

Experiments on TCP Re-Ordering

March 27th 2017

Introduction

The Transmission Control Protocol (TCP) is very sensitive to the behavior of packets sent end-to-end. Variations in arrival time (“jitter”) coupled with other anomalous events that can either change the ordering, or completely lose, sections of the data cause problems that are challenging to recover from. These challenges are propagated back to the user in the form of “low observed throughput”, e.g. things will still make it end-to-end, but will be “slower” than if no such perturbation occurred at all.

The following document outlines experimentation done to quantify the impact of ordering on TCP. This document does not attempt to find the exact inflection point, but does offer some useful datum points that will help understand the impact of this behavior on TCP performance.

Experimental Environment

Testing was conducted on the ESnet network using two geographically separated hosts:

- lbl-diskpt1.es.net – Berkeley, California
- bnl-diskpt1.es.net – Upton, New York

Each system* is configured with 96GB of RAM, and contains 2 Intel Xeon processors (3.50GHz), with 6 cores each. The network interface is dual port 10Gbps Myricom card.

These hosts are separated by a round trip latency (RTT) of approximately 70ms, with a symmetrical 35ms of one-way latency (OWD). The hosts are connected via a 100Gbps backbone network, and have a bottleneck connection of 10Gbps at the host and local switching level. Measurements will be performed with tools from the perfSONAR Toolkit†, in particular: ping, traceroute, tracepath, owamp, iperf3, and bwctl.

Base Measurements

Latency measurements were consistent with expectations:

```
[rootjz@lbl-diskpt1 ~]# ping -c 20 bnl-diskpt1.es.net
PING bnl-diskpt1.es.net (198.124.238.150) 56(84) bytes of data:
64 bytes from bnl-diskpt1.es.net (198.124.238.150): icmp_seq=1 ttl=55 time=70.9 ms
```

* <https://fasterdata.es.net/science-dmz/DTN/reference-implementation/>

† <http://www.perfsonar.net/>

```
64 bytes from bnl-diskpt1.es.net (198.124.238.150): icmp_seq=2 ttl=55 time=70.9 ms
...
64 bytes from bnl-diskpt1.es.net (198.124.238.150): icmp_seq=20 ttl=55 time=70.9 ms
```

```
--- bnl-diskpt1.es.net ping statistics ---
20 packets transmitted, 20 received, 0% packet loss, time 19107ms
rtt min/avg/max/mdev = 70.906/70.928/70.964/0.367 ms
```

And:

```
[rootjz@lbl-diskpt1 ~]# owping bnl-diskpt1.es.net
Approximately 13.5 seconds until results available
```

```
--- owping statistics from [198.129.77.102]:8790 to [bnl-diskpt1.es.net]:9809 ---
SID:      c67cee96d95bac6297574d6a4f480dfa
first:    2015-07-23T11:09:40.028
last:     2015-07-23T11:09:51.487
100 sent, 0 lost (0.000%), 0 duplicates
one-way delay min/median/max = 35.4/35.5/35.5 ms, (err=0.298 ms)
one-way jitter = 0 ms (P95-P50)
Hops = 9 (consistently)
no reordering
```

```
--- owping statistics from [bnl-diskpt1.es.net]:9664 to [198.129.77.102]:8774 ---
SID:      c6814d66d95bac62a0b5430ffa5a94c0
first:    2015-07-23T11:09:40.011
last:     2015-07-23T11:09:50.497
100 sent, 0 lost (0.000%), 0 duplicates
one-way delay min/median/max = 35.5/35.6/35.6 ms, (err=0.298 ms)
one-way jitter = 0 ms (P95-P50)
Hops = 9 (consistently)
no reordering
```

Bandwidth testing indicated a clean network path that reached a steady state of more than 9Gbps over the 70ms round trip path. Minor variations in performance were observed, most likely due to the congestion profile of the local and WAN network.

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available
```

```
SENDER START
Connecting to host 198.124.238.150, port 5719
[ 15] local 198.129.77.102 port 54464 connected to 198.124.238.150 port 5719
[ ID] Interval      Transfer      Bandwidth      Retr  Cwnd
[ 15] 0.00-1.00 sec  2.04 MBytes  17.1 Mbits/sec  0    280 KBytes
[ 15] 1.00-2.00 sec  8.64 MBytes  72.4 Mbits/sec  0    935 KBytes
[ 15] 2.00-3.00 sec  18.8 MBytes  158 Mbits/sec  0    1.73 MBytes
[ 15] 3.00-4.00 sec  36.2 MBytes  304 Mbits/sec  0    3.40 MBytes
[ 15] 4.00-5.00 sec  70.0 MBytes  587 Mbits/sec  0    6.56 MBytes
[ 15] 5.00-6.00 sec  125 MBytes  1049 Mbits/sec  0    11.1 MBytes
[ 15] 6.00-7.00 sec  219 MBytes  1835 Mbits/sec  0    19.7 MBytes
[ 15] 7.00-8.00 sec  396 MBytes  3324 Mbits/sec  0    37.1 MBytes
[ 15] 8.00-9.00 sec  704 MBytes  5903 Mbits/sec  0    64.7 MBytes
[ 15] 9.00-10.00 sec 1.03 GBytes  8871 Mbits/sec  0    84.9 MBytes
[ 15] 10.00-11.00 sec 1.10 GBytes  9446 Mbits/sec  0    84.9 MBytes
[ 15] 11.00-12.00 sec 1.10 GBytes  9429 Mbits/sec  0    84.9 MBytes
[ 15] 12.00-13.00 sec 1.10 GBytes  9426 Mbits/sec  0    84.9 MBytes
[ 15] 13.00-14.00 sec 1.09 GBytes  9395 Mbits/sec  0    85.1 MBytes
[ 15] 14.00-15.00 sec 1.09 GBytes  9332 Mbits/sec  0    85.3 MBytes
[ 15] 15.00-16.00 sec 1.08 GBytes  9290 Mbits/sec  0    85.3 MBytes
[ 15] 16.00-17.00 sec 1.08 GBytes  9259 Mbits/sec  0    85.4 MBytes
[ 15] 17.00-18.00 sec 1.06 GBytes  9081 Mbits/sec  0    85.4 MBytes
[ 15] 18.00-19.00 sec 1.08 GBytes  9269 Mbits/sec  0    85.4 MBytes
[ 15] 19.00-20.00 sec 1.09 GBytes  9322 Mbits/sec  0    85.4 MBytes
[ 15] 20.00-21.00 sec 1.08 GBytes  9280 Mbits/sec  0    85.4 MBytes
[ 15] 21.00-22.00 sec 1.08 GBytes  9235 Mbits/sec  0    85.5 MBytes
[ 15] 22.00-23.00 sec 1.07 GBytes  9157 Mbits/sec  0    85.5 MBytes
```

```

[ 15] 23.00-24.00 sec 1.06 GBytes 9112 Mbits/sec 0 85.5 MBytes
[ 15] 24.00-25.00 sec 1.05 GBytes 8997 Mbits/sec 0 85.5 MBytes
[ 15] 25.00-26.00 sec 1.06 GBytes 9091 Mbits/sec 0 85.5 MBytes
[ 15] 26.00-27.00 sec 1.06 GBytes 9144 Mbits/sec 0 85.5 MBytes
[ 15] 27.00-28.00 sec 1.05 GBytes 9028 Mbits/sec 0 85.5 MBytes
[ 15] 28.00-29.00 sec 1.05 GBytes 9018 Mbits/sec 0 85.5 MBytes
[ 15] 29.00-30.00 sec 1.04 GBytes 8955 Mbits/sec 0 85.5 MBytes
[ 15] 30.00-31.00 sec 1.06 GBytes 9133 Mbits/sec 0 85.5 MBytes
[ 15] 31.00-32.00 sec 1.06 GBytes 9123 Mbits/sec 0 85.5 MBytes
[ 15] 32.00-33.00 sec 1.10 GBytes 9426 Mbits/sec 0 85.5 MBytes
[ 15] 33.00-34.00 sec 1.10 GBytes 9469 Mbits/sec 0 85.5 MBytes
[ 15] 34.00-35.00 sec 1.10 GBytes 9458 Mbits/sec 0 85.5 MBytes
[ 15] 35.00-36.00 sec 1.10 GBytes 9458 Mbits/sec 0 85.5 MBytes
[ 15] 36.00-37.00 sec 1.10 GBytes 9458 Mbits/sec 0 85.5 MBytes
[ 15] 37.00-38.00 sec 1.10 GBytes 9416 Mbits/sec 0 85.5 MBytes
[ 15] 38.00-39.00 sec 1.08 GBytes 9238 Mbits/sec 0 85.5 MBytes
[ 15] 39.00-40.00 sec 1.07 GBytes 9196 Mbits/sec 0 85.5 MBytes
[ 15] 40.00-41.00 sec 1.07 GBytes 9227 Mbits/sec 0 85.5 MBytes
[ 15] 41.00-42.00 sec 1.07 GBytes 9217 Mbits/sec 0 85.5 MBytes
[ 15] 42.00-43.00 sec 1.07 GBytes 9185 Mbits/sec 0 85.5 MBytes
[ 15] 43.00-44.00 sec 1.06 GBytes 9069 Mbits/sec 0 85.5 MBytes
[ 15] 44.00-45.00 sec 1.04 GBytes 8914 Mbits/sec 0 85.5 MBytes
[ 15] 45.00-46.00 sec 1.03 GBytes 8871 Mbits/sec 0 85.5 MBytes
[ 15] 46.00-47.00 sec 1.00 GBytes 8630 Mbits/sec 0 85.5 MBytes
[ 15] 47.00-48.00 sec 1.00 GBytes 8630 Mbits/sec 0 85.5 MBytes
[ 15] 48.00-49.00 sec 1020 MBytes 8556 Mbits/sec 0 85.5 MBytes
[ 15] 49.00-50.00 sec 1.00 GBytes 8629 Mbits/sec 0 85.5 MBytes
[ 15] 50.00-51.00 sec 1.00 GBytes 8599 Mbits/sec 0 85.5 MBytes
[ 15] 51.00-52.00 sec 1.05 GBytes 9049 Mbits/sec 0 85.5 MBytes
[ 15] 52.00-53.00 sec 1.10 GBytes 9448 Mbits/sec 0 85.5 MBytes
[ 15] 53.00-54.00 sec 1.10 GBytes 9448 Mbits/sec 0 85.5 MBytes
[ 15] 54.00-55.00 sec 1.10 GBytes 9446 Mbits/sec 0 85.5 MBytes
[ 15] 55.00-56.00 sec 1.10 GBytes 9449 Mbits/sec 0 85.5 MBytes
[ 15] 56.00-57.00 sec 1.10 GBytes 9427 Mbits/sec 0 85.5 MBytes
[ 15] 57.00-58.00 sec 1.10 GBytes 9416 Mbits/sec 0 85.5 MBytes
[ 15] 58.00-59.00 sec 1.10 GBytes 9437 Mbits/sec 0 85.5 MBytes
[ 15] 59.00-60.00 sec 1.10 GBytes 9448 Mbits/sec 0 85.5 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 56.1 GBytes 8031 Mbits/sec 0
[ 15] 0.00-60.00 sec 56.1 GBytes 8031 Mbits/sec
sender
receiver

```

iperf Done.

SENDER END

In summary, the path is understood to be a “production” network, and thus will have other traffic beyond the testing that is being performed. This should not impact the results of this experiment, and is provided as a notion of transparency.

The tests cases below will be achieved using the netem[‡] tool to provide “out of order” behavior for the packets in a single stream of TCP. This sort of testing will exercise the ability of the end hosts to handle out of order streams of data, and thus how well they can keep up with a misbehaving network. In situations like this, the availability of faster processors and ample memory is a must.

There are two different ways to specify packet reordering in netem. The first method, “gap” uses a fixed sequence and reorders every Nth packet. A simple usage of this is:

```
tc qdisc change dev eth0 root netem gap 5 delay 10ms
```

[‡] <http://www.linuxfoundation.org/collaborate/workgroups/networking/netem>

This causes every 5th (10th, 15th, ...) packet to go to be sent immediately and every other packet to be delayed by 10ms. This is predictable and useful for base protocol testing, e.g. when verifying a process like reassembly of streams.

The second form reorder of re-ordering is more like real life - it causes a certain percentage of the packets to get mis-ordered.

```
tc qdisc change dev eth0 root netem delay 10ms reorder 25% 50%
```

In this example, 25% of packets (with a correlation of 50%) will get sent immediately, others will be delayed by 10ms.

Netem also has the ability to re-order packets if the random delay values are out of order. The following will cause some reordering:

```
tc qdisc change dev eth0 root netem delay 100ms 75ms
```

If the first packet gets a random delay of 100ms (100ms base - 0ms jitter) and the second packet is sent 1ms later and gets a delay of 50ms (100ms base - 50ms jitter); the second packet will be sent first. This is because the queue discipline tifo inside netem, keeps packets in order by time to send.

There are caveats to the use of netem re-ordering:

- Mixing forms of reordering may lead to unexpected results
- For any method of reordering to work, some delay is necessary
- If the delay is less than the inter-packet arrival time then no reordering will be seen

Lastly, it is necessary for us to determine the “number” of packets that we may encounter coming through the interface for a given speed. Since our network is understood to be 10Gbps, the NIC will behave in binary (e.g. “off” and “on”) fashion. Thus[§]:

```
[10,000,000,000 b/s / (1,500 B * 8 b/B)] == 833 333 p/s
```

We will now exercise various situations of testing following the ‘real world’ example that adds a slight delay for a percentage of out of order packets.

Test Case 1

In the first test case, we want to see what happens when a tiny amount of re-ordering (1% of packets) are sent out of order with a very small delay difference

[§]http://www.cisco.com/web/about/security/intelligence/network_performance_metrics.html

(.2ms) for our long 70ms path. We can do this by applying some netem definitions on each of the hosts in our test:

```
LBL Side:
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .2ms reorder 99%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1
```

```
BNL Side:
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .2ms reorder 99%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1
```

Note that we apply this on each side, and only for the specific netmask of the other host. This will prevent any production traffic from experiencing the same impacts of this test. After applying these changes, we run a similar bandwidth measurement test:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

