TCP Performance

These slides are created by Dr. Yih Huang of George Mason University. Students registered in Dr. Huang's courses at GMU can make a single machine-readable copy and print a single copy of each slide for their own reference, so long as each slide contains the copyright statement, and GMU facilities are not used to produce paper copies. Permission for any other use, either in machine-readable or printed form, must be obtained from the author in writing.

TCP Responsibilities in Congestion Control

- A TCP sources see two types of **loss indications**:
  - Triple duplicate ACKs (**TD**)
  - Time-outs (**TO**)
- A TD event cuts `cwnd` by half.
- A TO event sets `cwnd` to 1.
- With smaller window sizes, the source must "stop and wait" frequently and thus reduce traffic rate.

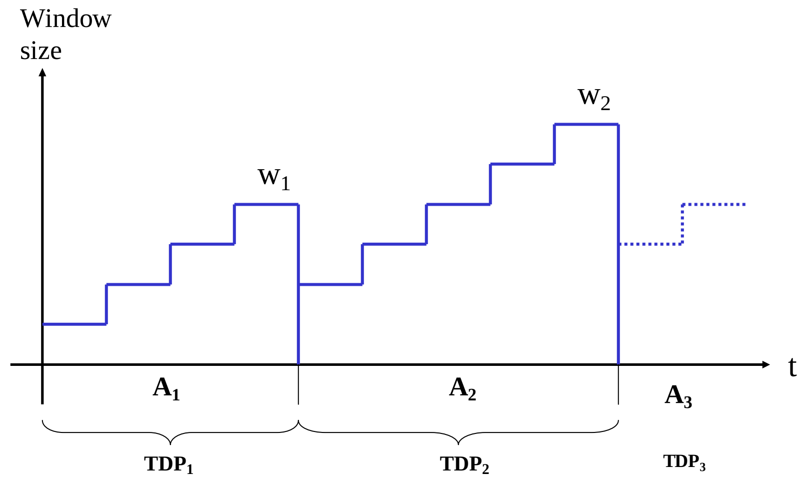
Assumptions/Simplifications

- The time needed to send all packets in a window is smaller than the round trip time.
- When a packet loss to the *k*-th packet in a round, the rest of the packets in that round are lost too.
  - This is largely due to the FIFO queueing of routers.
- These behaviors are generally, but not always, observed in the real world.

Analysis of TD-Only Scenarios

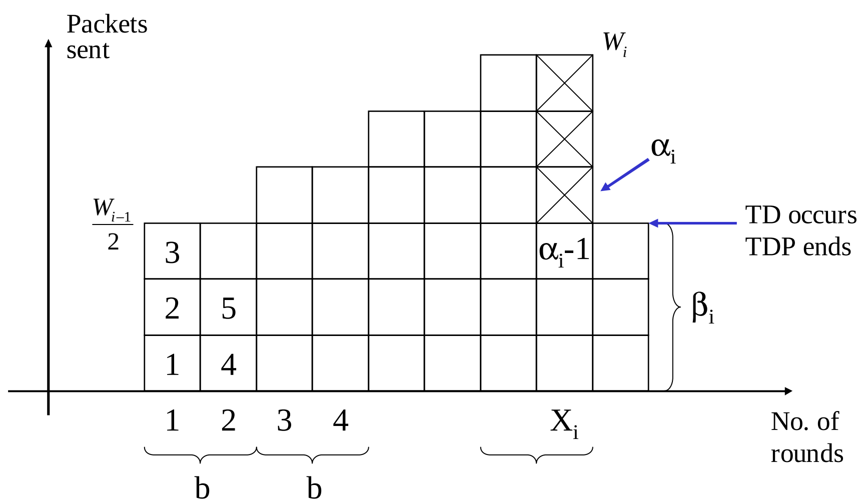
- TD Period (**TDP**): a period between two triple-duplicates indications.
- A_i : the duration of the *i*-th TD period, TDP_i
- Y_i : the no. of packets sent in TDP_i
- W_i : the window size (`cwnd`) at the end of TDP_i
- b : the no. of packets ack-ed per ACK.
  - In many TCP implementations, b=2.
- Our goal: throughput $B = E[Y] / E[A]$

Evolution of Window Size



CS 756 5

Packets Sent during a TD Period



CS 756 6

- A TD period starts immediately after a TD loss indication and thus `cwnd` is $W_{i-1}/2$.
- At each round, the window is incremented by $1/b$ and the no. of packets sent per round is incremented by 1 every b rounds.
- Let α_i be the first packet lost in TDP_i and X_i the round where this loss occurs.
- After packet α_i, $W_i - 1$ more packets are sent.

- We have $Y_i = \alpha_i + W_i - 1$
- And thus $E[Y] = E[\alpha] + E[W] - 1$
- Let p be the probability of packet loss.

$$P[\alpha = k] = (1-p)^{k-1} p \quad \Rightarrow \quad E[\alpha] = \frac{1}{p}$$

- It follows that

$$E[Y] = \frac{1-p}{p} + E[W]$$

- Next, we must figure out E[W] and E[A].

