Congestions and Control Mechanisms n Wired and Wireless Networks

MD Gulzar[1], B Mahender[2], Mr.B.Buchibabu[3]
[1](Asst professor, cmrec, Hyderabad, india Gulzar.arhan@gmail.com)
[2](Asst professor, cmrec, Hyderabad, india mahendra504@gmail.com)
[3](Asst Professor, cmrec, Hyderabad, india b.buchibabu007@gmail.com)

ABSTRACT: In today's world, network is becoming compulsory technology even in a normal human life ranging from a student to a professional. In every field networks are being used. In olden days networks are formed by connecting the computer systems through cables. As the days are going on more number of computer systems are connected to the network. As a result network size gets increased and networks like LAN, MAN, WAN came into existence. Huge networks with cables has so many disadvantages like, if the cable is damaged there will be no communication between the system, and since one cable may be shared between many users the communication may become slower, and the users may put heavy load on the communicating cable than it's capacity which may cause congestion. To overcome all these demerits of wired networks wireless networks are proposed. In wireless networks there are no communicating physical cables. The system will communicate with each other by using radio or infra red waves which travels through air. The wireless networks will also suffer from so many problems like node mobility, , high error rate, long latency and even congestions. In this paper we are going to discuss various congestion control mechanisms in wired and wireless networks and their merits and demerits.

Keywords: congestion, congestion control TCP, wired network, wireless network

I. Introduction

The main problem In todays network is slower communication, which may result due to so many reasons like client may be a slower than server or vice versa, or the communicating channel may have the less bandwidth than the requirement of the communuatingresoureces, or so many computers are connected to a single network or a server for ex: when so many clients are connected each may have their own requests and there may be only one server to fulfill these requests due to this every client needs to wait until it's request get executed, and the intermediate resources of a network like routers and swithes may be slower. Due to all these reasons network performance becomes degraded. One more problem in network is congestion, when this occurs, the communicating devices may be blocked until the congestion is cleared.

II. Congestion

Congestion in a network occurs when the offered load by the communicating devices is greater than the available bandwidth of the communicating mechanism. Congestion may occur either at the resources or at the communicating channel. When a fast sender is sending packets to a slow receiver its buffer or queue in which the packets are stored before processing gets filled and remaining in comingpackets get discarded. In each case the sender needs to be informed regarding the congestion either at the network or receiver so that it will slow down the process of sending packets. The congestion may be handled differently in wired and wireless networks which we discuss in following sections.

III. Congestion Control In Wired Networks

Form the control theory point of view , all the solutions are divided into two categories. They are open loop congestion control and closed loop congestion control.Open loop solutions attempt to solve the problem by good design, to make sure congestion does not occur in first place. Once the system is created corrections ca not be made in the middle [10].

Closed loop solutions are based on the concept of feedback loop. This approach has 3 parts.

1. Monitor the system to detect when and where congestion occurs.
2. Pass this information to places where action can be taken.

3. Adjust the system operation to correct the problem.

Open loop congestion control algorithms are further divided into two categories. They are which acts at source and which acts at destination. The closed loop algorithms are also divided into two categories. They are implicit feedback and explicit feedback.

In explicit feedback the packets are sent back from the point of congestion to warn the source. In implicit the source identifies the congestion by making local observations. Such as time needed for acknowledgments to come back. The presence of congestion means the load is greater than the resource can handle. We can increase the resources or decrease the load to control the congestion[10].

3.1 Congestion Control in TCP

In TCP congestion control , the sender window size or sender's buffer size is decided by the available buffer space in the receiver window(rwnd), as well as the network's bandwidth called as congestion window(cwnd) because if the network has low bandwidth it cannot deliver all packets sent by the source to destination. So the actual window size of the sender is minimum of these two [9].

TCP's general policy for handling congestion is basedon three phases

1. Slow start
2. Congestion avoidance
3. Congestion detection

In slow start phase , the sender starts with a very slow rate of transmission, but increases the rate rapidly to reach the threshold. When the threshold is reached the data rate is reduced to avoid congestion.

3.1.1 Slow Start: Exponential Increase

One of the algorithms used by TCP to control the congestion is "slow start". Here initially the sender's window size is equal to congestion window(cwnd), which is equal to one maximum segment size(MSS).And it is assumed that rwnd is much higher than cwnd, so that sender window size always equal to cwnd.In this it is assumed that each segment is acknowledged individually. The sender starts with cwnd=1 MSS, that it can send only one segment. After receipt of acknowledgment] for segment 1, the cwnd is increased by 1, now cwnd is 2. Now two more segments can be sent, and these two can be acknowledged, the cwnd becomes 4. Now 4 segments are sent, when acknowledgment is received the cwnd becomes 8.

