Effect of IW and Initial RTO changes

Ilpo Järvinen, Aki Nyrhinen, Aaron Yi Ding, Markku Kojo

Department of Computer Science
University of Helsinki

IETF79 / Beijing
Nov 11th 2010
Simulation study to evaluate effects of recently proposed changes:
- Initial Window change from 3 packets to 10 packets
- Initial RTO change from 3 seconds to 1 second

Focus on (typical) slow/moderate bit-rate wireless links like environments

Initially presented IW10 results in the last ICCRG meeting @ Maastricht
Test setup

- Links (bw/one-way propagation delay)
  - EGDE 160kbps/250ms, BDP = 7 pkts (6.7)
  - HSPA 2Mbps/70ms, BDP = 24 pkts (23.3)
  - LTE 50Mbps/15ms, BDP = 125 pkts

- No wireless errors, nor allocation / error related delays considered

- 11ms propagation delay from sender to wireless link

- Buffer (FIFO) sizes
  - BDP (Bandwidth Delay Product)
  - 2 \cdot BDP
  - 50 Packets (EDGE only)

- Workload: A burst of 1, 2, 6 or 18 simultaneous downstream TCP flows (total 180kB) competing against a similar later starting burst (another 180kB), 100 replications

- ns2 TCP SACK in use
Summary of IW10 Effects

- With small number of TCP flows, IW10 improves performance
- With larger number of flows, IW10 tends to decrease performance - Regardless of IW, too many flows clearly results in suboptimal performance
- Fairness for later starting traffic improves with IW10
- Fairness within both bursts worse with IW10
No changes
No spurious RTOs
RTOs with IW10 when \# of flows is 6+6 or 18+18
  But not in the beginning for the flow that completes last (not for the SYN nor the first packet)
  ⇒ IRTO has no effect
IRTO: HSPA (2Mbps/70ms, BDP≈23 Packets)

Observations

- When overloaded, small improvement for the longest cases among later starting traffic
- Opposite effect for the first starting burst (the shortest cases delayed)
- No changes due to IRTO1 with 1+1 or 2+2 flows
IRTO: EDGE (160kbps/250ms, BDP≈7 Packets)

Observations

- With large buffer, number of RTOs increase
  - Mostly spurious RTOs
  - Completion of the longest flow is delayed
- The same trend with larger number of flows
- When IW10 in use, the first starting burst is able to take advantage and completes unfairly early
IRT0: EDGE (160kbps/250ms, BDP\(\approx\)7 Packets)

**Observations**

- Mostly the same regardless of IRT0
- IW10+IRT01 becomes more fair
  - RTO occurred sooner for the later starting burst (a spurious one)
### RED Configuration

<table>
<thead>
<tr>
<th>Cfg</th>
<th>RED</th>
<th>REDok</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link</td>
<td>EDGE</td>
<td>HSPA/LTE</td>
</tr>
<tr>
<td>$w_q$</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>$\text{max}_p$</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$\text{th}_{\text{min}}$</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>$\text{th}_{\text{max}}$</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>Buffer size</td>
<td>$2 \cdot BDP$</td>
<td>$2 \cdot BDP$</td>
</tr>
</tbody>
</table>

- Large buffers with *RED* configuration were not tested
- Not useful because of $\text{avg} > \text{th}_{\text{max}}$ dropper
- *REDok* config aimed to highly varying load
  - Thus vastly different from “default configuration”
  - Aggressive enough to respond to slow start
  - Parameters are link characteristics dependent
IW10 slightly more aggressive
RED similar to FIFO behavior (too slow to react)
With BDP IW10 hurts itself due to self-congestion
Slightly smaller delays except for the highest end
Also REDok fails to control the delay increase IW10 imposes

Minimum values with REDok:
- IW10: 2.80s
- IW3: 2.06s
Again, RED reacts too slowly

IW10 less aggressive due to self-congestion ⇒ more bursty
With REDok, traffic regulation works without heavy tail-drop
⇒ IW10 shows to be significantly more aggressive

Maximum values with REDok:
- IW10: 0.429s
- IW3: 0.296s
Similar behavior observed:

- Self-congestion $\Rightarrow$ IW10 is less aggressive
  - Except for the very highest end (in some of the cases)
  - With low enough load, IW10 is slightly more aggressive
- IRT01 only slightly “shifts” curves
  - Only happening when IRT01 has some effect in the first place
  - Quite insignificant in numbers
- Actual shape of the delay curves vary per queue size and type, however, those differences are out of scope here
Conclusions

- Smaller initial RTO performs better when effective e2e RTT smaller than 1 second
- More controversial when e2e RTT is larger
- IW10, while improving elapsed times, imposes higher queuing delay than IW3
  - However, if self-congesting, IW3 is more aggressive in terms of queuing delay
  - AQM (RED) failed to control the increase in the queuing delay
Questions?
Queue/RED set bytes_ true
Queue/RED set queue_in_bytes_ true
Queue/RED set gentle_ false
Queue/RED set setbit_ false
Queue/RED set use_mark_p_ false
Queue/RED set mean_pktsize_ 1500
Queue/RED set idle_pktsize_ 1500
Queue/RED set q_weight_ $wq
Queue/RED set thresh_ $minth
Queue/RED set maxthresh_ $maxth
Queue/RED set linterm_ [expr 1.0/$maxp]
Queue/RED set wait_ false
6 Flows Elapsed Times (FIFO, RED, REDok and IW)

HSPA, Later Starting Burst

Elapsed time of the longest flow (s)
(median, quartiles, min-max)

1+1 flows
2+2 flows
6+6 flows
18+18 flows

IETF79 / Beijing Nov 11th 2010
Jain's fairness index between the bursts (median, quartiles, min-max)