```
SENDER START
Connecting to host 198.124.238.150, port 5716
[ 15] local 198.129.77.102 port 60430 connected to 198.124.238.150 port 5716
[ ID] Interval      Transfer      Bandwidth     Retr  Cwnd
[ 15] 0.00-1.00    sec  2.03 MBytes  17.0 Mbits/sec  1    358 KBytes
[ 15] 1.00-2.00    sec  22.8 MBytes  191 Mbits/sec  0    3.19 MBytes
[ 15] 2.00-3.00    sec  88.8 MBytes  744 Mbits/sec  0    9.65 MBytes
[ 15] 3.00-4.00    sec  231 MBytes  1940 Mbits/sec  0    24.3 MBytes
[ 15] 4.00-5.00    sec  575 MBytes  4823 Mbits/sec  0    59.2 MBytes
[ 15] 5.00-6.00    sec  1.03 GBytes  8850 Mbits/sec  0    81.4 MBytes
[ 15] 6.00-7.00    sec  1.09 GBytes  9406 Mbits/sec  0    81.4 MBytes
[ 15] 7.00-8.00    sec  1018 MBytes  8534 Mbits/sec  0    81.1 MBytes
[ 15] 8.00-9.00    sec  1.03 GBytes  8851 Mbits/sec  0    81.2 MBytes
[ 15] 9.00-10.00   sec  1.08 GBytes  9280 Mbits/sec  0    81.6 MBytes
[ 15] 10.00-11.00  sec  1.09 GBytes  9343 Mbits/sec  0    81.3 MBytes
[ 15] 11.00-12.00  sec  1015 MBytes  8514 Mbits/sec  0    81.4 MBytes
[ 15] 12.00-13.00  sec  1020 MBytes  8556 Mbits/sec  0    81.4 MBytes
[ 15] 13.00-14.00  sec  1.09 GBytes  9343 Mbits/sec  0    81.2 MBytes
[ 15] 14.00-15.00  sec  1.07 GBytes  9165 Mbits/sec  0    81.5 MBytes
[ 15] 15.00-16.00  sec  1001 MBytes  8399 Mbits/sec  0    81.3 MBytes
[ 15] 16.00-17.00  sec  1.08 GBytes  9238 Mbits/sec  0    81.5 MBytes
[ 15] 17.00-18.00  sec  995 MBytes  8346 Mbits/sec  0    81.6 MBytes
[ 15] 18.00-19.00  sec  1.08 GBytes  9291 Mbits/sec  0    81.5 MBytes
[ 15] 19.00-20.00  sec  1.06 GBytes  9134 Mbits/sec  0    81.1 MBytes
[ 15] 20.00-21.00  sec  1.07 GBytes  9206 Mbits/sec  0    81.5 MBytes
[ 15] 21.00-22.00  sec  1018 MBytes  8535 Mbits/sec  0    81.3 MBytes
[ 15] 22.00-23.00  sec  1.09 GBytes  9343 Mbits/sec  0    81.3 MBytes
[ 15] 23.00-24.00  sec  1.08 GBytes  9259 Mbits/sec  0    81.2 MBytes
[ 15] 24.00-25.00  sec  1.03 GBytes  8839 Mbits/sec  0    81.5 MBytes
[ 15] 25.00-26.00  sec  1.02 GBytes  8756 Mbits/sec  0    81.3 MBytes
[ 15] 26.00-27.00  sec  1024 MBytes  8587 Mbits/sec  0    81.1 MBytes
[ 15] 27.00-28.00  sec  965 MBytes  8095 Mbits/sec  0    81.3 MBytes
[ 15] 28.00-29.00  sec  1.03 GBytes  8840 Mbits/sec  0    81.3 MBytes
[ 15] 29.00-30.00  sec  1.01 GBytes  8661 Mbits/sec  0    81.4 MBytes
[ 15] 30.00-31.00  sec  955 MBytes  8011 Mbits/sec  0    81.3 MBytes
[ 15] 31.00-32.00  sec  1.09 GBytes  9353 Mbits/sec  0    81.4 MBytes
[ 15] 32.00-33.00  sec  1.10 GBytes  9448 Mbits/sec  0    81.4 MBytes
[ 15] 33.00-34.00  sec  1.10 GBytes  9416 Mbits/sec  0    81.6 MBytes
[ 15] 34.00-35.00  sec  1.10 GBytes  9427 Mbits/sec  0    81.4 MBytes
[ 15] 35.00-36.00  sec  1.10 GBytes  9416 Mbits/sec  0    81.5 MBytes
```

```

[ 15] 36.00-37.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.5 MBytes
[ 15] 37.00-38.00 sec 1.10 GBytes 9426 Mbits/sec 0 81.2 MBytes
[ 15] 38.00-39.00 sec 1.09 GBytes 9354 Mbits/sec 0 81.3 MBytes
[ 15] 39.00-40.00 sec 1.08 GBytes 9311 Mbits/sec 0 81.5 MBytes
[ 15] 40.00-41.00 sec 1.07 GBytes 9175 Mbits/sec 0 81.5 MBytes
[ 15] 41.00-42.00 sec 1.05 GBytes 8996 Mbits/sec 0 81.3 MBytes
[ 15] 42.00-43.00 sec 1.03 GBytes 8808 Mbits/sec 0 81.5 MBytes
[ 15] 43.00-44.00 sec 1.04 GBytes 8964 Mbits/sec 0 81.0 MBytes
[ 15] 44.00-45.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.4 MBytes
[ 15] 45.00-46.00 sec 1.10 GBytes 9417 Mbits/sec 0 81.3 MBytes
[ 15] 46.00-47.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 47.00-48.00 sec 1.08 GBytes 9301 Mbits/sec 0 81.1 MBytes
[ 15] 48.00-49.00 sec 1.07 GBytes 9195 Mbits/sec 0 81.1 MBytes
[ 15] 49.00-50.00 sec 1.05 GBytes 9039 Mbits/sec 0 81.3 MBytes
[ 15] 50.00-51.00 sec 1.03 GBytes 8852 Mbits/sec 0 81.8 MBytes
[ 15] 51.00-52.00 sec 1.02 GBytes 8797 Mbits/sec 0 81.3 MBytes
[ 15] 52.00-53.00 sec 1.05 GBytes 9008 Mbits/sec 0 81.5 MBytes
[ 15] 53.00-54.00 sec 1.10 GBytes 9436 Mbits/sec 0 81.3 MBytes
[ 15] 54.00-55.00 sec 1.10 GBytes 9418 Mbits/sec 0 81.4 MBytes
[ 15] 55.00-56.00 sec 1.10 GBytes 9426 Mbits/sec 0 81.6 MBytes
[ 15] 56.00-57.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.3 MBytes
[ 15] 57.00-58.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.3 MBytes
[ 15] 58.00-59.00 sec 1.09 GBytes 9406 Mbits/sec 0 81.5 MBytes
[ 15] 59.00-60.00 sec 1.10 GBytes 9436 Mbits/sec 0 81.5 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15]  0.00-60.00 sec 59.0 GBytes 8450 Mbits/sec 1
[ 15]  0.00-60.00 sec 59.0 GBytes 8450 Mbits/sec

```

iperf Done.

SENDER END

The data from this test shows two things:

1. Despite the application of the rule, we did not see any retransmissions of data packets (except during slow start – which is normal). This leads us to conclude that the small delay of certain data packets was within acceptable tolerances to TCP: it did not trigger the Selective Acknowledgment (SACK) procedure in the algorithm when a packet has not been seen.
2. The congestion window fluctuates with very small differences during the transfer. We don't see this in the control case, and we could conclude that because the algorithm is going through small stalls waiting on a delayed packet (e.g. around only .2ms) it makes small adjustments to the size of the window to compensate for possible network congestion.

We can conclude that this small amount of re-ordering/delay does not impact our performance in a meaningful way.

Test Case 2

In the second test case, we want to see what happens when a tiny amount of re-ordering (1% of packets) are sent out of order with a slightly larger, yet still small, delay difference (.7ms) for our long 70ms path. We can do this by applying some different netem definitions on each of the hosts in our test:

```

LBL Side:
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .7ms reorder 99%
100%

```

```
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1
```

BNL Side:

```
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .7ms reorder 99%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1
```

After applying these changes, we run two similar bandwidth measurement tests:

Run 1:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available
```

SENDER START

Connecting to host 198.124.238.150, port 5754

[15] local 198.129.77.102 port 35230 connected to 198.124.238.150 port 5754

[ID]	Interval	Transfer	Bandwidth	Retr	Cwnd
[15]	0.00-1.00	sec 2.86 MBytes	24.0 Mbits/sec	0	577 KBytes
[15]	1.00-2.00	sec 26.9 MBytes	225 Mbits/sec	0	3.23 MBytes
[15]	2.00-3.00	sec 85.0 MBytes	713 Mbits/sec	0	8.76 MBytes
[15]	3.00-4.00	sec 209 MBytes	1751 Mbits/sec	0	22.0 MBytes
[15]	4.00-5.00	sec 541 MBytes	4540 Mbits/sec	0	59.1 MBytes
[15]	5.00-6.00	sec 1.02 GBytes	8777 Mbits/sec	0	81.5 MBytes
[15]	6.00-7.00	sec 1.09 GBytes	9374 Mbits/sec	0	81.4 MBytes
[15]	7.00-8.00	sec 1.09 GBytes	9374 Mbits/sec	0	81.5 MBytes
[15]	8.00-9.00	sec 1.09 GBytes	9385 Mbits/sec	0	81.2 MBytes
[15]	9.00-10.00	sec 1.09 GBytes	9375 Mbits/sec	0	81.3 MBytes
[15]	10.00-11.00	sec 1.09 GBytes	9384 Mbits/sec	0	81.3 MBytes
[15]	11.00-12.00	sec 878 MBytes	7362 Mbits/sec	17381	66.5 MBytes
[15]	12.00-13.00	sec 829 MBytes	6952 Mbits/sec	5527	66.6 MBytes
[15]	13.00-14.00	sec 928 MBytes	7780 Mbits/sec	0	67.3 MBytes
[15]	14.00-15.00	sec 939 MBytes	7875 Mbits/sec	0	68.0 MBytes
[15]	15.00-16.00	sec 952 MBytes	7990 Mbits/sec	0	68.8 MBytes
[15]	16.00-17.00	sec 956 MBytes	8022 Mbits/sec	0	69.7 MBytes
[15]	17.00-18.00	sec 969 MBytes	8126 Mbits/sec	0	70.6 MBytes
[15]	18.00-19.00	sec 981 MBytes	8231 Mbits/sec	0	71.6 MBytes
[15]	19.00-20.00	sec 998 MBytes	8368 Mbits/sec	0	72.6 MBytes
[15]	20.00-21.00	sec 1014 MBytes	8504 Mbits/sec	0	73.8 MBytes
[15]	21.00-22.00	sec 1.00 GBytes	8630 Mbits/sec	0	75.0 MBytes
[15]	22.00-23.00	sec 1.03 GBytes	8808 Mbits/sec	0	76.2 MBytes
[15]	23.00-24.00	sec 1.04 GBytes	8903 Mbits/sec	0	77.6 MBytes
[15]	24.00-25.00	sec 1.06 GBytes	9070 Mbits/sec	0	79.0 MBytes
[15]	25.00-26.00	sec 1.08 GBytes	9237 Mbits/sec	0	80.4 MBytes
[15]	26.00-27.00	sec 1.10 GBytes	9417 Mbits/sec	0	82.0 MBytes
[15]	27.00-28.00	sec 1.09 GBytes	9395 Mbits/sec	0	83.6 MBytes
[15]	28.00-29.00	sec 1.09 GBytes	9364 Mbits/sec	0	82.7 MBytes
[15]	29.00-30.00	sec 1.10 GBytes	9447 Mbits/sec	0	81.7 MBytes
[15]	30.00-31.00	sec 1.10 GBytes	9479 Mbits/sec	0	83.6 MBytes
[15]	31.00-32.00	sec 1.10 GBytes	9447 Mbits/sec	0	82.8 MBytes
[15]	32.00-33.00	sec 1.11 GBytes	9532 Mbits/sec	0	82.1 MBytes
[15]	33.00-34.00	sec 1.11 GBytes	9500 Mbits/sec	0	81.5 MBytes
[15]	34.00-35.00	sec 1.11 GBytes	9500 Mbits/sec	0	81.6 MBytes
[15]	35.00-36.00	sec 1.09 GBytes	9353 Mbits/sec	0	83.3 MBytes
[15]	36.00-37.00	sec 1.10 GBytes	9479 Mbits/sec	0	83.0 MBytes
[15]	37.00-38.00	sec 1.11 GBytes	9511 Mbits/sec	0	82.8 MBytes
[15]	38.00-39.00	sec 1.11 GBytes	9500 Mbits/sec	0	82.7 MBytes
[15]	39.00-40.00	sec 1.11 GBytes	9521 Mbits/sec	0	82.6 MBytes
[15]	40.00-41.00	sec 1.12 GBytes	9585 Mbits/sec	0	82.6 MBytes
[15]	41.00-42.00	sec 1.12 GBytes	9594 Mbits/sec	0	82.7 MBytes
[15]	42.00-43.00	sec 1.12 GBytes	9594 Mbits/sec	0	82.8 MBytes
[15]	43.00-44.00	sec 1.12 GBytes	9604 Mbits/sec	0	82.9 MBytes
[15]	44.00-45.00	sec 1.12 GBytes	9595 Mbits/sec	0	83.0 MBytes
[15]	45.00-46.00	sec 1.12 GBytes	9605 Mbits/sec	0	83.2 MBytes
[15]	46.00-47.00	sec 1.12 GBytes	9585 Mbits/sec	0	82.0 MBytes
[15]	47.00-48.00	sec 1.12 GBytes	9584 Mbits/sec	0	81.4 MBytes

```

[ 15] 48.00-49.00 sec 1.12 GBytes 9584 Mbits/sec 0 82.0 MBytes
[ 15] 49.00-50.00 sec 1.12 GBytes 9594 Mbits/sec 0 82.8 MBytes
[ 15] 50.00-51.00 sec 1.12 GBytes 9605 Mbits/sec 0 83.5 MBytes
[ 15] 51.00-52.00 sec 1.11 GBytes 9542 Mbits/sec 0 81.5 MBytes
[ 15] 52.00-53.00 sec 1.12 GBytes 9584 Mbits/sec 0 82.7 MBytes
[ 15] 53.00-54.00 sec 1.12 GBytes 9605 Mbits/sec 0 83.2 MBytes
[ 15] 54.00-55.00 sec 1.11 GBytes 9573 Mbits/sec 0 81.7 MBytes
[ 15] 55.00-56.00 sec 1.12 GBytes 9615 Mbits/sec 0 83.2 MBytes
[ 15] 56.00-57.00 sec 1.11 GBytes 9563 Mbits/sec 0 81.6 MBytes
[ 15] 57.00-58.00 sec 1.12 GBytes 9605 Mbits/sec 0 83.4 MBytes
[ 15] 58.00-59.00 sec 1.11 GBytes 9563 Mbits/sec 0 81.9 MBytes
[ 15] 59.00-60.00 sec 1.12 GBytes 9584 Mbits/sec 0 83.7 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 59.5 GBytes 8516 Mbits/sec 22908
[ 15] 0.00-60.00 sec 59.5 GBytes 8516 Mbits/sec

```

iperf Done.

