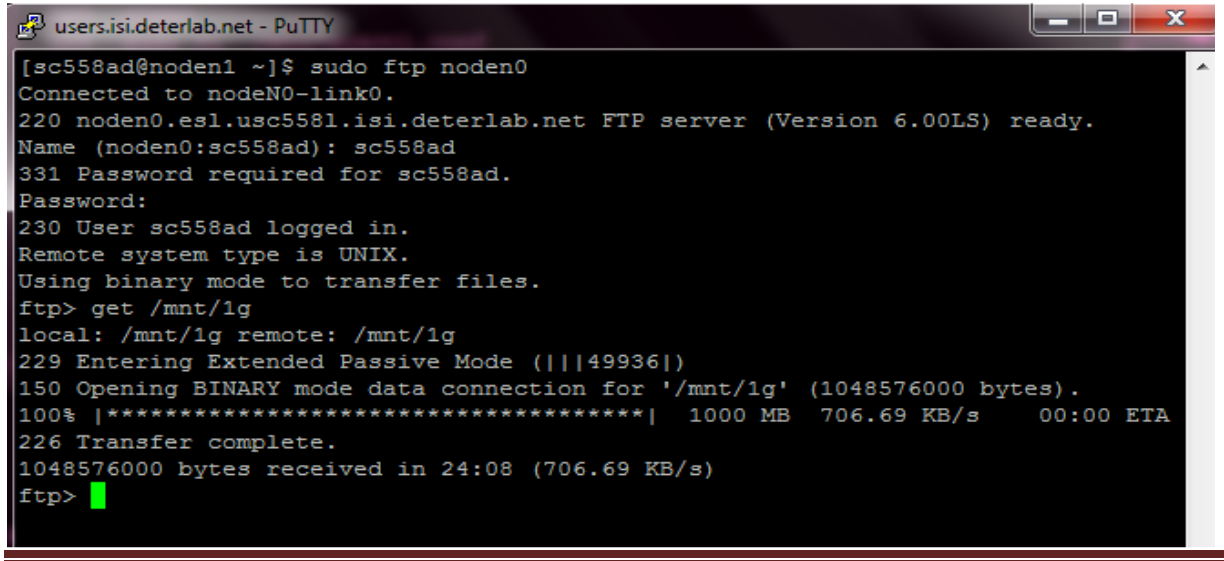
Save? (Check this box if you want these settings to be used next time the experiment is swapped in)



Showing some improvement

Changing it in all nodes (except receiver) and not only on the sender node.





Changed the following in the /etc/sysctl.conf

Changing tcp window size

net.inet.tcp.recvspace=32768
net.inet.tcp.sendspace=32768

Performance went down .

Experiment with Ubuntu 10.04 version:

We had initially started off with Ubuntu 10.04 as the reference links in the Lab8 pdf were related to Linux kernel tunings.

We found that the **default Algorithm** followed by tcp in ubuntu was **cubic**.

The default tcp throughput measured via ftp was 10984 KBps.

```
226 Transfer complete.
1048576000 bytes received in 93.22 secs (10984.3 kB/s)
ftp>
ftp> get /mnt/test2.img /mnt/test_here2.img
local: /mnt/test_here2.img remote: /mnt/test2.img
200 PORT command successful. Consider using PASV.
150 Opening BINARY mode data connection for /mnt/test2.img (1048576000 bytes).
226 Transfer complete.
1048576000 bytes received in 93.70 secs (10927.9 kB/s)
ftp> exit
```

After trying different modules, at loss = 0% and delay = 0ms, the performance was best given by veno algorithm and hence we decided to tweak tcp **veno** parameters.

Switching between the different algorithms could be easily done, by writing text to a /proc/ entry.

```
root@node0:~# echo "veno" > /proc/sys/net/ipv4/tcp_congestion_control

root@node0:~# cat /proc/sys/net/ipv4/tcp_congestion_control

veno
```

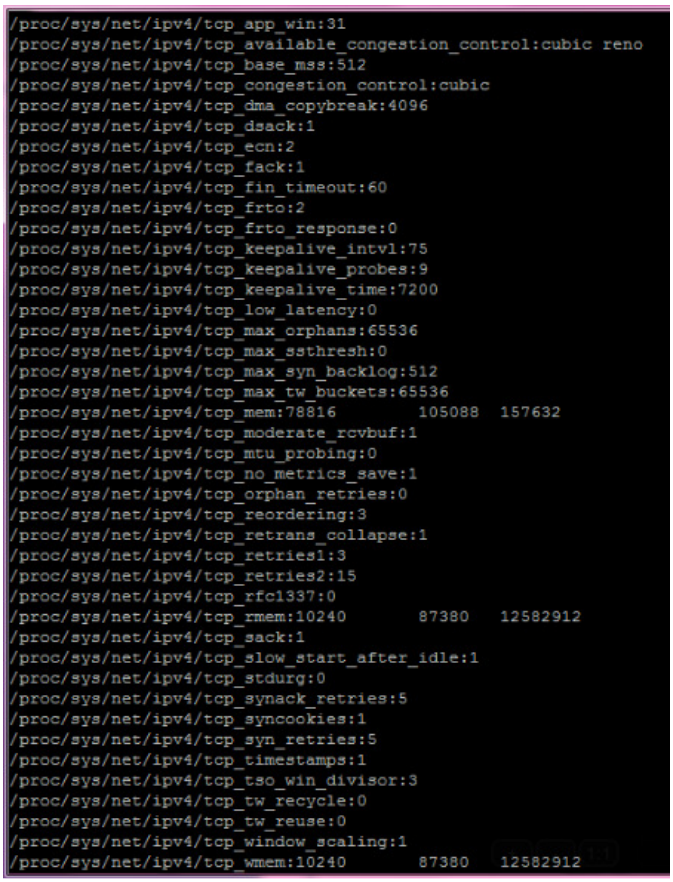
A list of available modules can be found here:

```
root@node0:~# ls /lib/modules/`uname -r`/kernel/net/ipv4/

ip_gre.ko  netfilter   tcp_cubic.ko   tcp_htcp.ko   tcp_lp.ko   tcp_vegas.koipip.ko
tcp_bic.ko  tcp_highspeed.ko  tcp_hybla.ko  tcp_scalable.ko  tcp_veno.ko
```