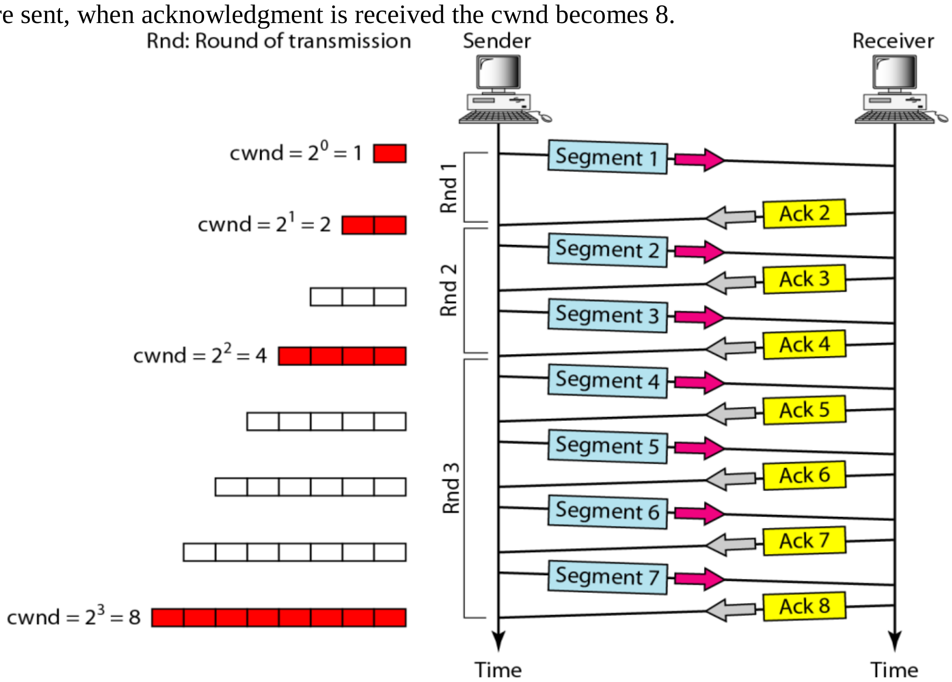

Figure1: Slow start

If we look at cwnd in terms of rounds, we find the rate is exponential as below.

Start->cwnd=1

After round 1->cwnd = 2^1 =2

After round 2 ->cwnd = 2^2= 4

After round 3 ->cwnd = 2^3 = 8

Slow start cannot continue indefinitely, there must be a threshold to stop these phase. The sender uses a variable ssthresh(slow start threshold). When the size of the window reaches to the threshold, slow start stops and next phase starts. In most implementations the value of ssthresh is 65,535 bytes.

3.1.2 Congestion Avoidance: Additive Increase

In slow start algorithm, the size of cwnd increase exponentially. To avoid congestion before it happens, the exponential growth must slow down. TCP defines algorithm called as congestion avoidance, which uses additive increase instead of exponential one [9]. When size of cwnd reaches to the threshold, slow start phase stops and additive phase begins.In this algorithm cwnd size will be incremented by one only when all segments in window are acknowledged.