SENDER END

Run 2:

```

[root@bz1-bl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available

```

SENDER START

```

Connecting to host 198.124.238.150, port 5756
[ 15] local 198.129.77.102 port 35068 connected to 198.124.238.150 port 5756
[ ID] Interval      Transfer      Bandwidth      Retr  Cwnd
[ 15] 0.00-1.00 sec 2.71 MBytes 22.8 Mbits/sec 1 524 KBytes
[ 15] 1.00-2.00 sec 21.1 MBytes 177 Mbits/sec 0 2.44 MBytes
[ 15] 2.00-3.00 sec 60.0 MBytes 503 Mbits/sec 0 5.90 MBytes
[ 15] 3.00-4.00 sec 135 MBytes 1132 Mbits/sec 0 13.9 MBytes
[ 15] 4.00-5.00 sec 349 MBytes 2926 Mbits/sec 0 36.9 MBytes
[ 15] 5.00-6.00 sec 825 MBytes 6920 Mbits/sec 0 81.4 MBytes
[ 15] 6.00-7.00 sec 1.09 GBytes 9375 Mbits/sec 0 81.4 MBytes
[ 15] 7.00-8.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.5 MBytes
[ 15] 8.00-9.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.2 MBytes
[ 15] 9.00-10.00 sec 1.09 GBytes 9383 Mbits/sec 0 81.4 MBytes
[ 15] 10.00-11.00 sec 1.09 GBytes 9365 Mbits/sec 0 81.2 MBytes
[ 15] 11.00-12.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 12.00-13.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 13.00-14.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 14.00-15.00 sec 1.09 GBytes 9385 Mbits/sec 0 81.3 MBytes
[ 15] 15.00-16.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.1 MBytes
[ 15] 16.00-17.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.4 MBytes
[ 15] 17.00-18.00 sec 1.09 GBytes 9384 Mbits/sec 0 81.3 MBytes
[ 15] 18.00-19.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.3 MBytes
[ 15] 19.00-20.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.1 MBytes
[ 15] 20.00-21.00 sec 1.09 GBytes 9384 Mbits/sec 0 81.3 MBytes
[ 15] 21.00-22.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 22.00-23.00 sec 1.09 GBytes 9365 Mbits/sec 0 81.0 MBytes
[ 15] 23.00-24.00 sec 1.09 GBytes 9373 Mbits/sec 0 81.4 MBytes
[ 15] 24.00-25.00 sec 1.09 GBytes 9365 Mbits/sec 0 81.4 MBytes
[ 15] 25.00-26.00 sec 1.09 GBytes 9373 Mbits/sec 0 81.3 MBytes
[ 15] 26.00-27.00 sec 1.09 GBytes 9365 Mbits/sec 0 81.5 MBytes
[ 15] 27.00-28.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.4 MBytes
[ 15] 28.00-29.00 sec 1.09 GBytes 9375 Mbits/sec 0 81.3 MBytes
[ 15] 29.00-30.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.5 MBytes
[ 15] 30.00-31.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.3 MBytes
[ 15] 31.00-32.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.3 MBytes
[ 15] 32.00-33.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 33.00-34.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.4 MBytes
[ 15] 34.00-35.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 35.00-36.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 36.00-37.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.1 MBytes
[ 15] 37.00-38.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.5 MBytes
[ 15] 38.00-39.00 sec 976 MBytes 8190 Mbits/sec 4588 54.7 MBytes
[ 15] 39.00-40.00 sec 1.00 GBytes 8619 Mbits/sec 0 81.2 MBytes
[ 15] 40.00-41.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.3 MBytes
[ 15] 41.00-42.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.1 MBytes

```



```

[ 15] 42.00-43.00 sec 1.09 GBytes 9386 Mbits/sec 0 81.2 MBytes
[ 15] 43.00-44.00 sec 1.09 GBytes 9364 Mbits/sec 0 81.1 MBytes
[ 15] 44.00-45.00 sec 1.09 GBytes 9384 Mbits/sec 0 81.3 MBytes
[ 15] 45.00-46.00 sec 1.09 GBytes 9375 Mbits/sec 0 81.4 MBytes
[ 15] 46.00-47.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 47.00-48.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.2 MBytes
[ 15] 48.00-49.00 sec 1.09 GBytes 9385 Mbits/sec 0 81.4 MBytes
[ 15] 49.00-50.00 sec 1.09 GBytes 9363 Mbits/sec 0 81.4 MBytes
[ 15] 50.00-51.00 sec 1.09 GBytes 9385 Mbits/sec 0 81.3 MBytes
[ 15] 51.00-52.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 52.00-53.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.2 MBytes
[ 15] 53.00-54.00 sec 1.09 GBytes 9373 Mbits/sec 0 81.1 MBytes
[ 15] 54.00-55.00 sec 1.09 GBytes 9375 Mbits/sec 0 81.3 MBytes
[ 15] 55.00-56.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.3 MBytes
[ 15] 56.00-57.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.1 MBytes
[ 15] 57.00-58.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.4 MBytes
[ 15] 58.00-59.00 sec 1.09 GBytes 9385 Mbits/sec 0 81.3 MBytes
[ 15] 59.00-60.00 sec 1.09 GBytes 9374 Mbits/sec 0 81.3 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15]  0.00-60.00 sec 60.1 GBytes 8598 Mbits/sec 4589
[ 15]  0.00-60.00 sec 60.1 GBytes 8598 Mbits/sec
sender
receiver

```

iperf Done.

SENDER END

The data from this test shows three things:

1. In both cases, the re-ordering/delay of some portion of the data stream most likely triggered the SACK code of TCP to be enabled. Internally this probably played out as:
 - a. The window was stalled waiting for a single packet, but more packets that would have occurred “after” the one we are waiting on kept arriving
 - b. The receiver will send what is called a “duplicate acknowledgment” to the sender for the last ‘in order’ packet it has received. For example, if you have seen numbers 1-5, and are missing number 6, the duplicate acknowledgement for number 5 will be sent each time you still don’t receive number 6.
 - c. Upon receiving the duplicate acknowledgement(s), the sender will assume the packet is loss and try to re-send it again
 - d. Upon receipt of 3 duplicate acknowledgements, the sender will re-send the entire window of packets – trying to recover from whatever has happened in the network (e.g. fearing the worse – that packet loss has happened).
 - e. Alternatively if the “timer”, a notion built into the TCP algorithm that records the length of time that must be spent waiting for a packet before we assume it is lost, could be triggered causing a SACK/Window transmission
 - f. The burst of several thousand retransmissions is indicative of either case d) or e): that is the resending of a window of data several MB in size. It has the desired impact: we recover from the out of order/delay and see the speed return to an acceptable value.

2. Because the probability of this event happening is so low (1% of the time), it does not occur often and thus allows us to recover. If this percentage were higher, we may see this event occur more often.
3. The difference in delay (.7ms) caused a slightly different behavior than in the previous experiment (.2ms). This is revealing more about the interworkings of the TCP algorithm. It may be the case that the “timer” that records the period of time between when a packet should have arrived, and when it is viewed as lost, is starting to be triggered by this percentage (e.g. ~1%) of the RTT delay.

We can conclude that this small amount of re-ordering/delay does impact our performance slightly, but recovery is possible.

Test Case 3

In the third test case, we want to see what happens when a still tiny amount of re-ordering (1% of packets) are sent out of order with a delay that is larger than the previous two (.8ms) for our long 70ms path. Here are the changes to netem:

LBL Side:

```
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .8ms reorder 99% 100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst 198.124.238.150/32 flowid 1:1
```

BNL Side:

```
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .8ms reorder 99% 100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst 198.129.77.102/32 flowid 1:1
```

After applying these changes, we run a similar bandwidth measurement test:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available
```

SENDER START

```
Connecting to host 198.124.238.150, port 5781
[ 15] local 198.129.77.102 port 33040 connected to 198.124.238.150 port 5781
[ ID] Interval      Transfer    Bandwidth  Retr  Cwnd
[ 15] 0.00-1.00    sec  1.99 MBytes  16.7 Mbits/sec    1   367 KBytes
[ 15] 1.00-2.00    sec  21.1 MBytes  177 Mbits/sec    0   2.65 MBytes
[ 15] 2.00-3.00    sec  66.2 MBytes  556 Mbits/sec    0   6.64 MBytes
[ 15] 3.00-4.00    sec  161 MBytes  1353 Mbits/sec    0   16.5 MBytes
[ 15] 4.00-5.00    sec  226 MBytes  1898 Mbits/sec   639   9.26 MBytes
[ 15] 5.00-6.00    sec  200 MBytes  1678 Mbits/sec    0   23.6 MBytes
[ 15] 6.00-7.00    sec  549 MBytes  4603 Mbits/sec    0   56.2 MBytes
[ 15] 7.00-8.00    sec  1.00 GBytes  8596 Mbits/sec    0   81.2 MBytes
[ 15] 8.00-9.00    sec  939 MBytes  7877 Mbits/sec  18246   83.7 MBytes
[ 15] 9.00-10.00   sec  850 MBytes  7130 Mbits/sec  30612   83.5 MBytes
[ 15] 10.00-11.00  sec  868 MBytes  7278 Mbits/sec  29056   83.2 MBytes
[ 15] 11.00-12.00  sec  1024 MBytes  8588 Mbits/sec  8166   81.3 MBytes
[ 15] 12.00-13.00  sec  1.09 GBytes  9364 Mbits/sec    0   81.5 MBytes
[ 15] 13.00-14.00  sec  916 MBytes  7686 Mbits/sec  16472   74.0 MBytes
[ 15] 14.00-15.00  sec  1.04 GBytes  8902 Mbits/sec  3962   81.1 MBytes
[ 15] 15.00-16.00  sec  996 MBytes  8357 Mbits/sec  11998   81.4 MBytes
[ 15] 16.00-17.00  sec  1.04 GBytes  8965 Mbits/sec  4401   83.1 MBytes
```

```

[ 15] 17.00-18.00 sec 1.05 GBytes 8997 Mbits/sec 8519 63.3 MBytes
[ 15] 18.00-19.00 sec 950 MBytes 7969 Mbits/sec 5547 54.7 MBytes
[ 15] 19.00-20.00 sec 1.00 GBytes 8598 Mbits/sec 52 81.4 MBytes
[ 15] 20.00-21.00 sec 1.06 GBytes 9144 Mbits/sec 3467 71.3 MBytes
[ 15] 21.00-22.00 sec 1.02 GBytes 8735 Mbits/sec 7967 83.2 MBytes
[ 15] 22.00-23.00 sec 868 MBytes 7277 Mbits/sec 29253 83.7 MBytes
[ 15] 23.00-24.00 sec 864 MBytes 7246 Mbits/sec 29888 83.2 MBytes
[ 15] 24.00-25.00 sec 1.04 GBytes 8902 Mbits/sec 6448 64.5 MBytes
[ 15] 25.00-26.00 sec 995 MBytes 8346 Mbits/sec 2806 81.4 MBytes
[ 15] 26.00-27.00 sec 1.04 GBytes 8976 Mbits/sec 3870 81.3 MBytes
[ 15] 27.00-28.00 sec 1.05 GBytes 8985 Mbits/sec 4115 81.3 MBytes
[ 15] 28.00-29.00 sec 1.09 GBytes 9354 Mbits/sec 0 81.2 MBytes
[ 15] 29.00-30.00 sec 1.09 GBytes 9354 Mbits/sec 0 81.4 MBytes
[ 15] 30.00-31.00 sec 980 MBytes 8221 Mbits/sec 8296 81.5 MBytes
[ 15] 31.00-32.00 sec 979 MBytes 8211 Mbits/sec 17762 83.7 MBytes
[ 15] 32.00-33.00 sec 865 MBytes 7256 Mbits/sec 29236 83.4 MBytes
[ 15] 33.00-34.00 sec 928 MBytes 7780 Mbits/sec 18165 81.3 MBytes
[ 15] 34.00-35.00 sec 1.05 GBytes 8997 Mbits/sec 6359 69.5 MBytes
[ 15] 35.00-36.00 sec 985 MBytes 8263 Mbits/sec 20382 51.0 MBytes
[ 15] 36.00-37.00 sec 919 MBytes 7707 Mbits/sec 7363 78.3 MBytes
[ 15] 37.00-38.00 sec 1.04 GBytes 8944 Mbits/sec 4174 81.1 MBytes
[ 15] 38.00-39.00 sec 1.01 GBytes 8651 Mbits/sec 12664 83.9 MBytes
[ 15] 39.00-40.00 sec 870 MBytes 7299 Mbits/sec 28898 84.2 MBytes
[ 15] 40.00-41.00 sec 868 MBytes 7277 Mbits/sec 28846 65.4 MBytes
[ 15] 41.00-42.00 sec 1.00 GBytes 8630 Mbits/sec 4380 82.3 MBytes
[ 15] 42.00-43.00 sec 1.05 GBytes 9049 Mbits/sec 4569 72.9 MBytes
[ 15] 43.00-44.00 sec 961 MBytes 8064 Mbits/sec 17368 83.4 MBytes
[ 15] 44.00-45.00 sec 860 MBytes 7214 Mbits/sec 28767 83.3 MBytes
[ 15] 45.00-46.00 sec 868 MBytes 7277 Mbits/sec 29091 63.1 MBytes
[ 15] 46.00-47.00 sec 862 MBytes 7234 Mbits/sec 29858 65.9 MBytes
[ 15] 47.00-48.00 sec 889 MBytes 7456 Mbits/sec 26992 73.4 MBytes
[ 15] 48.00-49.00 sec 1.04 GBytes 8902 Mbits/sec 3729 75.9 MBytes
[ 15] 49.00-50.00 sec 1.04 GBytes 8934 Mbits/sec 3710 81.2 MBytes
[ 15] 50.00-51.00 sec 982 MBytes 8242 Mbits/sec 10743 51.8 MBytes
[ 15] 51.00-52.00 sec 981 MBytes 8231 Mbits/sec 0 81.5 MBytes
[ 15] 52.00-53.00 sec 992 MBytes 8326 Mbits/sec 12120 81.4 MBytes
[ 15] 53.00-54.00 sec 990 MBytes 8305 Mbits/sec 9450 65.2 MBytes
[ 15] 54.00-55.00 sec 1.05 GBytes 8986 Mbits/sec 0 81.5 MBytes
[ 15] 55.00-56.00 sec 908 MBytes 7613 Mbits/sec 27412 83.6 MBytes
[ 15] 56.00-57.00 sec 861 MBytes 7225 Mbits/sec 29992 84.3 MBytes
[ 15] 57.00-58.00 sec 861 MBytes 7225 Mbits/sec 28895 84.0 MBytes
[ 15] 58.00-59.00 sec 864 MBytes 7246 Mbits/sec 29373 83.9 MBytes
[ 15] 59.00-60.00 sec 864 MBytes 7246 Mbits/sec 29299 83.7 MBytes
-----
[ ID] Interval Transfer Bandwidth Retr sender
[ 15] 0.00-60.00 sec 51.8 GBytes 7415 Mbits/sec 733378 receiver
[ 15] 0.00-60.00 sec 51.8 GBytes 7414 Mbits/sec

```

iperf Done.

SENDER END

In this test case, we start to see some very chaotic behavior. The data from this test shows:

1. As in test case #2, we are seeing numerous retransmissions occurring
2. The retransmissions may be causing a “snowball” behavior, e.g. the initial trigger of a slightly delayed packet causes the eventual re-sending of the data again. There is inevitably duplication that occurs as a result of this, would could trigger more SACK behavior
3. The fact that the delay is just beyond 1% of the total RTT (.8ms on a 70ms path) seems to indicate that the “timer” that records when a packet is lost or not is being triggered more regularly.
4. Despite all the extra “work” that is performed, note that our interval throughput (and indeed, the overall average), is still very high. Achieving

nearly 7.5 Gbps over the 60 seconds is still very good, and most users may never notice that there is a “problem” with the network. This is still far less than our ideal case of more than 9Gbps, however.

We can conclude this small amount of transmission, with a delay that is just beyond 1% of the RTT of the end-to-end path, causes significant behavior change in TCP. Due to the capabilities of the host, a fast recovery is possible with minimal impact on the overall throughput.

Test Case 4

In the third test case, we depart slightly from our previous attempts. We will try to reorder more of the packets initially – up to 25% of them and exercise a very small delay interval (.2ms) for our long 70ms path. We can do this by applying some netem definitions on each of the hosts in our test:

LBL Side:

```
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .2ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1
```

BNL Side:

```
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .2ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1
```

After applying these changes, we run two similar bandwidth measurement tests:

Run 1:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

```
Connecting to host 198.124.238.150, port 5758
[ 15] local 198.129.77.102 port 35779 connected to 198.124.238.150 port 5758
[ ID] Interval      Transfer      Bandwidth      Retr  Cwnd
[ 15] 0.00-1.00    sec  3.04 MBytes  25.5 Mbits/sec    0   533 KBytes
[ 15] 1.00-2.00    sec  25.2 MBytes  211 Mbits/sec    0  3.54 MBytes
[ 15] 2.00-3.00    sec  119 MBytes  996 Mbits/sec    0  14.7 MBytes
[ 15] 3.00-4.00    sec  432 MBytes  3628 Mbits/sec    0  50.3 MBytes
[ 15] 4.00-5.00    sec  625 MBytes  5243 Mbits/sec  4870  42.4 MBytes
[ 15] 5.00-6.00    sec  592 MBytes  4970 Mbits/sec    0  42.4 MBytes
[ 15] 6.00-7.00    sec  595 MBytes  4991 Mbits/sec    0  42.5 MBytes
[ 15] 7.00-8.00    sec  596 MBytes  5002 Mbits/sec    0  42.7 MBytes
[ 15] 8.00-9.00    sec  600 MBytes  5033 Mbits/sec    0  43.1 MBytes
[ 15] 9.00-10.00   sec  602 MBytes  5054 Mbits/sec    0  43.7 MBytes
[ 15] 10.00-11.00  sec  582 MBytes  4886 Mbits/sec    0  44.5 MBytes
[ 15] 11.00-12.00  sec  584 MBytes  4897 Mbits/sec    0  45.3 MBytes
[ 15] 12.00-13.00  sec  591 MBytes  4960 Mbits/sec    0  46.3 MBytes
[ 15] 13.00-14.00  sec  601 MBytes  5044 Mbits/sec    0  47.5 MBytes
[ 15] 14.00-15.00  sec  611 MBytes  5128 Mbits/sec    0  48.8 MBytes
[ 15] 15.00-16.00  sec  621 MBytes  5211 Mbits/sec    0  50.1 MBytes
[ 15] 16.00-17.00  sec  634 MBytes  5316 Mbits/sec    0  51.7 MBytes
[ 15] 17.00-18.00  sec  645 MBytes  5411 Mbits/sec    0  53.2 MBytes
[ 15] 18.00-19.00  sec  659 MBytes  5526 Mbits/sec    0  54.9 MBytes
```

[15]	19.00-20.00	sec	670 MBytes	5620 Mbits/sec	0	56.7 MBytes
[15]	20.00-21.00	sec	685 MBytes	5746 Mbits/sec	0	58.4 MBytes
[15]	21.00-22.00	sec	695 MBytes	5830 Mbits/sec	0	60.4 MBytes
[15]	22.00-23.00	sec	710 MBytes	5956 Mbits/sec	0	62.3 MBytes
[15]	23.00-24.00	sec	721 MBytes	6050 Mbits/sec	0	64.4 MBytes
[15]	24.00-25.00	sec	735 MBytes	6166 Mbits/sec	0	66.4 MBytes
[15]	25.00-26.00	sec	746 MBytes	6260 Mbits/sec	0	68.5 MBytes
[15]	26.00-27.00	sec	758 MBytes	6354 Mbits/sec	0	70.8 MBytes
[15]	27.00-28.00	sec	770 MBytes	6459 Mbits/sec	0	72.9 MBytes
[15]	28.00-29.00	sec	965 MBytes	8095 Mbits/sec	0	75.3 MBytes
[15]	29.00-30.00	sec	1.04 GBytes	8955 Mbits/sec	0	77.7 MBytes
[15]	30.00-31.00	sec	1.08 GBytes	9238 Mbits/sec	0	80.1 MBytes
[15]	31.00-32.00	sec	1.11 GBytes	9521 Mbits/sec	0	82.5 MBytes
[15]	32.00-33.00	sec	1.12 GBytes	9626 Mbits/sec	0	81.7 MBytes
[15]	33.00-34.00	sec	1.12 GBytes	9626 Mbits/sec	0	81.4 MBytes
[15]	34.00-35.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.9 MBytes
[15]	35.00-36.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.8 MBytes
[15]	36.00-37.00	sec	1.12 GBytes	9626 Mbits/sec	0	83.7 MBytes
[15]	37.00-38.00	sec	1.12 GBytes	9615 Mbits/sec	0	83.4 MBytes
[15]	38.00-39.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.6 MBytes
[15]	39.00-40.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.8 MBytes
[15]	40.00-41.00	sec	1.12 GBytes	9626 Mbits/sec	0	84.1 MBytes
[15]	41.00-42.00	sec	1.12 GBytes	9605 Mbits/sec	0	81.3 MBytes
[15]	42.00-43.00	sec	1.12 GBytes	9636 Mbits/sec	0	81.9 MBytes
[15]	43.00-44.00	sec	1.12 GBytes	9626 Mbits/sec	0	82.5 MBytes
[15]	44.00-45.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.3 MBytes
[15]	45.00-46.00	sec	1.12 GBytes	9626 Mbits/sec	0	84.1 MBytes
[15]	46.00-47.00	sec	1.12 GBytes	9605 Mbits/sec	0	81.7 MBytes
[15]	47.00-48.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.1 MBytes
[15]	48.00-49.00	sec	1.12 GBytes	9605 Mbits/sec	0	81.9 MBytes
[15]	49.00-50.00	sec	1.12 GBytes	9605 Mbits/sec	0	82.0 MBytes
[15]	50.00-51.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.6 MBytes
[15]	51.00-52.00	sec	1.12 GBytes	9594 Mbits/sec	0	81.6 MBytes
[15]	52.00-53.00	sec	1.12 GBytes	9594 Mbits/sec	0	83.5 MBytes
[15]	53.00-54.00	sec	1.11 GBytes	9553 Mbits/sec	0	81.7 MBytes
[15]	54.00-55.00	sec	1.12 GBytes	9636 Mbits/sec	0	83.7 MBytes
[15]	55.00-56.00	sec	1.12 GBytes	9605 Mbits/sec	0	82.2 MBytes
[15]	56.00-57.00	sec	1.12 GBytes	9594 Mbits/sec	0	81.4 MBytes
[15]	57.00-58.00	sec	1.12 GBytes	9615 Mbits/sec	0	83.6 MBytes
[15]	58.00-59.00	sec	1.12 GBytes	9606 Mbits/sec	0	82.8 MBytes
[15]	59.00-60.00	sec	1.11 GBytes	9573 Mbits/sec	0	82.3 MBytes

[ID]	Interval	Transfer	Bandwidth	Retr	sender	receiver
[15]	0.00-60.00	sec 51.3 GBytes	7351 Mbits/sec	4870		
[15]	0.00-60.00	sec 51.3 GBytes	7350 Mbits/sec			

iperf Done.

SENDER END

Run 2:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

Connecting to host 198.124.238.150, port 5760

[15] local 198.129.77.102 port 56206 connected to 198.124.238.150 port 5760

[ID]	Interval	Transfer	Bandwidth	Retr	Cwnd
[15]	0.00-1.00	sec 1.44 MBytes	12.1 Mbits/sec	0	236 KBytes
[15]	1.00-2.00	sec 12.8 MBytes	108 Mbits/sec	0	1.95 MBytes
[15]	2.00-3.00	sec 65.0 MBytes	545 Mbits/sec	0	7.42 MBytes
[15]	3.00-4.00	sec 196 MBytes	1646 Mbits/sec	0	21.2 MBytes
[15]	4.00-5.00	sec 504 MBytes	4226 Mbits/sec	0	51.7 MBytes
[15]	5.00-6.00	sec 1002 MBytes	8410 Mbits/sec	0	81.5 MBytes
[15]	6.00-7.00	sec 1.10 GBytes	9437 Mbits/sec	0	81.2 MBytes
[15]	7.00-8.00	sec 1.10 GBytes	9448 Mbits/sec	0	81.3 MBytes
[15]	8.00-9.00	sec 1.10 GBytes	9437 Mbits/sec	0	81.5 MBytes
[15]	9.00-10.00	sec 1.10 GBytes	9427 Mbits/sec	0	81.4 MBytes
[15]	10.00-11.00	sec 1.10 GBytes	9447 Mbits/sec	0	81.1 MBytes
[15]	11.00-12.00	sec 1.10 GBytes	9428 Mbits/sec	0	81.5 MBytes
[15]	12.00-13.00	sec 1.10 GBytes	9448 Mbits/sec	0	81.5 MBytes

```

[ 15] 13.00-14.00 sec 1.10 GBytes 9435 Mbits/sec 0 81.4 MBytes
[ 15] 14.00-15.00 sec 1.10 GBytes 9439 Mbits/sec 0 81.4 MBytes
[ 15] 15.00-16.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.5 MBytes
[ 15] 16.00-17.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.4 MBytes
[ 15] 17.00-18.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.2 MBytes
[ 15] 18.00-19.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.4 MBytes
[ 15] 19.00-20.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.2 MBytes
[ 15] 20.00-21.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.2 MBytes
[ 15] 21.00-22.00 sec 1.10 GBytes 9448 Mbits/sec 0 81.5 MBytes
[ 15] 22.00-23.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.5 MBytes
[ 15] 23.00-24.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.3 MBytes
[ 15] 24.00-25.00 sec 1.10 GBytes 9448 Mbits/sec 0 81.2 MBytes
[ 15] 25.00-26.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.6 MBytes
[ 15] 26.00-27.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.6 MBytes
[ 15] 27.00-28.00 sec 1.10 GBytes 9435 Mbits/sec 0 81.5 MBytes
[ 15] 28.00-29.00 sec 1.10 GBytes 9439 Mbits/sec 0 81.4 MBytes
[ 15] 29.00-30.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.4 MBytes
[ 15] 30.00-31.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.1 MBytes
[ 15] 31.00-32.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.5 MBytes
[ 15] 32.00-33.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.4 MBytes
[ 15] 33.00-34.00 sec 1.10 GBytes 9437 Mbits/sec 0 81.2 MBytes
[ 15] 34.00-35.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.5 MBytes
[ 15] 35.00-36.00 sec 1.10 GBytes 9447 Mbits/sec 0 81.2 MBytes
[ 15] 36.00-37.00 sec 1.10 GBytes 9427 Mbits/sec 0 81.0 MBytes
[ 15] 37.00-38.00 sec 1.09 GBytes 9406 Mbits/sec 0 81.4 MBytes
[ 15] 38.00-39.00 sec 1.04 GBytes 8955 Mbits/sec 0 81.4 MBytes
[ 15] 39.00-40.00 sec 902 MBytes 7571 Mbits/sec 0 81.3 MBytes
[ 15] 40.00-41.00 sec 851 MBytes 7141 Mbits/sec 1712 60.4 MBytes
[ 15] 41.00-42.00 sec 724 MBytes 6071 Mbits/sec 17394 57.4 MBytes
[ 15] 42.00-43.00 sec 821 MBytes 6889 Mbits/sec 0 60.4 MBytes
[ 15] 43.00-44.00 sec 865 MBytes 7256 Mbits/sec 0 63.6 MBytes
[ 15] 44.00-45.00 sec 911 MBytes 7644 Mbits/sec 0 67.2 MBytes
[ 15] 45.00-46.00 sec 961 MBytes 8064 Mbits/sec 0 70.8 MBytes
[ 15] 46.00-47.00 sec 1012 MBytes 8493 Mbits/sec 0 74.4 MBytes
[ 15] 47.00-48.00 sec 1.04 GBytes 8913 Mbits/sec 0 77.9 MBytes
[ 15] 48.00-49.00 sec 1.08 GBytes 9301 Mbits/sec 0 81.4 MBytes
[ 15] 49.00-50.00 sec 1.11 GBytes 9573 Mbits/sec 0 81.4 MBytes
[ 15] 50.00-51.00 sec 1.12 GBytes 9636 Mbits/sec 0 81.8 MBytes
[ 15] 51.00-52.00 sec 1.12 GBytes 9626 Mbits/sec 0 82.1 MBytes
[ 15] 52.00-53.00 sec 1.12 GBytes 9626 Mbits/sec 0 82.3 MBytes
[ 15] 53.00-54.00 sec 1.12 GBytes 9637 Mbits/sec 0 82.7 MBytes
[ 15] 54.00-55.00 sec 1.12 GBytes 9636 Mbits/sec 0 83.1 MBytes
[ 15] 55.00-56.00 sec 1.12 GBytes 9636 Mbits/sec 0 83.6 MBytes
[ 15] 56.00-57.00 sec 1.12 GBytes 9595 Mbits/sec 0 81.4 MBytes
[ 15] 57.00-58.00 sec 1.12 GBytes 9605 Mbits/sec 0 82.1 MBytes
[ 15] 58.00-59.00 sec 1.12 GBytes 9626 Mbits/sec 0 83.0 MBytes
[ 15] 59.00-60.00 sec 1.12 GBytes 9636 Mbits/sec 0 84.0 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 59.3 GBytes 8484 Mbits/sec 19106
[ 15] 0.00-60.00 sec 59.3 GBytes 8484 Mbits/sec
sender
receiver

```

iperf Done.

SENDER END

The data from this test shows:

1. The higher probability (25% of the time) that a packet is delayed does cause our congestion window to fluctuate a bit, as it adjusts to the perceived packet loss.
2. The blips of retransmissions mimic what we saw in previous cases (e.g. #2) and could be explained the same way. We most likely get 'unlucky' and have a string of delayed out of order packets that pause our window long enough to trigger a re-sending or SACK compensation. We do recover and maintain a high speed after each of these.

We can conclude that this higher probability of an out of order packet is causing some problems for TCP. It is working harder, but since the delay is minimal (.2ms on a 70ms RTT path) it can recover faster and get back to a higher bandwidth measurement.

Test Case 5

Building on test case #4, we do the same thing with our reordering percentage (25%), but increase the delay to .7ms for our 70ms path:

LBL Side:

```
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .7ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1
```

BNL Side:

```
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .7ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1
```

After applying these changes, we run two similar bandwidth measurement tests:

Run 1:

