TCP Congestion control

TCP uses a congestion window and a congestion policy that avoid congestion. Previously, we assumed that only the receiver can dictate the sender's window size. We ignored another entity here, the network. If the network cannot deliver the data as fast as it is created by the sender, it must tell the sender to slow down. In other words, in addition to the receiver, the network is a second entity that determines the size of the sender's window.

Congestion policy in TCP –

1. Slow Start Phase: starts slowly increment is exponential to threshold
2. Congestion Avoidance Phase: After reaching the threshold increment is by 1
3. Congestion Detection Phase: Sender goes back to Slow start phase or Congestion avoidance phase.

Slow Start Phase : exponential increment – In this phase after every RTT the congestion window size increments exponentially.

Initially cwnd = 1
After 1 RTT, cwnd = 2^(1) = 2
2 RTT, cwnd = 2^(2) = 4
3 RTT, cwnd = 2^(3) = 8

Congestion Avoidance Phase : additive increment – This phase starts after the threshold value also denoted as ssthresh. The size of cwnd(congestion window) increases additive. After each RTT cwnd = cwnd + 1.

Initially cwnd = i
After 1 RTT, cwnd = i+1
2 RTT, cwnd = i+2
3 RTT, cwnd = i+3

Congestion Detection Phase : multiplicative decrement – If congestion occurs, the congestion window size is decreased. The only way a sender can guess that congestion has occurred is the need to retransmit a segment. Retransmission is needed to recover a missing packet that is assumed to have been dropped by a router due to congestion. Retransmission can occur in one of two cases: when the RTO timer times out or when three duplicate ACKs are received.

Case 1 : Retransmission due to Timeout – In this case congestion possibility is high.
(a) ssthresh is reduced to half of the current window size.
(b) set cwnd = 1
(c) start with slow start phase again.

Case 2 : Retransmission due to 3 Acknowledgement Duplicates – In this case congestion possibility is less.
(a) ssthresh value reduces to half of the current window size.
(b) set cwnd= ssthresh
(c) start with congestion avoidance phase

1. Congestion window size

It is known to the sender, and the bytes sent should never exceed the congestion window size. If the sending rate goes over the congestion window size, packets are lost and the sender is forced to retransmit them.

2. Receiver window size

It is an advertisement from the receiver about the number of bytes it can receive before dropping packets. As such, the sender must always send bytes less than or equal to the receiver window size to avoid retransmission of packets.

Hence, the formula for sender window size is:

Sender window size = Min(Congestion window size, Receiver window size)

Congestion control in TCP

TCP's mechanism for congestion control comprises three phases:

1. Slow start

The sender starts by setting the congestion window equal to the maximum segment size (1 MSS) and increases its size by one after each acknowledgment is received. The size of the congestion window hence increases exponentially in this phase, as shown below:

The congestion window size increases exponentially during slow start phase

The illustration shows how the congestion window size changes exponentially during the slow-start phase. The slow-start phase continues until the **slow-start threshold** is reached, which is defined as follows:

- Slow-start threshold (SSH) = (Receiver window size / Maximum Segment Size) / 2

2. Congestion avoidance phase

This phase starts once the slow-start threshold is reached and continues until the congestion window becomes equal to the receiver window size. In this phase, the sender cautiously increases the congestion window by one after receiving an acknowledgment to avoid congestion. The congestion window increments as follows:

- Congestion window size = congestion window size + 1

The progression of transmission up until now is shown below:

The transmission of data until receiver window size is reached

3. Congestion Detection Phase

In the last phase, TCP promptly acts when a packet loss is detected. Its reaction is based on the two types of congestion given below:

<u>1. Mild congestion</u>

Congestion is mild when the sender receives three duplicate acknowledgments (ACKs) from the receiver. It indicates the possibility of only a few packets being dropped, which can be perceived through the ACKs received. When this happens, the following happens:

- The congestion window is halved.
- The slow-start threshold is set to this new congestion window value.
- The sender continues with the congestion avoidance phase until a loss is detected.

2. Severe congestion

This case occurs when the retransmission timer of the sender expires. When this happens, the following happens:

- The slow-start phase is started until the first loss is detected after the retransmission timer expires.
- The sender sets the slow-start threshold to halve the current congestion window size.
- Congestion window size is reduced to 1 MSS by the sender.
- The slow-start phase is resumed until the congestion window reaches the slow-start threshold.
- The sender then goes into the congestion avoidance phase until a loss of segments is detected through the retransmission timer expiry.