           Simulation Studies of Increased Initial TCP Window Size

Status of this Memo

Copyright Notice

Abstract

   An increase in the permissible initial window size of a TCP
   connection, from one segment to three or four segments, has been
   under discussion in the tcp-impl working group. This document covers
   some simulation studies of the effects of increasing the initial
   window size of TCP. Both long-lived TCP connections (file transfers)
   and short-lived web-browsing style connections were modeled. The
   simulations were performed using the publicly available ns-2
   simulator and our custom models and files are also available.
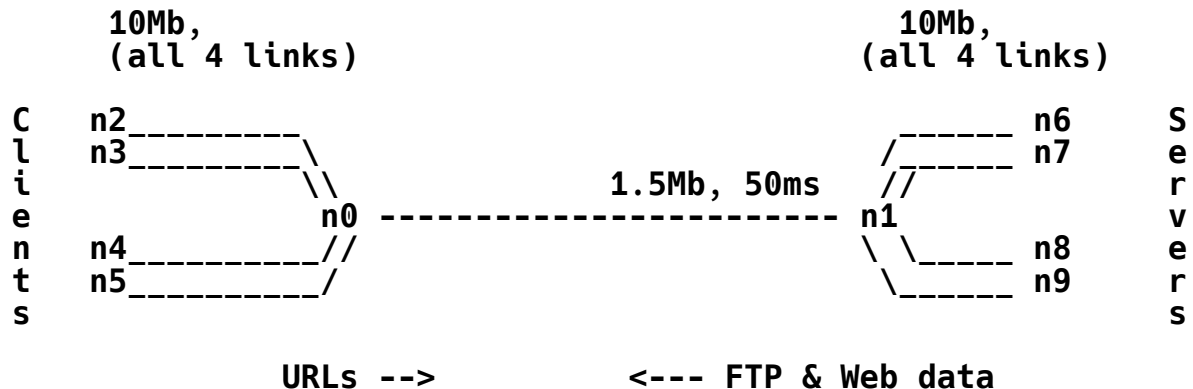
1. Introduction

   We present results from a set of simulations with increased TCP
   initial window (IW). The main objectives were to explore the
   conditions under which the larger IW was a "win" and to determine the
   effects, if any, the larger IW might have on other traffic flows
   using an IW of one segment.

   This study was inspired by discussions at the Munich IETF tcp-impl
   and tcp-sat meetings. A proposal to increase the IW size to about 4K
   bytes (4380 bytes in the case of 1460 byte segments) was discussed.
   Concerns about both the utility of the increase and its effect on
   other traffic were raised. Some studies were presented showing the
   positive effects of increased IW on individual connections, but no
   studies were shown with a wide variety of simultaneous traffic flows.
   It appeared that some of the questions being raised could be
   addressed in an ns-2 simulation. Early results from our simulations
   were previously posted to the tcp-impl mailing list and presented at
   the tcp-impl WG meeting at the December 1997 IETF.

## 2. Model and Assumptions

We simulated a network topology with a bottleneck link as shown:

```
     10Mb,                                      10Mb,
     (all 4 links)                             (all 4 links)

  C   n2_____                        _____ n6    S
  l   n3_____\                      /_____ n7    e
  i            \\         1.5Mb, 50ms  //               r
  e            n0 ---------------------- n1             v
  n   n4_____//                    \ \_____ n8      e
  t   n5_____/                      _____ n9      r
  s                                                     s

        URLs -->              <--- FTP & Web data
```

File downloading and web-browsing clients are attached to the nodes (n2-n5) on the left-hand side. These clients are served by the FTP and Web servers attached to the nodes (n6-n9) on the right-hand side. The links to and from those nodes are at 10 Mbps. The bottleneck link is between n1 and n0. All links are bi-directional, but only ACKs, SYNs, FINs, and URLs are flowing from left to right. Some simulations were also performed with data traffic flowing from right to left simultaneously, but it had no effect on the results.

In the simulations we assumed that all ftps transferred 1-MB files and that all web pages had exactly three embedded URLs. The web clients are browsing quite aggressively, requesting a new page after a random delay uniformly distributed between 1 and 5 seconds. This is not meant to realistically model a single user's web-browsing pattern, but to create a reasonably heavy traffic load whose individual tcp connections accurately reflect real web traffic. Some discussion of these models as used in earlier studies is available in references [3] and [4].