```
[root@l1bl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

Connecting to host 198.124.238.150, port 5762

[15] local 198.129.77.102 port 51021 connected to 198.124.238.150 port 5762

[ID]	Interval		Transfer	Bandwidth	Retr	Cwnd
[15]	0.00-1.00	sec	2.07 MBytes	17.3 Mbits/sec	0	358 KBytes
[15]	1.00-2.00	sec	18.2 MBytes	153 Mbits/sec	0	2.24 MBytes
[15]	2.00-3.00	sec	61.2 MBytes	514 Mbits/sec	0	6.43 MBytes
[15]	3.00-4.00	sec	158 MBytes	1321 Mbits/sec	0	16.0 MBytes
[15]	4.00-5.00	sec	404 MBytes	3387 Mbits/sec	0	43.2 MBytes
[15]	5.00-6.00	sec	915 MBytes	7676 Mbits/sec	0	81.5 MBytes
[15]	6.00-7.00	sec	1.09 GBytes	9374 Mbits/sec	0	81.2 MBytes
[15]	7.00-8.00	sec	1.09 GBytes	9374 Mbits/sec	0	81.3 MBytes
[15]	8.00-9.00	sec	1.09 GBytes	9383 Mbits/sec	0	81.3 MBytes
[15]	9.00-10.00	sec	1.09 GBytes	9365 Mbits/sec	0	81.4 MBytes
[15]	10.00-11.00	sec	1.09 GBytes	9375 Mbits/sec	0	81.3 MBytes
[15]	11.00-12.00	sec	1.09 GBytes	9384 Mbits/sec	0	81.3 MBytes
[15]	12.00-13.00	sec	1.09 GBytes	9374 Mbits/sec	0	81.4 MBytes
[15]	13.00-14.00	sec	1.09 GBytes	9374 Mbits/sec	0	81.2 MBytes
[15]	14.00-15.00	sec	1.09 GBytes	9375 Mbits/sec	0	81.2 MBytes
[15]	15.00-16.00	sec	1.09 GBytes	9385 Mbits/sec	0	81.1 MBytes
[15]	16.00-17.00	sec	891 MBytes	7476 Mbits/sec	15641	63.8 MBytes
[15]	17.00-18.00	sec	921 MBytes	7728 Mbits/sec	0	67.2 MBytes
[15]	18.00-19.00	sec	942 MBytes	7907 Mbits/sec	0	68.0 MBytes
[15]	19.00-20.00	sec	942 MBytes	7906 Mbits/sec	0	68.8 MBytes
[15]	20.00-21.00	sec	962 MBytes	8074 Mbits/sec	0	69.7 MBytes
[15]	21.00-22.00	sec	968 MBytes	8116 Mbits/sec	0	70.6 MBytes
[15]	22.00-23.00	sec	981 MBytes	8231 Mbits/sec	0	71.6 MBytes
[15]	23.00-24.00	sec	996 MBytes	8358 Mbits/sec	0	72.7 MBytes
[15]	24.00-25.00	sec	1018 MBytes	8535 Mbits/sec	0	73.8 MBytes
[15]	25.00-26.00	sec	1.01 GBytes	8651 Mbits/sec	0	75.0 MBytes
[15]	26.00-27.00	sec	1.02 GBytes	8766 Mbits/sec	0	76.2 MBytes

[15]	27.00-28.00	sec	1.04	GBytes	8902	Mbits/sec	0	77.5	MBytes
[15]	28.00-29.00	sec	1.05	GBytes	9049	Mbits/sec	0	78.8	MBytes
[15]	29.00-30.00	sec	1.07	GBytes	9196	Mbits/sec	0	80.2	MBytes
[15]	30.00-31.00	sec	1.09	GBytes	9395	Mbits/sec	0	81.6	MBytes
[15]	31.00-32.00	sec	1.10	GBytes	9416	Mbits/sec	0	83.1	MBytes
[15]	32.00-33.00	sec	1.10	GBytes	9427	Mbits/sec	0	82.0	MBytes
[15]	33.00-34.00	sec	1.09	GBytes	9384	Mbits/sec	0	82.3	MBytes
[15]	34.00-35.00	sec	1.10	GBytes	9437	Mbits/sec	0	82.8	MBytes
[15]	35.00-36.00	sec	1.10	GBytes	9448	Mbits/sec	0	81.8	MBytes
[15]	36.00-37.00	sec	1.11	GBytes	9542	Mbits/sec	0	81.5	MBytes
[15]	37.00-38.00	sec	1.12	GBytes	9595	Mbits/sec	0	83.0	MBytes
[15]	38.00-39.00	sec	1.11	GBytes	9573	Mbits/sec	0	82.3	MBytes
[15]	39.00-40.00	sec	1.11	GBytes	9574	Mbits/sec	0	81.8	MBytes
[15]	40.00-41.00	sec	1.10	GBytes	9489	Mbits/sec	0	81.4	MBytes
[15]	41.00-42.00	sec	1.11	GBytes	9542	Mbits/sec	0	82.4	MBytes
[15]	42.00-43.00	sec	1.11	GBytes	9542	Mbits/sec	0	83.2	MBytes
[15]	43.00-44.00	sec	1.11	GBytes	9521	Mbits/sec	0	82.9	MBytes
[15]	44.00-45.00	sec	1.11	GBytes	9553	Mbits/sec	0	82.7	MBytes
[15]	45.00-46.00	sec	1.12	GBytes	9584	Mbits/sec	0	82.4	MBytes
[15]	46.00-47.00	sec	1.11	GBytes	9573	Mbits/sec	0	82.2	MBytes
[15]	47.00-48.00	sec	1.12	GBytes	9584	Mbits/sec	0	82.3	MBytes
[15]	48.00-49.00	sec	1.11	GBytes	9574	Mbits/sec	0	82.2	MBytes
[15]	49.00-50.00	sec	1.11	GBytes	9573	Mbits/sec	0	82.5	MBytes
[15]	50.00-51.00	sec	1.11	GBytes	9563	Mbits/sec	0	82.6	MBytes
[15]	51.00-52.00	sec	1.12	GBytes	9584	Mbits/sec	0	82.8	MBytes
[15]	52.00-53.00	sec	1.11	GBytes	9563	Mbits/sec	0	83.1	MBytes
[15]	53.00-54.00	sec	1.11	GBytes	9542	Mbits/sec	0	83.3	MBytes
[15]	54.00-55.00	sec	1.11	GBytes	9511	Mbits/sec	0	82.0	MBytes
[15]	55.00-56.00	sec	1.11	GBytes	9542	Mbits/sec	0	81.3	MBytes
[15]	56.00-57.00	sec	1.10	GBytes	9458	Mbits/sec	748	66.6	MBytes
[15]	57.00-58.00	sec	955	MBytes	8011	Mbits/sec	0	70.4	MBytes
[15]	58.00-59.00	sec	1001	MBytes	8399	Mbits/sec	0	74.1	MBytes
[15]	59.00-60.00	sec	1.03	GBytes	8839	Mbits/sec	0	78.0	MBytes

[ID]	Interval	Transfer	Bandwidth	Retr	sender
[15]	0.00-60.00	sec 58.9 GBytes	8434 Mbits/sec	16389	sender
[15]	0.00-60.00	sec 58.9 GBytes	8433 Mbits/sec		receiver

iperf Done.

SENDER END

Run 2:

```
[root@jz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

```
Connecting to host 198.124.238.150, port 5770
[ 15] local 198.129.77.102 port 60218 connected to 198.124.238.150 port 5770
[ ID] Interval      Transfer      Bandwidth    Retr  Cwnd
[ 15] 0.00-1.00    sec 4.16 MBytes  34.9 Mbits/sec  2    926 KBytes
[ 15] 1.00-2.00    sec 270 MBytes  2267 Mbits/sec  0    76.9 MBytes
[ 15] 2.00-3.00    sec 784 MBytes  6575 Mbits/sec  22693 66.1 MBytes
[ 15] 3.00-4.00    sec 862 MBytes  7235 Mbits/sec  1049 67.0 MBytes
[ 15] 4.00-5.00    sec 931 MBytes  7812 Mbits/sec  0    68.6 MBytes
[ 15] 5.00-6.00    sec 824 MBytes  6910 Mbits/sec  3769 52.7 MBytes
[ 15] 6.00-7.00    sec 904 MBytes  7581 Mbits/sec  476 67.1 MBytes
[ 15] 7.00-8.00    sec 938 MBytes  7864 Mbits/sec  0    68.4 MBytes
[ 15] 8.00-9.00    sec 952 MBytes  7990 Mbits/sec  0    69.9 MBytes
[ 15] 9.00-10.00   sec 978 MBytes  8200 Mbits/sec  0    71.7 MBytes
[ 15] 10.00-11.00  sec 1002 MBytes  8410 Mbits/sec  0    73.7 MBytes
[ 15] 11.00-12.00  sec 1.01 GBytes  8671 Mbits/sec  0    76.0 MBytes
[ 15] 12.00-13.00  sec 1.04 GBytes  8934 Mbits/sec  0    78.6 MBytes
[ 15] 13.00-14.00  sec 1.07 GBytes  9207 Mbits/sec  0    80.8 MBytes
[ 15] 14.00-15.00  sec 1.09 GBytes  9322 Mbits/sec  0    81.3 MBytes
[ 15] 15.00-16.00  sec 1.08 GBytes  9311 Mbits/sec  0    81.1 MBytes
[ 15] 16.00-17.00  sec 1.09 GBytes  9322 Mbits/sec  0    81.3 MBytes
[ 15] 17.00-18.00  sec 1.09 GBytes  9332 Mbits/sec  0    81.3 MBytes
[ 15] 18.00-19.00  sec 1.09 GBytes  9332 Mbits/sec  0    81.0 MBytes
[ 15] 19.00-20.00  sec 1.03 GBytes  8829 Mbits/sec  5942 81.7 MBytes
[ 15] 20.00-21.00  sec 919 MBytes  7707 Mbits/sec  23341 63.9 MBytes
```



```

[ 15] 21.00-22.00 sec 894 MBytes 7498 Mbits/sec 8349 69.2 MBytes
[ 15] 22.00-23.00 sec 989 MBytes 8293 Mbits/sec 0 74.8 MBytes
[ 15] 23.00-24.00 sec 1.04 GBytes 8945 Mbits/sec 0 80.7 MBytes
[ 15] 24.00-25.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.0 MBytes
[ 15] 25.00-26.00 sec 1.09 GBytes 9331 Mbits/sec 0 81.2 MBytes
[ 15] 26.00-27.00 sec 1.09 GBytes 9333 Mbits/sec 0 81.3 MBytes
[ 15] 27.00-28.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.3 MBytes
[ 15] 28.00-29.00 sec 1.09 GBytes 9343 Mbits/sec 0 81.5 MBytes
[ 15] 29.00-30.00 sec 1.09 GBytes 9333 Mbits/sec 0 81.4 MBytes
[ 15] 30.00-31.00 sec 1.09 GBytes 9321 Mbits/sec 0 81.7 MBytes
[ 15] 31.00-32.00 sec 1.09 GBytes 9333 Mbits/sec 0 81.2 MBytes
[ 15] 32.00-33.00 sec 1.09 GBytes 9342 Mbits/sec 0 81.2 MBytes
[ 15] 33.00-34.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.2 MBytes
[ 15] 34.00-35.00 sec 1.09 GBytes 9333 Mbits/sec 0 80.9 MBytes
[ 15] 35.00-36.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.1 MBytes
[ 15] 36.00-37.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.5 MBytes
[ 15] 37.00-38.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.3 MBytes
[ 15] 38.00-39.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.4 MBytes
[ 15] 39.00-40.00 sec 1.09 GBytes 9341 Mbits/sec 0 81.2 MBytes
[ 15] 40.00-41.00 sec 1.09 GBytes 9323 Mbits/sec 0 81.3 MBytes
[ 15] 41.00-42.00 sec 1.09 GBytes 9352 Mbits/sec 0 81.4 MBytes
[ 15] 42.00-43.00 sec 1.08 GBytes 9312 Mbits/sec 0 81.3 MBytes
[ 15] 43.00-44.00 sec 1.09 GBytes 9343 Mbits/sec 0 81.8 MBytes
[ 15] 44.00-45.00 sec 1.08 GBytes 9311 Mbits/sec 0 81.3 MBytes
[ 15] 45.00-46.00 sec 1.09 GBytes 9332 Mbits/sec 0 80.8 MBytes
[ 15] 46.00-47.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.2 MBytes
[ 15] 47.00-48.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.4 MBytes
[ 15] 48.00-49.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.8 MBytes
[ 15] 49.00-50.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.4 MBytes
[ 15] 50.00-51.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.1 MBytes
[ 15] 51.00-52.00 sec 1.09 GBytes 9332 Mbits/sec 0 81.5 MBytes
[ 15] 52.00-53.00 sec 1.09 GBytes 9333 Mbits/sec 0 81.4 MBytes
[ 15] 53.00-54.00 sec 1.09 GBytes 9321 Mbits/sec 0 81.4 MBytes
[ 15] 54.00-55.00 sec 1.09 GBytes 9322 Mbits/sec 0 81.3 MBytes
[ 15] 55.00-56.00 sec 1.08 GBytes 9311 Mbits/sec 0 81.3 MBytes
[ 15] 56.00-57.00 sec 1.07 GBytes 9217 Mbits/sec 3030 81.8 MBytes
[ 15] 57.00-58.00 sec 870 MBytes 7298 Mbits/sec 26918 82.4 MBytes
[ 15] 58.00-59.00 sec 939 MBytes 7875 Mbits/sec 12277 78.9 MBytes
[ 15] 59.00-60.00 sec 1.08 GBytes 9248 Mbits/sec 0 81.2 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 60.3 GBytes 8629 Mbits/sec 107846
[ 15] 0.00-60.00 sec 60.3 GBytes 8629 Mbits/sec
sender
receiver

```

iperf Done.

SENDER END

As in some of the previous cases, we are starting to see the perfect storm of bad behavior becoming more evident:

1. The higher probability of an out of order packet, coupled with a longer delay, means that when we do take a hit we fall harder and must spend more time in recovery
2. These hits to performance occur more frequently, and may cause problems if the behavior happens again in short succession
3. We still are able to show relatively high performance due to the performance of the host (e.g. ample memory, fast processors), but the impact of the performance is noticeable on a macro level.

We can conclude that we are starting to find the threshold for probability of dropping a packet, and the delay that it suffers. TCP is starting to work much harder than it needs to, and this is reflective in the throughput shown in bwctl.

Test Case 6

We will take the settings from test #5, and go a step further. We will keep the 25% of packets being out of order metric, and increase the delay to .8ms on a 70ms RTT path:

LBL Side:

```
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .8ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1
```

BNL Side:

```
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .8ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1
```

After applying these changes, we run two similar bandwidth measurement tests:

Run 1:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

Connecting to host 198.124.238.150, port 5773