We retained the traffic settings (0ms delay and 0% loss) and changed the tcp parameters :

After some research we found that we could do a few settings to make the performance a little better.

The default TCP parameters are as follows:



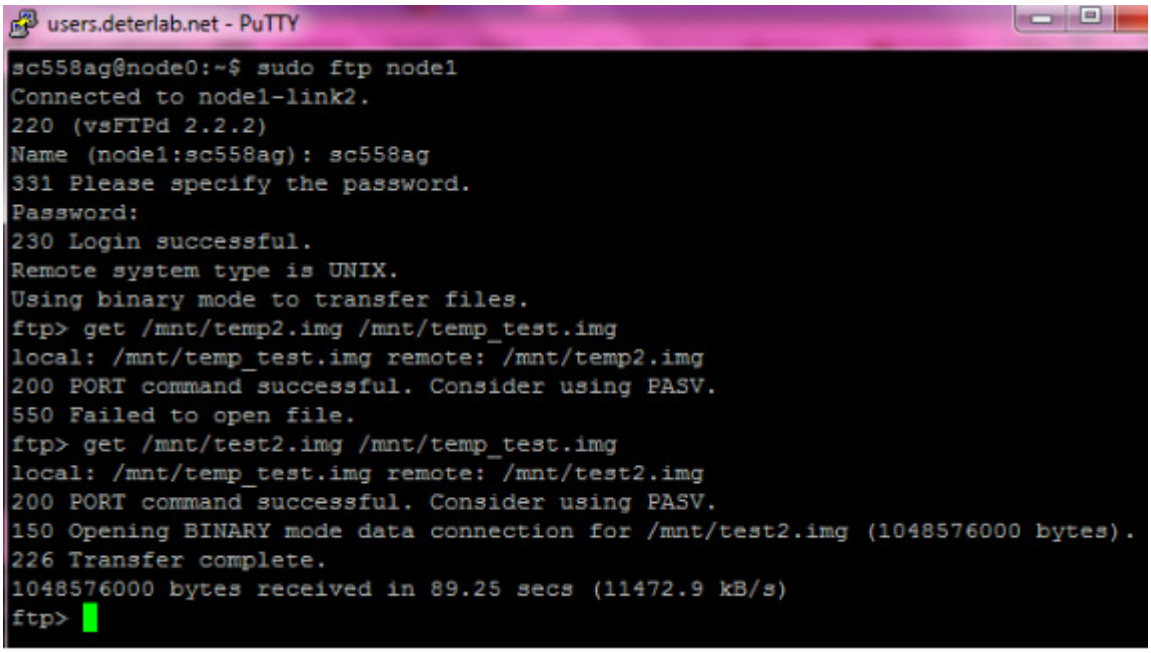
Change made to TCP parameters is as follows:

Changes made in /**proc/sys/net/core/** and /**proc/sys/net/ipv4**/.

We disabled timestamps to **reduce 12 byte overhead**.

We changed the default and maximum window size settings. And since the window sizes were still above 65536, we kept the window scaling enabled.

```
echo 256960 > /proc/sys/net/core/rmem_default
echo 256960 > /proc/sys/net/core/rmem_max
echo 256960 > /proc/sys/net/core/wmem_default
echo 256960 > /proc/sys/net/core/wmem_max
echo 0 > /proc/sys/net/ipv4/tcp_timestamps
echo 1 > /proc/sys/net/ipv4/tcp_sack
echo 1 > /proc/sys/net/ipv4/tcp_window_scaling
```



The throughput increased from **10984 KBps to 11472 KBps!**

Next:
The default traffic settings were removed and the delay was set to 20 ms RTT and a loss of 10% was added and we found that the throughput was reduced to 289.8KBps.

But just like the above settings we removed the 12 byte overhead for timestamps and changed the window size settings and We took multiple readings and we found that this did give **an increase in throughput by about 289 KBps to 341 KBps!**

Part II – modifying and compiling Kernel code

We tried compiling the source code on kernel in FBSD8 but we could not observe any change.

We will try working on this until the demo.