The maximum tcp window was set to 11 packets, maximum packet (or segment) size to 1460 bytes, and buffer sizes were set at 25 packets. (The ns-2 TCPs require setting window sizes and buffer sizes in number of packets. In our tcp-full code some of the internal parameters have been set to be byte-oriented, but external values must still be set in number of packets.)  In our simulations, we varied the number of data segments sent into a new TCP connection (or initial window) from one to four, keeping all segments at 1460 bytes. A dropped packet causes a restart window of one segment to be used, just as in current practice.

For ns-2 users: The tcp-full code was modified to use an
"application" class and three application client-server pairs were
written: a simple file transfer (ftp), a model of http1.0 style web
connection and a very rough model of http1.1 style web connection.
The required files and scripts for these simulations are available
under the contributed code section on the ns-simulator web page at
the sites ftp://ftp.ee.lbl.gov/IW.{tar, tar.Z} or http://www-
nrg.ee.lbl.gov/floyd/tcp_init_win.html.

Simulations were run with 8, 16, 32 web clients and a number of ftp
clients ranging from 0 to 3. The IW was varied from 1 to 4, though
the 4-packet case lies beyond what is currently recommended. The
figures of merit used were goodput, the median page delay seen by the
web clients and the median file transfer delay seen by the ftp
clients. The simulated run time was rather large, 360 seconds, to
ensure an adequate sample. (Median values remained the same for
simulations with larger run times and can be considered stable)

## 3. Results

In our simulations, we varied the number of file transfer clients in
order to change the congestion of the link. Recall that our ftp
clients continuously request 1 Mbyte transfers, so the link
utilization is over 90% when even a single ftp client is present.
When three file transfer clients are running simultaneously, the
resultant congestion is somewhat pathological, making the values
recorded stable. Though all connections use the same initial window,
the effect of increasing the IW on a 1 Mbyte file transfer is not
detectable, thus we focus on the web browsing connections.  (In the
tables, we use "webs" to indicate number of web clients and "ftps" to
indicate the number of file transfer clients attached.) Table 1 shows
the median delays experienced by the web transfers with an increase
in the TCP IW.  There is clearly an improvement in transfer delays
for the web connections with increase in the IW, in many cases on the
order of 30%.  The steepness of the performance improvement going
from an IW of 1 to an IW of 2 is mainly due to the distribution of
files fetched by each URL (see references [1] and [2]); the median
size of both primary and in-line URLs fits completely into two
packets. If file distributions change, the shape of this curve may
also change.

Table 1. Median web page delay

| #Webs | #FTPs | IW=1 (s) | IW=2 | IW=3 (% decrease) | IW=4 |
|-------|-------|----------|------|------|------|
| 8 | 0 | 0.56 | 14.3 | 17.9 | 16.1 |
| 8 | 1 | 1.06 | 18.9 | 25.5 | 32.1 |
| 8 | 2 | 1.18 | 16.1 | 17.1 | 28.9 |
| 8 | 3 | 1.26 | 11.9 | 19.0 | 27.0 |
| 16 | 0 | 0.64 | 11.0 | 15.6 | 18.8 |
| 16 | 1 | 1.04 | 17.3 | 24.0 | 35.6 |
| 16 | 2 | 1.22 | 17.2 | 20.5 | 25.4 |
| 16 | 3 | 1.31 | 10.7 | 21.4 | 22.1 |
| 32 | 0 | 0.92 | 17.6 | 28.6 | 21.0 |
| 32 | 1 | 1.19 | 19.6 | 25.0 | 26.1 |
| 32 | 2 | 1.43 | 23.8 | 35.0 | 33.6 |
| 32 | 3 | 1.56 | 19.2 | 29.5 | 33.3 |

Table 2 shows the bottleneck link utilization and packet drop
percentage of the same experiment. Packet drop rates did increase
with IW, but in all cases except that of the single most pathological
overload, the increase in drop percentage was less than 1%. A
decrease in packet drop percentage is observed in some overloaded
situations, specifically when ftp transfers consumed most of the link
bandwidth and a large number of web transfers shared the remaining
bandwidth of the link. In this case, the web transfers experience
severe packet loss and some of the IW=4 web clients suffer multiple
packet losses from the same window, resulting in longer recovery
times than when there is a single packet loss in a window. During the
recovery time, the connections are inactive which alleviates
congestion and thus results in a decrease in the packet drop
percentage. It should be noted that such observations were made only
in extremely overloaded scenarios.