```
[ 15] local 198.129.77.102 port 55427 connected to 198.124.238.150 port 5773
[ ID] Interval          Transfer      Bandwidth    Retr  Cwnd
[ 15] 0.00-1.00 sec      3.00 MBytes  25.1 Mbits/sec    0   577 KBytes
[ 15] 1.00-2.00 sec     22.1 MBytes  185 Mbits/sec    0   2.56 MBytes
[ 15] 2.00-3.00 sec     65.0 MBytes  545 Mbits/sec    0   6.70 MBytes
[ 15] 3.00-4.00 sec     161 MBytes  1353 Mbits/sec    0  16.7 MBytes
[ 15] 4.00-5.00 sec     404 MBytes  3387 Mbits/sec    0  42.9 MBytes
[ 15] 5.00-6.00 sec     910 MBytes  7634 Mbits/sec    0  81.0 MBytes
[ 15] 6.00-7.00 sec     1.06 GBytes 9081 Mbits/sec    0  81.4 MBytes
[ 15] 7.00-8.00 sec     1.06 GBytes 9091 Mbits/sec    0  81.2 MBytes
[ 15] 8.00-9.00 sec     1.06 GBytes 9091 Mbits/sec    0  81.3 MBytes
[ 15] 9.00-10.00 sec    1.06 GBytes 9102 Mbits/sec    0  81.3 MBytes
[ 15] 10.00-11.00 sec    1.06 GBytes 9079 Mbits/sec    0  81.4 MBytes
[ 15] 11.00-12.00 sec    1.06 GBytes 9070 Mbits/sec    0  81.3 MBytes
[ 15] 12.00-13.00 sec    1.05 GBytes 9061 Mbits/sec    0  81.2 MBytes
[ 15] 13.00-14.00 sec    904 MBytes 7581 Mbits/sec  4447  64.6 MBytes
[ 15] 14.00-15.00 sec    1.02 GBytes 8798 Mbits/sec    0  81.4 MBytes
[ 15] 15.00-16.00 sec    1.05 GBytes 9060 Mbits/sec    0  81.3 MBytes
[ 15] 16.00-17.00 sec    1.05 GBytes 9059 Mbits/sec    0  81.4 MBytes
[ 15] 17.00-18.00 sec    1.05 GBytes 9060 Mbits/sec    0  81.4 MBytes
[ 15] 18.00-19.00 sec    1.05 GBytes 9049 Mbits/sec    0  81.3 MBytes
[ 15] 19.00-20.00 sec    1.07 GBytes 9227 Mbits/sec    0  81.4 MBytes
[ 15] 20.00-21.00 sec    964 MBytes 8085 Mbits/sec  16542  70.6 MBytes
[ 15] 21.00-22.00 sec    988 MBytes 8283 Mbits/sec  13925  66.1 MBytes
[ 15] 22.00-23.00 sec    872 MBytes 7319 Mbits/sec  30394  66.1 MBytes
[ 15] 23.00-24.00 sec    874 MBytes 7328 Mbits/sec  30787  63.3 MBytes
[ 15] 24.00-25.00 sec    876 MBytes 7352 Mbits/sec  30581  62.8 MBytes
[ 15] 25.00-26.00 sec    868 MBytes 7277 Mbits/sec  30259  82.5 MBytes
[ 15] 26.00-27.00 sec    871 MBytes 7309 Mbits/sec  29128  83.7 MBytes
[ 15] 27.00-28.00 sec    872 MBytes 7319 Mbits/sec  29082  84.1 MBytes
[ 15] 28.00-29.00 sec    868 MBytes 7277 Mbits/sec  29434  83.9 MBytes
[ 15] 29.00-30.00 sec    882 MBytes 7403 Mbits/sec  18207  73.9 MBytes
[ 15] 30.00-31.00 sec    1.04 GBytes 8913 Mbits/sec  4153  81.0 MBytes
[ 15] 31.00-32.00 sec   1008 MBytes 8451 Mbits/sec  4685  74.9 MBytes
[ 15] 32.00-33.00 sec    1.04 GBytes 8924 Mbits/sec  7836  64.1 MBytes
[ 15] 33.00-34.00 sec    875 MBytes 7340 Mbits/sec  29663  82.5 MBytes
[ 15] 34.00-35.00 sec    868 MBytes 7277 Mbits/sec  28906  83.6 MBytes
```

[15]	35.00-36.00	sec	875 MBytes	7340 Mbits/sec	29158	84.2 MBytes
[15]	36.00-37.00	sec	795 MBytes	6669 Mbits/sec	26648	65.2 MBytes
[15]	37.00-38.00	sec	765 MBytes	6417 Mbits/sec	26293	83.0 MBytes
[15]	38.00-39.00	sec	851 MBytes	7141 Mbits/sec	23426	58.3 MBytes
[15]	39.00-40.00	sec	981 MBytes	8231 Mbits/sec	10041	45.8 MBytes
[15]	40.00-41.00	sec	660 MBytes	5537 Mbits/sec	909	73.0 MBytes
[15]	41.00-42.00	sec	1016 MBytes	8525 Mbits/sec	8454	83.7 MBytes
[15]	42.00-43.00	sec	876 MBytes	7350 Mbits/sec	28213	82.7 MBytes
[15]	43.00-44.00	sec	872 MBytes	7319 Mbits/sec	9070	67.9 MBytes
[15]	44.00-45.00	sec	954 MBytes	8001 Mbits/sec	0	70.8 MBytes
[15]	45.00-46.00	sec	1005 MBytes	8431 Mbits/sec	0	73.9 MBytes
[15]	46.00-47.00	sec	996 MBytes	8357 Mbits/sec	0	76.9 MBytes
[15]	47.00-48.00	sec	988 MBytes	8284 Mbits/sec	0	79.8 MBytes
[15]	48.00-49.00	sec	1009 MBytes	8462 Mbits/sec	0	82.7 MBytes
[15]	49.00-50.00	sec	971 MBytes	8147 Mbits/sec	0	82.9 MBytes
[15]	50.00-51.00	sec	1.09 GBytes	9395 Mbits/sec	0	83.5 MBytes
[15]	51.00-52.00	sec	1.09 GBytes	9343 Mbits/sec	0	81.2 MBytes
[15]	52.00-53.00	sec	1021 MBytes	8567 Mbits/sec	251	72.6 MBytes
[15]	53.00-54.00	sec	949 MBytes	7959 Mbits/sec	520	70.1 MBytes
[15]	54.00-55.00	sec	989 MBytes	8294 Mbits/sec	0	73.8 MBytes
[15]	55.00-56.00	sec	1.01 GBytes	8640 Mbits/sec	0	77.5 MBytes
[15]	56.00-57.00	sec	1.01 GBytes	8703 Mbits/sec	0	81.3 MBytes
[15]	57.00-58.00	sec	1.01 GBytes	8672 Mbits/sec	0	81.9 MBytes
[15]	58.00-59.00	sec	1022 MBytes	8577 Mbits/sec	0	82.6 MBytes
[15]	59.00-60.00	sec	1009 MBytes	8462 Mbits/sec	0	83.5 MBytes

[ID]	Interval	Transfer	Bandwidth	Retr	sender
[15]	0.00-60.00	sec 53.0 GBytes	7587 Mbits/sec	501012	receiver
[15]	0.00-60.00	sec 53.0 GBytes	7587 Mbits/sec		

iperf Done.

SENDER END

Run 2:

```
[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available
```

SENDER START

```
Connecting to host 198.124.238.150, port 5775
[ 15] local 198.129.77.102 port 52559 connected to 198.124.238.150 port 5775
```

[ID]	Interval	Transfer	Bandwidth	Retr	Cwnd
[15]	0.00-1.00	sec 1.47 MBytes	12.3 Mbits/sec	5	227 KBytes
[15]	1.00-2.00	sec 15.1 MBytes	126 Mbits/sec	0	1.88 MBytes
[15]	2.00-3.00	sec 48.8 MBytes	409 Mbits/sec	0	4.78 MBytes
[15]	3.00-4.00	sec 101 MBytes	849 Mbits/sec	247	4.71 MBytes
[15]	4.00-5.00	sec 100 MBytes	839 Mbits/sec	76	11.8 MBytes
[15]	5.00-6.00	sec 268 MBytes	2244 Mbits/sec	0	28.2 MBytes
[15]	6.00-7.00	sec 549 MBytes	4603 Mbits/sec	1023	17.1 MBytes
[15]	7.00-8.00	sec 176 MBytes	1478 Mbits/sec	1222	10.6 MBytes
[15]	8.00-9.00	sec 162 MBytes	1363 Mbits/sec	518	18.1 MBytes
[15]	9.00-10.00	sec 268 MBytes	2244 Mbits/sec	1022	18.7 MBytes
[15]	10.00-11.00	sec 415 MBytes	3481 Mbits/sec	0	48.5 MBytes
[15]	11.00-12.00	sec 919 MBytes	7707 Mbits/sec	3701	82.9 MBytes
[15]	12.00-13.00	sec 851 MBytes	7141 Mbits/sec	29884	83.6 MBytes
[15]	13.00-14.00	sec 872 MBytes	7319 Mbits/sec	28802	82.2 MBytes
[15]	14.00-15.00	sec 871 MBytes	7309 Mbits/sec	31084	66.1 MBytes
[15]	15.00-16.00	sec 881 MBytes	7392 Mbits/sec	30090	66.6 MBytes
[15]	16.00-17.00	sec 878 MBytes	7361 Mbits/sec	30710	63.0 MBytes
[15]	17.00-18.00	sec 951 MBytes	7979 Mbits/sec	15131	81.4 MBytes
[15]	18.00-19.00	sec 1004 MBytes	8421 Mbits/sec	4734	81.4 MBytes
[15]	19.00-20.00	sec 1.03 GBytes	8860 Mbits/sec	9945	63.9 MBytes
[15]	20.00-21.00	sec 878 MBytes	7361 Mbits/sec	33750	47.1 MBytes
[15]	21.00-22.00	sec 844 MBytes	7078 Mbits/sec	23632	71.4 MBytes
[15]	22.00-23.00	sec 1.00 GBytes	8619 Mbits/sec	9580	65.1 MBytes
[15]	23.00-24.00	sec 881 MBytes	7392 Mbits/sec	29533	64.3 MBytes
[15]	24.00-25.00	sec 955 MBytes	8011 Mbits/sec	15070	81.1 MBytes
[15]	25.00-26.00	sec 1.09 GBytes	9353 Mbits/sec	0	81.3 MBytes
[15]	26.00-27.00	sec 1.05 GBytes	9007 Mbits/sec	3651	81.2 MBytes
[15]	27.00-28.00	sec 1.05 GBytes	9018 Mbits/sec	3891	81.3 MBytes
[15]	28.00-29.00	sec 966 MBytes	8106 Mbits/sec	17276	48.6 MBytes

```

[ 15] 29.00-30.00 sec 834 MBytes 6993 Mbits/sec 472 81.3 MBytes
[ 15] 30.00-31.00 sec 999 MBytes 8379 Mbits/sec 4843 81.4 MBytes
[ 15] 31.00-32.00 sec 1.09 GBytes 9353 Mbits/sec 0 81.3 MBytes
[ 15] 32.00-33.00 sec 981 MBytes 8231 Mbits/sec 8725 83.4 MBytes
[ 15] 33.00-34.00 sec 875 MBytes 7340 Mbits/sec 28915 67.2 MBytes
[ 15] 34.00-35.00 sec 959 MBytes 8043 Mbits/sec 16795 81.3 MBytes
[ 15] 35.00-36.00 sec 906 MBytes 7602 Mbits/sec 27307 84.1 MBytes
[ 15] 36.00-37.00 sec 876 MBytes 7349 Mbits/sec 28687 83.4 MBytes
[ 15] 37.00-38.00 sec 914 MBytes 7666 Mbits/sec 22307 81.8 MBytes
[ 15] 38.00-39.00 sec 1.05 GBytes 8997 Mbits/sec 5079 66.4 MBytes
[ 15] 39.00-40.00 sec 1.02 GBytes 8787 Mbits/sec 2731 81.4 MBytes
[ 15] 40.00-41.00 sec 1.06 GBytes 9070 Mbits/sec 0 81.5 MBytes
[ 15] 41.00-42.00 sec 1.06 GBytes 9071 Mbits/sec 0 81.2 MBytes
[ 15] 42.00-43.00 sec 1.06 GBytes 9122 Mbits/sec 4267 65.9 MBytes
[ 15] 43.00-44.00 sec 949 MBytes 7959 Mbits/sec 5765 61.7 MBytes
[ 15] 44.00-45.00 sec 1005 MBytes 8431 Mbits/sec 423 81.3 MBytes
[ 15] 45.00-46.00 sec 1.06 GBytes 9081 Mbits/sec 0 81.5 MBytes
[ 15] 46.00-47.00 sec 1.06 GBytes 9070 Mbits/sec 0 81.5 MBytes
[ 15] 47.00-48.00 sec 1.06 GBytes 9091 Mbits/sec 0 81.3 MBytes
[ 15] 48.00-49.00 sec 1.06 GBytes 9080 Mbits/sec 0 81.2 MBytes
[ 15] 49.00-50.00 sec 1.06 GBytes 9070 Mbits/sec 5161 63.3 MBytes
[ 15] 50.00-51.00 sec 1.00 GBytes 8609 Mbits/sec 0 81.4 MBytes
[ 15] 51.00-52.00 sec 1020 MBytes 8556 Mbits/sec 9140 68.5 MBytes
[ 15] 52.00-53.00 sec 719 MBytes 6029 Mbits/sec 8394 80.9 MBytes
[ 15] 53.00-54.00 sec 1.05 GBytes 9018 Mbits/sec 0 81.2 MBytes
[ 15] 54.00-55.00 sec 1.05 GBytes 9028 Mbits/sec 0 81.4 MBytes
[ 15] 55.00-56.00 sec 1.05 GBytes 9049 Mbits/sec 0 81.4 MBytes
[ 15] 56.00-57.00 sec 1.06 GBytes 9134 Mbits/sec 0 81.0 MBytes
[ 15] 57.00-58.00 sec 1.06 GBytes 9144 Mbits/sec 0 81.4 MBytes
[ 15] 58.00-59.00 sec 1.05 GBytes 9039 Mbits/sec 0 81.3 MBytes
[ 15] 59.00-60.00 sec 1.05 GBytes 9039 Mbits/sec 0 81.2 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 49.5 GBytes 7092 Mbits/sec 503588
[ 15] 0.00-60.00 sec 49.5 GBytes 7092 Mbits/sec
sender
receiver

```

iperf Done.

SENDER END

The data from this test is very stark: the combination of a long delay and frequently occurring incidents of out of order packets is causing significant problems for our transfer. On a macro scale, we can see that performance is recoverable due to the fast hosts, but they are spending a lot of time working out the network behavior.

Test Case 7

We will take the settings from test #6, and go a step further still, given we are starting to see the limits of the hosts and TCP. We will keep the 25% of packets being out of order metric, and increase the delay to .9ms on a 70ms RTT path:

LBL Side:

```

sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay .9ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1

```

BNL Side:

```

sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay .9ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1

```

After applying these changes, we run two similar bandwidth measurement tests:

Run 1:

```
[root@l1-bl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available
```

SENDER START

Connecting to host 198.124.238.150, port 5778

[15] local 198.129.77.102 port 55174 connected to 198.124.238.150 port 5778

[ID]	Interval		Transfer	Bandwidth	Retr	Cwnd
[15]	0.00-1.00	sec	2.24 MBytes	18.8 Mb/s	0	384 KBytes
[15]	1.00-2.00	sec	26.0 MBytes	218 Mb/s	0	3.29 MBytes
[15]	2.00-3.00	sec	81.2 MBytes	682 Mb/s	0	8.47 MBytes
[15]	3.00-4.00	sec	178 MBytes	1489 Mb/s	0	16.8 MBytes
[15]	4.00-5.00	sec	241 MBytes	2024 Mb/s	987	8.73 MBytes
[15]	5.00-6.00	sec	182 MBytes	1531 Mb/s	274	22.6 MBytes
[15]	6.00-7.00	sec	376 MBytes	3156 Mb/s	1146	12.7 MBytes
[15]	7.00-8.00	sec	140 MBytes	1174 Mb/s	1238	9.02 MBytes
[15]	8.00-9.00	sec	186 MBytes	1562 Mb/s	218	22.7 MBytes
[15]	9.00-10.00	sec	521 MBytes	4373 Mb/s	0	60.3 MBytes
[15]	10.00-11.00	sec	768 MBytes	6438 Mb/s	3538	31.3 MBytes
[15]	11.00-12.00	sec	679 MBytes	5694 Mb/s	1287	53.9 MBytes
[15]	12.00-13.00	sec	745 MBytes	6250 Mb/s	1702	82.1 MBytes
[15]	13.00-14.00	sec	960 MBytes	8053 Mb/s	6504	64.1 MBytes
[15]	14.00-15.00	sec	1.03 GBytes	8840 Mb/s	3128	64.8 MBytes
[15]	15.00-16.00	sec	936 MBytes	7854 Mb/s	16145	83.7 MBytes
[15]	16.00-17.00	sec	841 MBytes	7057 Mb/s	29270	82.9 MBytes
[15]	17.00-18.00	sec	845 MBytes	7088 Mb/s	29190	61.0 MBytes
[15]	18.00-19.00	sec	878 MBytes	7361 Mb/s	18354	81.3 MBytes
[15]	19.00-20.00	sec	938 MBytes	7864 Mb/s	11923	67.7 MBytes
[15]	20.00-21.00	sec	1018 MBytes	8535 Mb/s	4558	81.4 MBytes
[15]	21.00-22.00	sec	878 MBytes	7362 Mb/s	19089	64.6 MBytes
[15]	22.00-23.00	sec	1012 MBytes	8493 Mb/s	4133	81.4 MBytes
[15]	23.00-24.00	sec	1.02 GBytes	8745 Mb/s	3276	81.3 MBytes
[15]	24.00-25.00	sec	975 MBytes	8179 Mb/s	7512	80.3 MBytes
[15]	25.00-26.00	sec	1020 MBytes	8556 Mb/s	7013	69.8 MBytes
[15]	26.00-27.00	sec	986 MBytes	8273 Mb/s	4248	83.8 MBytes
[15]	27.00-28.00	sec	1.01 GBytes	8703 Mb/s	3851	81.3 MBytes
[15]	28.00-29.00	sec	995 MBytes	8347 Mb/s	11645	83.8 MBytes
[15]	29.00-30.00	sec	928 MBytes	7781 Mb/s	8045	70.5 MBytes
[15]	30.00-31.00	sec	982 MBytes	8241 Mb/s	4475	81.8 MBytes
[15]	31.00-32.00	sec	1.03 GBytes	8818 Mb/s	3131	81.4 MBytes
[15]	32.00-33.00	sec	991 MBytes	8315 Mb/s	4200	81.5 MBytes
[15]	33.00-34.00	sec	911 MBytes	7644 Mb/s	5466	81.3 MBytes
[15]	34.00-35.00	sec	1006 MBytes	8441 Mb/s	7281	83.2 MBytes
[15]	35.00-36.00	sec	990 MBytes	8305 Mb/s	4266	81.4 MBytes
[15]	36.00-37.00	sec	899 MBytes	7539 Mb/s	9274	60.4 MBytes
[15]	37.00-38.00	sec	920 MBytes	7717 Mb/s	9865	44.4 MBytes
[15]	38.00-39.00	sec	816 MBytes	6847 Mb/s	2203	69.0 MBytes
[15]	39.00-40.00	sec	980 MBytes	8221 Mb/s	4180	80.2 MBytes
[15]	40.00-41.00	sec	1.02 GBytes	8735 Mb/s	7275	64.7 MBytes
[15]	41.00-42.00	sec	1000 MBytes	8389 Mb/s	4144	81.2 MBytes
[15]	42.00-43.00	sec	959 MBytes	8043 Mb/s	8476	71.2 MBytes
[15]	43.00-44.00	sec	986 MBytes	8273 Mb/s	4322	81.6 MBytes
[15]	44.00-45.00	sec	1002 MBytes	8410 Mb/s	10077	63.7 MBytes
[15]	45.00-46.00	sec	998 MBytes	8368 Mb/s	5312	81.3 MBytes
[15]	46.00-47.00	sec	990 MBytes	8305 Mb/s	8648	66.8 MBytes
[15]	47.00-48.00	sec	992 MBytes	8326 Mb/s	4769	66.4 MBytes
[15]	48.00-49.00	sec	972 MBytes	8158 Mb/s	5601	56.0 MBytes
[15]	49.00-50.00	sec	854 MBytes	7162 Mb/s	1005	81.4 MBytes
[15]	50.00-51.00	sec	950 MBytes	7969 Mb/s	16608	69.8 MBytes
[15]	51.00-52.00	sec	960 MBytes	8053 Mb/s	4950	83.6 MBytes
[15]	52.00-53.00	sec	1010 MBytes	8472 Mb/s	7349	72.4 MBytes
[15]	53.00-54.00	sec	972 MBytes	8158 Mb/s	4798	82.4 MBytes
[15]	54.00-55.00	sec	974 MBytes	8168 Mb/s	14520	64.7 MBytes
[15]	55.00-56.00	sec	986 MBytes	8272 Mb/s	2712	64.7 MBytes
[15]	56.00-57.00	sec	948 MBytes	7949 Mb/s	7503	72.9 MBytes
[15]	57.00-58.00	sec	921 MBytes	7728 Mb/s	20405	63.5 MBytes
[15]	58.00-59.00	sec	830 MBytes	6963 Mb/s	27416	66.3 MBytes
[15]	59.00-60.00	sec	826 MBytes	6931 Mb/s	28182	71.7 MBytes

```

-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec  48.0 GBytes  6877 Mbits/sec  446687
[ 15] 0.00-60.00 sec  48.0 GBytes  6877 Mbits/sec

```

iperf Done.

SENDER END

Run 2:

```

[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 67 seconds until test results available

```

SENDER START

Connecting to host 198.124.238.150, port 5780

[15] local 198.129.77.102 port 42780 connected to 198.124.238.150 port 5780

```

[ ID] Interval      Transfer      Bandwidth      Retr  Cwnd
[ 15] 0.00-1.00 sec  2.63 MBytes  22.0 Mbits/sec    1   516 KBytes
[ 15] 1.00-2.00 sec  23.9 MBytes  200 Mbits/sec    0   3.10 MBytes
[ 15] 2.00-3.00 sec  76.2 MBytes  640 Mbits/sec    0   7.91 MBytes
[ 15] 3.00-4.00 sec  90.0 MBytes  755 Mbits/sec  196   8.94 MBytes
[ 15] 4.00-5.00 sec  124 MBytes  1038 Mbits/sec  148   7.99 MBytes
[ 15] 5.00-6.00 sec  206 MBytes  1730 Mbits/sec    0  22.6 MBytes
[ 15] 6.00-7.00 sec  356 MBytes  2988 Mbits/sec  643  27.2 MBytes
[ 15] 7.00-8.00 sec  442 MBytes  3712 Mbits/sec  2714  23.4 MBytes
[ 15] 8.00-9.00 sec  436 MBytes  3660 Mbits/sec  6176  45.0 MBytes
[ 15] 9.00-10.00 sec  441 MBytes  3701 Mbits/sec  3362  20.2 MBytes
[ 15] 10.00-11.00 sec  281 MBytes  2359 Mbits/sec  3680  27.1 MBytes
[ 15] 11.00-12.00 sec  452 MBytes  3796 Mbits/sec  1209  42.9 MBytes
[ 15] 12.00-13.00 sec  686 MBytes  5757 Mbits/sec  1106  73.0 MBytes
[ 15] 13.00-14.00 sec  886 MBytes  7434 Mbits/sec  3003  82.8 MBytes
[ 15] 14.00-15.00 sec  871 MBytes  7309 Mbits/sec  3471  81.4 MBytes
[ 15] 15.00-16.00 sec  1004 MBytes  8419 Mbits/sec  7557  81.3 MBytes
[ 15] 16.00-17.00 sec  921 MBytes  7729 Mbits/sec  12722  76.0 MBytes
[ 15] 17.00-18.00 sec  994 MBytes  8336 Mbits/sec  7647  62.2 MBytes
[ 15] 18.00-19.00 sec  958 MBytes  8032 Mbits/sec  12464  83.7 MBytes
[ 15] 19.00-20.00 sec  848 MBytes  7109 Mbits/sec  29326  84.2 MBytes
[ 15] 20.00-21.00 sec  930 MBytes  7801 Mbits/sec  17083  66.8 MBytes
[ 15] 21.00-22.00 sec  896 MBytes  7518 Mbits/sec  5469  68.3 MBytes
[ 15] 22.00-23.00 sec  989 MBytes  8294 Mbits/sec  2627  81.3 MBytes
[ 15] 23.00-24.00 sec  985 MBytes  8263 Mbits/sec  10957  74.9 MBytes
[ 15] 24.00-25.00 sec  928 MBytes  7780 Mbits/sec  18857  65.4 MBytes
[ 15] 25.00-26.00 sec  965 MBytes  8094 Mbits/sec  10950  57.8 MBytes
[ 15] 26.00-27.00 sec  994 MBytes  8337 Mbits/sec  3770  82.3 MBytes
[ 15] 27.00-28.00 sec  1.02 GBytes  8756 Mbits/sec  3378  81.3 MBytes
[ 15] 28.00-29.00 sec  1.01 GBytes  8682 Mbits/sec  6941  66.8 MBytes
[ 15] 29.00-30.00 sec  1010 MBytes  8472 Mbits/sec  4007  80.9 MBytes
[ 15] 30.00-31.00 sec  1001 MBytes  8399 Mbits/sec  7651  82.1 MBytes
[ 15] 31.00-32.00 sec  981 MBytes  8230 Mbits/sec  4496  81.1 MBytes
[ 15] 32.00-33.00 sec  985 MBytes  8264 Mbits/sec  10254  76.2 MBytes
[ 15] 33.00-34.00 sec  998 MBytes  8368 Mbits/sec  4096  81.6 MBytes
[ 15] 34.00-35.00 sec  992 MBytes  8325 Mbits/sec  4508  81.3 MBytes
[ 15] 35.00-36.00 sec  991 MBytes  8316 Mbits/sec  3816  81.2 MBytes
[ 15] 36.00-37.00 sec  1011 MBytes  8483 Mbits/sec  7828  57.6 MBytes
[ 15] 37.00-38.00 sec  1.01 GBytes  8693 Mbits/sec  4097  63.4 MBytes
[ 15] 38.00-39.00 sec  1000 MBytes  8389 Mbits/sec  3397  81.3 MBytes
[ 15] 39.00-40.00 sec  965 MBytes  8095 Mbits/sec  8496  70.1 MBytes
[ 15] 40.00-41.00 sec  1.04 GBytes  8902 Mbits/sec    0  81.4 MBytes
[ 15] 41.00-42.00 sec  1.02 GBytes  8756 Mbits/sec  5566  63.2 MBytes
[ 15] 42.00-43.00 sec  875 MBytes  7340 Mbits/sec  14778  81.5 MBytes
[ 15] 43.00-44.00 sec  996 MBytes  8357 Mbits/sec  8586  65.7 MBytes
[ 15] 44.00-45.00 sec  985 MBytes  8263 Mbits/sec  587  81.4 MBytes
[ 15] 45.00-46.00 sec  1.02 GBytes  8724 Mbits/sec  5784  64.7 MBytes
[ 15] 46.00-47.00 sec  1005 MBytes  8431 Mbits/sec  5054  81.1 MBytes
[ 15] 47.00-48.00 sec  1019 MBytes  8546 Mbits/sec  4210  65.9 MBytes
[ 15] 48.00-49.00 sec  1015 MBytes  8515 Mbits/sec  4517  81.2 MBytes
[ 15] 49.00-50.00 sec  1.02 GBytes  8734 Mbits/sec  4285  77.0 MBytes
[ 15] 50.00-51.00 sec  980 MBytes  8221 Mbits/sec  8356  61.0 MBytes
[ 15] 51.00-52.00 sec  1.02 GBytes  8724 Mbits/sec    0  81.3 MBytes
[ 15] 52.00-53.00 sec  855 MBytes  7172 Mbits/sec  27114  83.4 MBytes
[ 15] 53.00-54.00 sec  846 MBytes  7099 Mbits/sec  29512  83.6 MBytes

```

```

[ 15] 54.00-55.00 sec 844 MBytes 7078 Mbits/sec 29307 83.7 MBytes
[ 15] 55.00-56.00 sec 844 MBytes 7078 Mbits/sec 28369 83.9 MBytes
[ 15] 56.00-57.00 sec 832 MBytes 6983 Mbits/sec 28921 63.6 MBytes
[ 15] 57.00-58.00 sec 946 MBytes 7938 Mbits/sec 5237 75.0 MBytes
[ 15] 58.00-59.00 sec 972 MBytes 8158 Mbits/sec 4530 83.8 MBytes
[ 15] 59.00-60.00 sec 984 MBytes 8252 Mbits/sec 5161 81.3 MBytes
-----
[ ID] Interval      Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 47.9 GBytes 6859 Mbits/sec 457977
[ 15] 0.00-60.00 sec 47.9 GBytes 6859 Mbits/sec

```

iperf Done.

SENDER END

We can conclude that the trend from tests #6 and #7 will continue, and the overall performance average (as seen in the individual intervals) is dropping. The higher number of out of order packets, and increasing delay at which they arrive, cause much consternation for TCP and we are re-sending significant amounts of data to compensate. We keep our bandwidth high, but the hosts are working very hard to do so.

Test Case 8

This experiment goes back where we left off in test #3: we will only out of order 1% of the packets, but we will increase our delay to a significant point: 1ms for a 70ms path. As we have seen so far, the delay of the packet matters a lot, because this relates to TCP's internal timer mechanism:

```

LBL Side:
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay 1ms reorder 99% 100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst 198.124.238.150/32 flowid 1:1

```

```

BNL Side:
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay 1ms reorder 99% 100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst 198.129.77.102/32 flowid 1:1

```

After applying these changes, we run a similar bandwidth measurement test:

```

[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 66 seconds until test results available

```