- $r_{i,j}$: the duration (round trip time) of the *j*-th round in TDP_i.
- The duration of TDP_i is $A_i = \sum_{j=1}^{X_i+1} r_{i,j}$
- We consider $r_{i,j}$ be a random variable independent of the size of congestion window, and thus independent of *i* and *j*.
- It follows that $E[A] = (E[X]+1)RTT$ where *RTT*=E[*r*] is the average round trip time.

- During TDP_i , the window size increases between $W_{i-1}/2$ and W_i linearly with slope 1/*b*, that is,

$$W_i = \frac{W_{i-1}}{2} + \frac{X_i}{b} \quad \Rightarrow \quad E[W] = \frac{2}{b} E[X]$$

- The fact that Y_i packets are transmitted in TDP_i is expressed by

$$Y_i = \sum_{k=0}^{X_i/b-1} (\frac{W_{i-1}}{2} + k)b + \beta_i$$

$$= \frac{X_i W_{i-1}}{2} + \frac{X_i}{2}(\frac{X_i}{b} - 1) + \beta_i$$

$$= \frac{X_i}{2}(\frac{W_{i-1}}{2} + W_i - 1) + \beta_i$$

Average Window Size

- Thus $\frac{1-p}{p} + E[W] = \frac{E[X]}{2}(\frac{E[W]}{2} + E[W] - 1) + E[\beta]$
- Assuming β_i be uniformly distributed 1 and W_i, and thus E[β]=E[W]/2, we have

$$E[W] = \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2}$$

- For small values of p,

$$E[W] \approx \sqrt{\frac{8}{3bp}}$$

Average TCP Throughput

- It follows that

$$E[X] = \frac{2+b}{6} + \sqrt{\frac{2b(1-p)}{3p} + \left(\frac{2+b}{6}\right)^2}$$

$$B(p) = \frac{E[Y]}{E[A]} = \frac{\frac{1-p}{p} + E[W]}{E[A]} = \frac{\frac{1-p}{p} + \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2}}{RTT\left(\frac{2+b}{6} + \sqrt{\frac{2b(1-p)}{3p} + \left(\frac{2+b}{6}\right)^2} + 1\right)}$$

- For small p,

$$B(p) \approx \frac{1}{RTT}\sqrt{\frac{3}{2bp}}$$

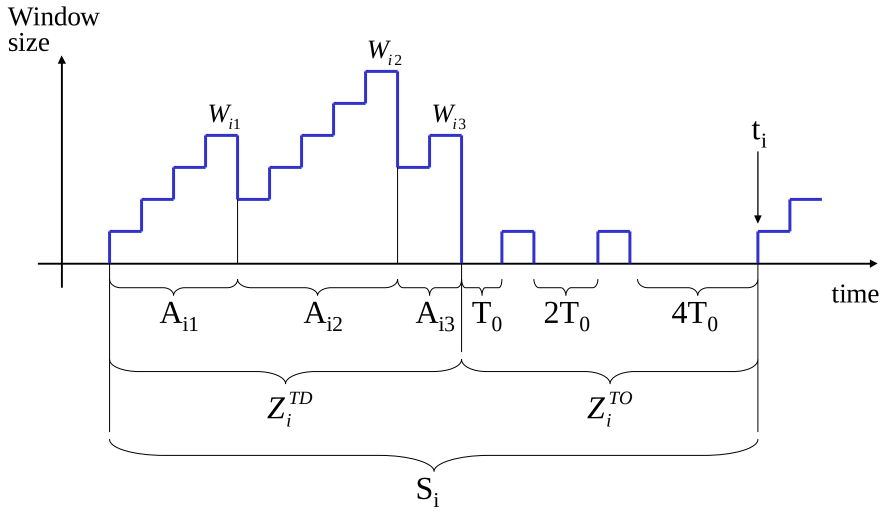
Discussion

- TCP favors flows with short RTT.
  - Find a server near you for downloading
- The relationship between packet loss rate *p* and throughput is not linear.
  - When *p* is increased 4 times, throughput drops to half.
- Large *b* values is bad for throughput.
  - Hence *b*=2 in common implementations
  - Why not *b*=1 ?

Taking Timeouts into Account

- Timeout (TO): the timer of a missing ACK fires before three duplicate ACKs are received.
- The initial TO period is denoted as T_0.
- After a TO, `cwnd` is reduced to 1, allowing for the retransmission only for the lost packet.
- If the retransmission fails (another TO), the TO period is set to $2T_0$.
- If the retransmission fails a second time, the TO period is set to $4T_0$.
- The maximum TO period is 64 T_0.