Table 2. Link utilization and packet drop rates

|        |        | Percentage Link Utilization | | | | Packet drop rate | | | |
|--------|--------|------|------|------|------|------|------|------|------|
| #Webs  | #FTPs  | IW=1 | IW=2 | IW=3 | IW=4 | IW=1 | IW=2 | IW=3 | IW=4 |
| 8      | 0      | 34   | 37   | 38   | 39   | 0.0  | 0.0  | 0.0  | 0.0  |
| 8      | 1      | 95   | 92   | 93   | 92   | 0.6  | 1.2  | 1.4  | 1.3  |
| 8      | 2      | 98   | 97   | 97   | 96   | 1.8  | 2.3  | 2.3  | 2.7  |
| 8      | 3      | 98   | 98   | 98   | 98   | 2.6  | 3.0  | 3.5  | 3.5  |
| 16     | 0      | 67   | 69   | 69   | 67   | 0.1  | 0.5  | 0.8  | 1.0  |
| 16     | 1      | 96   | 95   | 93   | 92   | 2.1  | 2.6  | 2.9  | 2.9  |
| 16     | 2      | 98   | 98   | 97   | 96   | 3.5  | 3.6  | 4.2  | 4.5  |
| 16     | 3      | 99   | 99   | 98   | 98   | 4.5  | 4.7  | 5.2  | 4.9  |
| 32     | 0      | 92   | 87   | 85   | 84   | 0.1  | 0.5  | 0.8  | 1.0  |
| 32     | 1      | 98   | 97   | 96   | 96   | 2.1  | 2.6  | 2.9  | 2.9  |
| 32     | 2      | 99   | 99   | 98   | 98   | 3.5  | 3.6  | 4.2  | 4.5  |
| 32     | 3      | 100  | 99   | 99   | 98   | 9.3  | 8.4  | 7.7  | 7.6  |

To get a more complete picture of performance, we computed the
network power, goodput divided by median delay (in Mbytes/ms), and
plotted it against IW for all scenarios. (Each scenario is uniquely
identified by its number of webs and number of file transfers.) We
plot these values in Figure 1 (in the pdf version), illustrating a
general advantage to increasing IW. When a large number of web
clients is combined with ftps, particularly multiple ftps,
pathological cases result from the extreme congestion. In these
cases, there appears to be no particular trend to the results of
increasing the IW, in fact simulation results are not particularly
stable.

To get a clearer picture of what is happening across all the tested
scenarios, we normalized the network power values for the non-
pathological scenario by the network power for that scenario at IW of
one. These results are plotted in Figure 2. As IW is increased from
one to four, network power increased by at least 15%, even in a
congested scenario dominated by bulk transfer traffic. In simulations
where web traffic has a dominant share of the available bandwidth,
the increase in network power was up to 60%.

The increase in network power at higher initial window sizes is due
to an increase in throughput and a decrease in the delay. Since the
(slightly) increased drop rates were accompanied by better
performance, drop rate is clearly not an indicator of user level
performance.

The gains in performance seen by the web clients need to be balanced
against the performance the file transfers are seeing. We computed
ftp network power and show this in Table 3.  It appears that the
improvement in network power seen by the web connections has
negligible effect on the concurrent file transfers. It can be
observed from the table that there is a small variation in the
network power of file transfers with an increase in the size of IW
but no particular trend can be seen. It can be concluded that the
network power of file transfers essentially remained the same.
However, it should be noted that a larger IW does allow web transfers
to gain slightly more bandwidth than with a smaller IW. This could
mean fewer bytes transferred for FTP applications or a slight
decrease in network power as computed by us.