```

SENDER START
Connecting to host 198.124.238.150, port 5786
[ 15] local 198.129.77.102 port 50538 connected to 198.124.238.150 port 5786
[ ID] Interval      Transfer      Bandwidth      Retr  Cwnd
[ 15] 0.00-1.00 sec 2.42 MBytes 20.3 Mbits/sec 0 489 KBytes
[ 15] 1.00-2.00 sec 22.2 MBytes 186 Mbits/sec 0 2.72 MBytes
[ 15] 2.00-3.00 sec 58.8 MBytes 493 Mbits/sec 0 5.67 MBytes
[ 15] 3.00-4.00 sec 108 MBytes 902 Mbits/sec 636 10.4 MBytes
[ 15] 4.00-5.00 sec 148 MBytes 1237 Mbits/sec 820 9.08 MBytes
[ 15] 5.00-6.00 sec 105 MBytes 881 Mbits/sec 846 5.84 MBytes
[ 15] 6.00-7.00 sec 92.5 MBytes 776 Mbits/sec 1147 6.47 MBytes
[ 15] 7.00-8.00 sec 110 MBytes 923 Mbits/sec 912 6.86 MBytes
[ 15] 8.00-9.00 sec 118 MBytes 986 Mbits/sec 908 7.19 MBytes
[ 15] 9.00-10.00 sec 169 MBytes 1416 Mbits/sec 835 13.9 MBytes

```

[15]	10.00-11.00	sec	150 MBytes	1258 Mbits/sec	1550	9.39 MBytes
[15]	11.00-12.00	sec	111 MBytes	933 Mbits/sec	1533	7.44 MBytes
[15]	12.00-13.00	sec	86.2 MBytes	724 Mbits/sec	1469	5.69 MBytes
[15]	13.00-14.00	sec	112 MBytes	944 Mbits/sec	912	7.59 MBytes
[15]	14.00-15.00	sec	108 MBytes	902 Mbits/sec	944	8.17 MBytes
[15]	15.00-16.00	sec	110 MBytes	923 Mbits/sec	1349	8.38 MBytes
[15]	16.00-17.00	sec	101 MBytes	849 Mbits/sec	1164	6.75 MBytes
[15]	17.00-18.00	sec	100 MBytes	839 Mbits/sec	1422	6.69 MBytes
[15]	18.00-19.00	sec	121 MBytes	1017 Mbits/sec	545	13.8 MBytes
[15]	19.00-20.00	sec	205 MBytes	1720 Mbits/sec	1424	11.5 MBytes
[15]	20.00-21.00	sec	134 MBytes	1122 Mbits/sec	1418	9.00 MBytes
[15]	21.00-22.00	sec	134 MBytes	1122 Mbits/sec	1538	9.49 MBytes
[15]	22.00-23.00	sec	140 MBytes	1174 Mbits/sec	1033	8.30 MBytes
[15]	23.00-24.00	sec	135 MBytes	1132 Mbits/sec	1485	9.71 MBytes
[15]	24.00-25.00	sec	121 MBytes	1017 Mbits/sec	1188	11.2 MBytes
[15]	25.00-26.00	sec	149 MBytes	1248 Mbits/sec	1449	10.2 MBytes
[15]	26.00-27.00	sec	194 MBytes	1625 Mbits/sec	467	14.9 MBytes
[15]	27.00-28.00	sec	120 MBytes	1007 Mbits/sec	3196	7.21 MBytes
[15]	28.00-29.00	sec	102 MBytes	860 Mbits/sec	739	8.30 MBytes
[15]	29.00-30.00	sec	136 MBytes	1143 Mbits/sec	426	12.0 MBytes
[15]	30.00-31.00	sec	189 MBytes	1583 Mbits/sec	1338	16.9 MBytes
[15]	31.00-32.00	sec	215 MBytes	1804 Mbits/sec	1378	14.7 MBytes
[15]	32.00-33.00	sec	230 MBytes	1929 Mbits/sec	1021	20.1 MBytes
[15]	33.00-34.00	sec	265 MBytes	2223 Mbits/sec	1909	14.1 MBytes
[15]	34.00-35.00	sec	166 MBytes	1395 Mbits/sec	1609	9.01 MBytes
[15]	35.00-36.00	sec	149 MBytes	1248 Mbits/sec	633	16.3 MBytes
[15]	36.00-37.00	sec	228 MBytes	1908 Mbits/sec	2712	17.3 MBytes
[15]	37.00-38.00	sec	226 MBytes	1898 Mbits/sec	2390	13.6 MBytes
[15]	38.00-39.00	sec	165 MBytes	1384 Mbits/sec	1138	16.4 MBytes
[15]	39.00-40.00	sec	212 MBytes	1783 Mbits/sec	1044	18.2 MBytes
[15]	40.00-41.00	sec	239 MBytes	2003 Mbits/sec	1934	11.4 MBytes
[15]	41.00-42.00	sec	146 MBytes	1227 Mbits/sec	2279	9.98 MBytes
[15]	42.00-43.00	sec	120 MBytes	1007 Mbits/sec	1245	8.06 MBytes
[15]	43.00-44.00	sec	108 MBytes	902 Mbits/sec	470	8.10 MBytes
[15]	44.00-45.00	sec	114 MBytes	954 Mbits/sec	1566	6.78 MBytes
[15]	45.00-46.00	sec	100 MBytes	839 Mbits/sec	1306	6.81 MBytes
[15]	46.00-47.00	sec	102 MBytes	860 Mbits/sec	1044	7.63 MBytes
[15]	47.00-48.00	sec	141 MBytes	1185 Mbits/sec	418	17.7 MBytes
[15]	48.00-49.00	sec	252 MBytes	2118 Mbits/sec	1540	18.7 MBytes
[15]	49.00-50.00	sec	244 MBytes	2045 Mbits/sec	1712	14.4 MBytes
[15]	50.00-51.00	sec	148 MBytes	1237 Mbits/sec	2647	8.40 MBytes
[15]	51.00-52.00	sec	93.8 MBytes	786 Mbits/sec	1168	5.99 MBytes
[15]	52.00-53.00	sec	91.2 MBytes	765 Mbits/sec	1163	5.96 MBytes
[15]	53.00-54.00	sec	96.2 MBytes	807 Mbits/sec	871	7.23 MBytes
[15]	54.00-55.00	sec	124 MBytes	1038 Mbits/sec	749	12.2 MBytes
[15]	55.00-56.00	sec	188 MBytes	1573 Mbits/sec	1279	12.4 MBytes
[15]	56.00-57.00	sec	160 MBytes	1342 Mbits/sec	1535	8.83 MBytes
[15]	57.00-58.00	sec	146 MBytes	1227 Mbits/sec	1150	9.98 MBytes
[15]	58.00-59.00	sec	119 MBytes	996 Mbits/sec	1505	8.12 MBytes
[15]	59.00-60.00	sec	116 MBytes	975 Mbits/sec	1375	7.66 MBytes

[ID]	Interval		Transfer	Bandwidth	Retr	
[15]	0.00-60.00	sec	8.20 GBytes	1174 Mbits/sec	73483	sender
[15]	0.00-60.00	sec	8.17 GBytes	1170 Mbits/sec		receiver

iperf Done.

SENDER END

The data from this test is conclusive – the delay of an out of order packet matters. Even for a small percentage (1% of all packets), the fact that we must wait more than 1% of the RTT for the path seems to cause absolute havoc for TCP and the hosts. All of the time is spend recovering in this example, and our bandwidth result (for each interval, and overall) is extremely low.

Test Case 9

Similar to case #8, we can use the same overall delay of an out of order packet (1ms), and have a higher number of out of order packets (25%):


```

LBL Side:
sudo /sbin/tc qdisc delete dev eth2.912 root
sudo /sbin/tc qdisc add dev eth2.912 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.912 parent 1:1 handle 10: netem delay 1ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.912 protocol ip parent 1:0 prio 3 u32 match ip dst
198.124.238.150/32 flowid 1:1

```

```

BNL Side:
sudo /sbin/tc qdisc delete dev eth2.917 root
sudo /sbin/tc qdisc add dev eth2.917 root handle 1: prio
sudo /sbin/tc qdisc add dev eth2.917 parent 1:1 handle 10: netem delay 1ms reorder 75%
100%
sudo /sbin/tc filter add dev eth2.917 protocol ip parent 1:0 prio 3 u32 match ip dst
198.129.77.102/32 flowid 1:1

```

After applying these changes, we run a similar bandwidth measurement test:

```

[rootjz@lbl-diskpt1 ~]# bwctl -T iperf3 -f m -t 60 -i 1 -c bnl-diskpt1.es.net
bwctl: Using tool: iperf3
bwctl: 65 seconds until test results available

```

```

SENDER START
Connecting to host 198.124.238.150, port 5785
[ 15] local 198.129.77.102 port 37616 connected to 198.124.238.150 port 5785
[ ID] Interval      Transfer      Bandwidth     Retr  Cwnd
[ 15] 0.00-1.00    sec  3.92 MBytes  32.9 Mbits/sec  2    743 KBytes
[ 15] 1.00-2.00    sec  31.9 MBytes  268 Mbits/sec  0    3.96 MBytes
[ 15] 2.00-3.00    sec  90.0 MBytes  755 Mbits/sec  0    8.74 MBytes
[ 15] 3.00-4.00    sec  150 MBytes  1258 Mbits/sec  380  4.65 MBytes
[ 15] 4.00-5.00    sec  81.2 MBytes  682 Mbits/sec  674  8.89 MBytes
[ 15] 5.00-6.00    sec  120 MBytes  1007 Mbits/sec  677  12.5 MBytes
[ 15] 6.00-7.00    sec  141 MBytes  1185 Mbits/sec  1428  7.69 MBytes
[ 15] 7.00-8.00    sec  112 MBytes  944 Mbits/sec  1279  7.82 MBytes
[ 15] 8.00-9.00    sec  112 MBytes  944 Mbits/sec  1244  7.13 MBytes
[ 15] 9.00-10.00   sec  112 MBytes  944 Mbits/sec  1458  7.58 MBytes
[ 15] 10.00-11.00  sec  109 MBytes  912 Mbits/sec  1224  7.89 MBytes
[ 15] 11.00-12.00  sec  111 MBytes  933 Mbits/sec  1598  7.05 MBytes
[ 15] 12.00-13.00  sec  106 MBytes  891 Mbits/sec  1879  6.94 MBytes
[ 15] 13.00-14.00  sec  108 MBytes  902 Mbits/sec  963  6.58 MBytes
[ 15] 14.00-15.00  sec  101 MBytes  849 Mbits/sec  1224  6.76 MBytes
[ 15] 15.00-16.00  sec  95.0 MBytes  797 Mbits/sec  1084  6.11 MBytes
[ 15] 16.00-17.00  sec  101 MBytes  849 Mbits/sec  1046  8.76 MBytes
[ 15] 17.00-18.00  sec  104 MBytes  870 Mbits/sec  1340  7.13 MBytes
[ 15] 18.00-19.00  sec  105 MBytes  881 Mbits/sec  724  10.8 MBytes
[ 15] 19.00-20.00  sec  171 MBytes  1437 Mbits/sec  675  12.5 MBytes
[ 15] 20.00-21.00  sec  131 MBytes  1101 Mbits/sec  1913  6.44 MBytes
[ 15] 21.00-22.00  sec  114 MBytes  954 Mbits/sec  553  12.6 MBytes
[ 15] 22.00-23.00  sec  156 MBytes  1311 Mbits/sec  1530  8.38 MBytes
[ 15] 23.00-24.00  sec  184 MBytes  1541 Mbits/sec  195  18.9 MBytes
[ 15] 24.00-25.00  sec  324 MBytes  2716 Mbits/sec  369  11.8 MBytes
[ 15] 25.00-26.00  sec  132 MBytes  1111 Mbits/sec  1697  8.60 MBytes
[ 15] 26.00-27.00  sec  139 MBytes  1164 Mbits/sec  673  9.23 MBytes
[ 15] 27.00-28.00  sec  140 MBytes  1174 Mbits/sec  1119  8.99 MBytes
[ 15] 28.00-29.00  sec  120 MBytes  1007 Mbits/sec  1688  6.70 MBytes
[ 15] 29.00-30.00  sec  105 MBytes  881 Mbits/sec  880  7.18 MBytes
[ 15] 30.00-31.00  sec  112 MBytes  944 Mbits/sec  827  9.75 MBytes
[ 15] 31.00-32.00  sec  119 MBytes  996 Mbits/sec  940  7.22 MBytes
[ 15] 32.00-33.00  sec  110 MBytes  923 Mbits/sec  1104  8.37 MBytes
[ 15] 33.00-34.00  sec  115 MBytes  965 Mbits/sec  1350  9.55 MBytes
[ 15] 34.00-35.00  sec  134 MBytes  1122 Mbits/sec  790  10.7 MBytes
[ 15] 35.00-36.00  sec  122 MBytes  1028 Mbits/sec  1545  10.2 MBytes
[ 15] 36.00-37.00  sec  134 MBytes  1122 Mbits/sec  1435  9.15 MBytes
[ 15] 37.00-38.00  sec  112 MBytes  944 Mbits/sec  1436  7.76 MBytes
[ 15] 38.00-39.00  sec  112 MBytes  944 Mbits/sec  994  7.94 MBytes
[ 15] 39.00-40.00  sec  146 MBytes  1227 Mbits/sec  915  10.2 MBytes
[ 15] 40.00-41.00  sec  132 MBytes  1111 Mbits/sec  1614  8.00 MBytes
[ 15] 41.00-42.00  sec  109 MBytes  912 Mbits/sec  1144  9.80 MBytes
[ 15] 42.00-43.00  sec  105 MBytes  881 Mbits/sec  1700  5.31 MBytes

```

```

[ 15] 43.00-44.00 sec 75.0 MBytes 629 Mbits/sec 693 5.59 MBytes
[ 15] 44.00-45.00 sec 85.0 MBytes 713 Mbits/sec 891 5.85 MBytes
[ 15] 45.00-46.00 sec 91.2 MBytes 765 Mbits/sec 1011 6.32 MBytes
[ 15] 46.00-47.00 sec 156 MBytes 1311 Mbits/sec 436 12.3 MBytes
[ 15] 47.00-48.00 sec 189 MBytes 1583 Mbits/sec 1896 15.0 MBytes
[ 15] 48.00-49.00 sec 151 MBytes 1269 Mbits/sec 1617 10.3 MBytes
[ 15] 49.00-50.00 sec 116 MBytes 975 Mbits/sec 1103 7.84 MBytes
[ 15] 50.00-51.00 sec 115 MBytes 965 Mbits/sec 750 9.26 MBytes
[ 15] 51.00-52.00 sec 108 MBytes 902 Mbits/sec 1408 6.59 MBytes
[ 15] 52.00-53.00 sec 108 MBytes 902 Mbits/sec 1051 7.56 MBytes
[ 15] 53.00-54.00 sec 121 MBytes 1017 Mbits/sec 1296 8.35 MBytes
[ 15] 54.00-55.00 sec 118 MBytes 986 Mbits/sec 1525 8.33 MBytes
[ 15] 55.00-56.00 sec 104 MBytes 870 Mbits/sec 1216 8.76 MBytes
[ 15] 56.00-57.00 sec 144 MBytes 1206 Mbits/sec 1030 8.77 MBytes
[ 15] 57.00-58.00 sec 129 MBytes 1080 Mbits/sec 1295 7.89 MBytes
[ 15] 58.00-59.00 sec 125 MBytes 1049 Mbits/sec 1182 10.4 MBytes
[ 15] 59.00-60.00 sec 132 MBytes 1111 Mbits/sec 1358 8.82 MBytes
-----
[ ID] Interval          Transfer      Bandwidth      Retr
[ 15] 0.00-60.00 sec 7.07 GBytes 1013 Mbits/sec 65068
[ 15] 0.00-60.00 sec 7.03 GBytes 1007 Mbits/sec

```

iperf Done.

SENDER END

In summary – “the ball game is over”. The high frequency of out of order packets, and the long delay they experience, wrecks TCP. The hosts spend more time recovering than doing useful work, and thus reflect a lower performance as a result.

At this point the addition of more delay, or higher frequencies of out of ordering will only show poor performance.

Conclusions

Delay is a significant factor in the operation of TCP. We know this because as the delay increases, we must either devote more memory to our sliding window to keep our bandwidth high, or ‘plateau’ on performance expectation.

Delay, when coupled with a higher probability of an out of order packet, kills TCP performance. This delay is a function of the complete RTT between hosts, and triggers internal TCP algorithm behavior.

As we increase the probability of an out of ordering happening, we see problems. As the delay increases when a missing packet may arrive, we see problems. Recovery is possible up to a certain limit.

The exercises above can describe many real world situations:

- Traversal of a device that may paralyze the process of processing a single TCP stream (a firewall, packet shapper)
- Traverseal of a device that may split traffic between links that are viewed as being equivalent (e.g. a bonded link on a host, a LAG on a network device)
- Queing somewhere in the path due to processing delays
- Network asymmetry and adaptative protocols to balance load

These results are experienced through simulation, and may be slightly different in practice.