Evolution of Window Size



Notations

- Z_i^{TO} : the duration of a sequence of TO.
- Z_i^{TD} : the time interval between two consecutive TO sequences.
- Define $S_i = Z_i^{TD} + Z_i^{TO}$
- M_i: the no. of packets sent during S_i.
- n_i: the no. of TD periods in interval Z_i^{TD}
- Y_{ij}: the no. of packets sent in the j-th TDP in Z_i^{TD}
- A_{ij}: the duration of TDP_{ij}
- X_{ij}: the no. of rounds in TDP_{ij}
- W_{ij}: the window size at the end of TDP_{ij}
- R_i: the no. of packets sent in Z_i^{TO}

- Notice that the “packet counts” (Y_{ij}, M_i and R_i) includes retransmissions and our results are for **throughput**, not **goodput**. We have

$$M_i = \sum_{j=1}^{n_i} Y_{ij} + R_i \quad \Rightarrow \quad E[M] = E\left[\sum_{j=1}^{n_i} Y_{ij}\right] + E[R]$$

$$S_i = \sum_{j=1}^{n_i} A_{ij} + Z_i^{TO} \quad \Rightarrow \quad E[S] = E\left[\sum_{j=1}^{n_i} A_{ij}\right] + E[Z^{TO}]$$

- Assuming that n_i is an independent sequence of random variable and independent of Y_{ij} and A_{ij}:

$$E\left[\sum_{j=1}^{n_i} Y_{ij}\right] = E[n]\times E[Y] \quad \text{and} \quad E\left[\sum_{j=1}^{n_i} A_{ij}\right] = E[n]\times E[A]$$

- Let Q be the probability that a TDP ends with a TO. We have $Q=1/E[n]$.
- The throughput can be expressed as

$$B = \frac{E[Y] + Q\times E[R]}{E[A] + Q\times E[Z^{TO}]}$$

- Since Y_{ij} and A_{ij} do not depend on TO, we can use previous results of $E[Y]$ and $E[A]$.
- We still need to figure out Q, $E[R]$, and $E[Z^{TO}]$.

- In the penultimate round, packet k+1 is lost.
- The probability that the first k packets are ACKed in a round of w packets, given there is one or more losses in the round is

$$A(w,k) = \frac{(1-p)^k p}{1-(1-p)^w}$$

- The probability that n packets are sent and m of them are acknowledged in the last round is

$$C(n,m) = \begin{cases} (1-p)^m p, & m < n \\ (1-p)^n, & m = n \end{cases}$$

- In the last round, the probability a loss in a window of size w causes TO is given by

$$Q'(w) = \begin{cases} 1 & w \le 3 \\ \sum_{k=0}^{2} A(w,k) + \sum_{k=3}^{w} A(w,k) \sum_{m=0}^{2} C(k,m) & w > 3 \end{cases}$$

$\sum_{k=0}^{2} A(w,k)$: < 3 packets successfully sent in the penultimate round, and thus < 3 sent in the last round. No way to produce 3 duplicate ACKs

$\sum_{k=3}^{w} A(w,k)$: ≥ 3 packets successfully sent in the penultimate round, allowing ≥ 3 sent in the last round

$\sum_{m=0}^{2} C(k,m)$: But < 3 last round packets get thru to cause duplicate ACKs. No way to produce 3 duplicate ACKs

Solving *Q*

- After some fun with algebra, we have

$$Q'(w) = \min\left(1, \frac{(1-(1-p)^3)(1+(1-p)^3(1-(1-p)^{w-3}))}{(1-(1-p)^w}\right)$$

- *Q*, the probability a TO occurs at the end of a TDP is $Q'(E[W])$, where $E[W]$ has been solved previously.

Solving *E*[*R*]

- Next, we find $E[R]$, the average no. of packets sent during an Z^{TO}.
- In a Z^{TO} there are *k*-1 consecutive losses followed by a successful transmission, that is,

$$P[R=k] = p^k(1-p)$$

- Thus, $E[R] = \dfrac{1}{1-p}$

Solving $E[Z^{TO}]$

- The first six TO have length $2^{i-1}T_0, \quad i = 1 \ldots 6$
- Following TO have length $64T_0$
- The duration of a sequence of k TO is

$$L_k = \begin{cases} (2^k - 1)T_0 & k \le 6 \\ (63 + 64(k-6))T_0 & k > 6 \end{cases}$$

- Thus,

$$E[Z^{TO}] = \sum_{k=1}^{\infty} L_k P[R = k]$$
$$= T_0 \frac{1 + p + 2p^2 + 4p^3 + 8p^4 + 16p^5 + 32p^6}{1-p}$$

Solving $B(p)$

$$B(p) = \frac{E[Y] + Q \times E[R]}{E[A] + Q \times E[Z^{TO}]}$$
$$= \frac{\frac{1-p}{p} + E[W] + Q'(E[W])\frac{1}{1-p}}{RTT(E[X]+1) + Q'(E[W])E[Z^{TO}]}$$
$$\approx \frac{1}{RTT\sqrt{\frac{2bp}{3}} + T_0 \min\left(1, 3\sqrt{\frac{3bp}{8}}\right) p(1+32p^2)}$$