Table 3. Network power of file transfers with an increase in the TCP
         IW size

| #Webs | #FTPs | IW=1 | IW=2 | IW=3 | IW=4 |
|-------|-------|------|------|------|------|
| 8     | 1     | 4.7  | 4.2  | 4.2  | 4.2  |
| 8     | 2     | 3.0  | 2.8  | 3.0  | 2.8  |
| 8     | 3     | 2.2  | 2.2  | 2.2  | 2.2  |
| 16    | 1     | 2.3  | 2.4  | 2.4  | 2.5  |
| 16    | 2     | 1.8  | 2.0  | 1.8  | 1.9  |
| 16    | 3     | 1.4  | 1.6  | 1.5  | 1.7  |
| 32    | 1     | 0.7  | 0.9  | 1.3  | 0.9  |
| 32    | 2     | 0.8  | 1.0  | 1.3  | 1.1  |
| 32    | 3     | 0.7  | 1.0  | 1.2  | 1.0  |

The above simulations all used http1.0 style web connections, thus, a
natural question is to ask how results are affected by migration to
http1.1. A rough model of this behavior was simulated by using one
connection to send all of the information from both the primary URL
and the three embedded, or in-line, URLs. Since the transfer size is
now made up of four web files, the steep improvement in performance
between an IW of 1 and an IW of two, noted in the previous results,
has been smoothed. Results are shown in Tables 4 & 5 and Figs. 3 & 4.
Occasionally an increase in IW from 3 to 4 decreases the network
power owing to a non-increase or a slight decrease in the throughput.
TCP connections opening up with a higher window size into a very
congested network might experience some packet drops and consequently
a slight decrease in the throughput. This indicates that increase of
the initial window sizes to further higher values (>4) may not always
result in a favorable network performance. This can be seen clearly
in Figure 4 where the network power shows a decrease for the two
highly congested cases.

Table 4. Median web page delay for http1.1

| #Webs | #FTPs | IW=1 (s) | IW=2 | IW=3 (% decrease) | IW=4 |
|-------|-------|----------|------|-------------------|------|
| 8     | 0     | 0.47     | 14.9 | 19.1              | 21.3 |
| 8     | 1     | 0.84     | 17.9 | 19.0              | 25.0 |
| 8     | 2     | 0.99     | 11.5 | 17.3              | 23.0 |
| 8     | 3     | 1.04     | 12.1 | 20.2              | 28.3 |
| 16    | 0     | 0.54     | 07.4 | 14.8              | 20.4 |
| 16    | 1     | 0.89     | 14.6 | 21.3              | 27.0 |
| 16    | 2     | 1.02     | 14.7 | 19.6              | 25.5 |
| 16    | 3     | 1.11     | 09.0 | 17.0              | 18.9 |
| 32    | 0     | 0.94     | 16.0 | 29.8              | 36.2 |
| 32    | 1     | 1.23     | 12.2 | 28.5              | 21.1 |
| 32    | 2     | 1.39     | 06.5 | 13.7              | 12.2 |
| 32    | 3     | 1.46     | 04.0 | 11.0              | 15.0 |

Table 5. Network power of file transfers with an increase in the
         TCP IW size

| #Webs | #FTPs | IW=1 | IW=2 | IW=3 | IW=4 |
|-------|-------|------|------|------|------|
| 8     | 1     | 4.2  | 4.2  | 4.2  | 3.7  |
| 8     | 2     | 2.7  | 2.5  | 2.6  | 2.3  |
| 8     | 3     | 2.1  | 1.9  | 2.0  | 2.0  |
| 16    | 1     | 1.8  | 1.8  | 1.5  | 1.4  |
| 16    | 2     | 1.5  | 1.2  | 1.1  | 1.5  |
| 16    | 3     | 1.0  | 1.0  | 1.0  | 1.0  |
| 32    | 1     | 0.3  | 0.3  | 0.5  | 0.3  |
| 32    | 2     | 0.4  | 0.3  | 0.4  | 0.4  |
| 32    | 3     | 0.4  | 0.3  | 0.4  | 0.5  |

For further insight, we returned to the http1.0 model and mixed some
web-browsing connections with IWs of one with those using IWs of
three. In this experiment, we first simulated a total of 16 web-
browsing connections, all using IW of one. Then the clients were
split into two groups of 8 each, one of which uses IW=1 and the other
used IW=3.

We repeated the simulations for a total of 32 and 64 web-browsing
clients, splitting those into groups of 16 and 32 respectively. Table
6 shows these results.  We report the goodput (in Mbytes), the web
page delays (in milli seconds), the percent utilization of the link
and the percent of packets dropped.

Table 6. Results for half-and-half scenario

| Median Page Delays and Goodput (MB) | | | | Link Utilization (%) & Drops (%) | | | |
|---|---|---|---|---|---|---|---|
| #Webs | IW=1 | | IW=3 | | IW=1 | | IW=3 | |

| #Webs | G.put | dly | G.put | dly | L.util | Drops | L.util | Drops |
|---|---|---|---|---|---|---|---|---|
| 16 | 35.5 | 0.64 | 36.4 | 0.54 | 67 | 0.1 | 69 | 0.7 |
| 8/8 | 16.9 | 0.67 | 18.9 | 0.52 | 68 | 0.5 | | |
| 32 | 48.9 | 0.91 | 44.7 | 0.68 | 92 | 3.5 | 85 | 4.3 |
| 16/16 | 22.8 | 0.94 | 22.9 | 0.71 | 89 | 4.6 | | |
| 64 | 51.9 | 1.50 | 47.6 | 0.86 | 98 | 13.0 | 91 | 8.6 |
| 32/32 | 29.0 | 1.40 | 22.0 | 1.20 | 98 | 12.0 | | |

   Unsurprisingly, the non-split experiments are consistent with our
   earlier results, clients with IW=3 outperform clients with IW=1. The
   results of running the 8/8 and 16/16 splits show that running a
   mixture of IW=3 and IW=1 has no negative effect on the IW=1
   conversations, while IW=3 conversations maintain their performance.
   However, the 32/32 split shows that web-browsing connections with
   IW=3 are adversely affected. We believe this is due to the
   pathological dynamics of this extremely congested scenario. Since
   embedded URLs open their connections simultaneously, very large
   number of TCP connections are arriving at the bottleneck link
   resulting in multiple packet losses for the IW=3 conversations. The
   myriad problems of this simultaneous opening strategy is, of course,
   part of the motivation for the development of http1.1.

4. Discussion

   The indications from these results are that increasing the initial
   window size to 3 packets (or 4380 bytes) helps to improve perceived
   performance. Many further variations on these simulation scenarios
   are possible and we've made our simulation models and scripts
   available in order to facilitate others' experiments.

   We also used the RED queue management included with ns-2 to perform
   some other simulation studies. We have not reported on those results
   here since we don't consider the studies complete. We found that by
   adding RED to the bottleneck link, we achieved similar performance
   gains (with an IW of 1) to those we found with increased IWs without
   RED. Others may wish to investigate this further.

   Although the simulation sets were run for a T1 link, several
   scenarios with varying levels of congestion and varying number of web
   and ftp clients were analyzed. It is reasonable to expect that the
   results would scale for links with higher bandwidth. However,

interested readers could investigate this aspect further.

We also used the RED queue management included with ns-2 to perform some other simulation studies. We have not reported on those results here since we don't consider the studies complete. We found that by adding RED to the bottleneck link, we achieved similar performance gains (with an IW of 1) to those we found with increased IWs without RED. Others may wish to investigate this further.

## 5. References

[1] B. Mah, "An Empirical Model of HTTP Network Traffic", Proceedings of INFOCOM '97, Kobe, Japan, April 7-11, 1997.

[2] C.R. Cunha, A. Bestavros, M.E. Crovella, "Characteristics of WWW Client-based Traces", Boston University Computer Science Technical Report BU-CS-95-010, July 18, 1995.

[3] K.M. Nichols and M. Laubach, "Tiers of Service for Data Access in a HFC Architecture", Proceedings of SCTE Convergence Conference, January, 1997.

[4] K.M. Nichols, "Improving Network Simulation with Feedback", available from knichols@baynetworks.com

## 6. Acknowledgements

This work benefited from discussions with and comments from Van Jacobson.

## 7. Security Considerations

This document discusses a simulation study of the effects of a proposed change to TCP. Consequently, there are no security considerations directly related to the document. There are also no known security considerations associated with the proposed change.

## 8. Authors' Addresses

Kedarnath Poduri
Bay Networks
4401 Great America Parkway
SC01-04
Santa Clara, CA 95052-8185

Phone: +1-408-495-2463
Fax:   +1-408-495-1299
EMail: kpoduri@Baynetworks.com


Kathleen Nichols
Bay Networks
4401 Great America Parkway
SC01-04
Santa Clara, CA 95052-8185

EMail: knichols@baynetworks.com

Full Copyright Statement