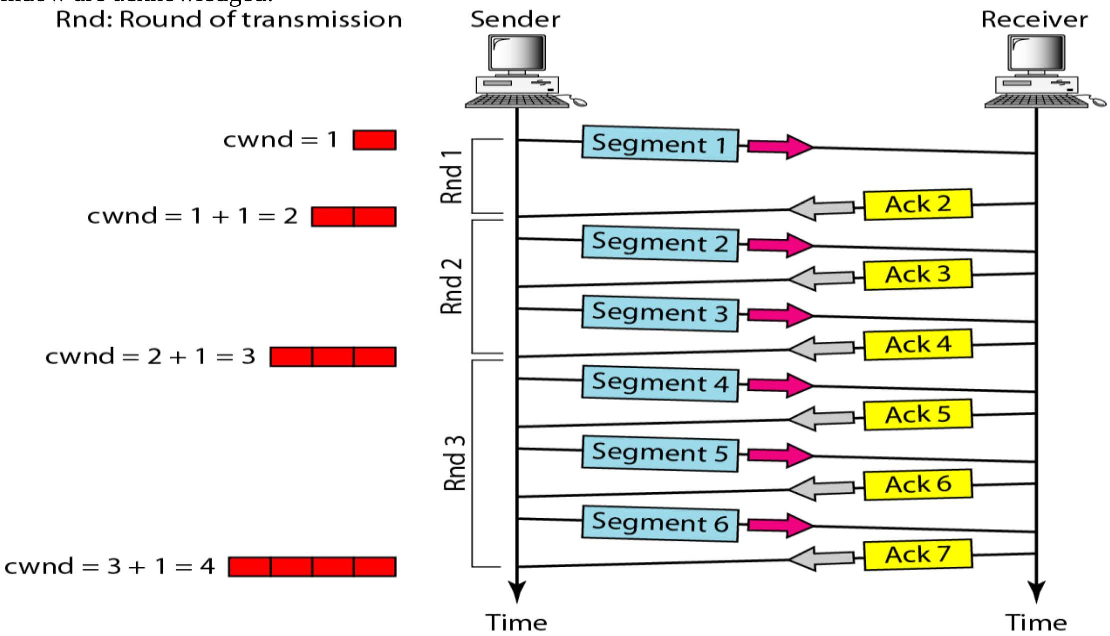

Figure 2 : congestion avoidance

If we look cwnd in terms of rounds

Start ->cwnd=1
Round 1 ->cwnd=1+1=2
Round 2 ->cwnd=2+1=3
Round 2 ->cwnd=3+1=4

In congestion avoidance phase the size of the cwnd increase additively until congestion is detected.

3.1.3 Congestion Detection

When congestion occurs , the congestion window size must bedecreased. The sender can guess about the congestion by the need of retransmission of the segment. Retransmission occurs in one of the two cases.

1. When a timer times out.
2. When 3 acknowledgments are achieved.

In both cases threshold will be dropped to one half, called as multiplicative decrease [9]

3.1.4 TCP Congestion Policy Summery

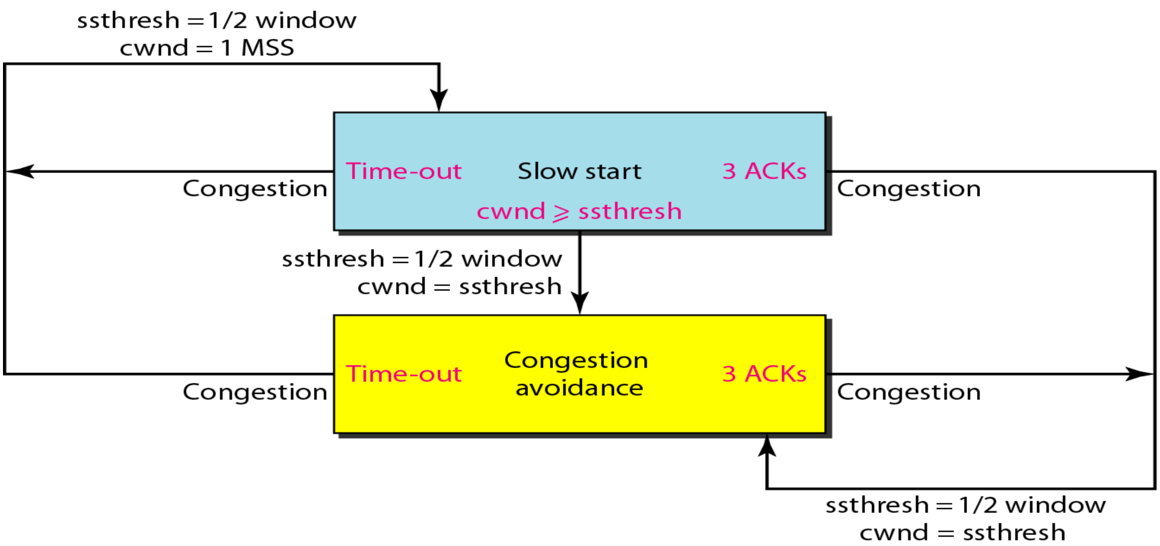

Figure3: TCP congestion policy

3.2 TCP -Tahoe and TCP -Reno

TCP-Tahoe is the old version of TCP. In this case cwnd size is reduced to 1 MSS. in each case of time out and 3 acknowledgments are arrived. TCP-Reno is the new version of TCP, in congestion detection phase when a time out occurs cwnd is decreased to 1MSS, and when three acknowledgments are arrived cwnd is reduced half of cwnd [8].

The reason for TCP Reno is , when 3 acknowledgments are re received means 3 packets are delivered , which means network has capacity to forward some acknowledgments. This is called as fast recovery.
The following fig shows the performance evaluation of both Reno and Tahoe[8].

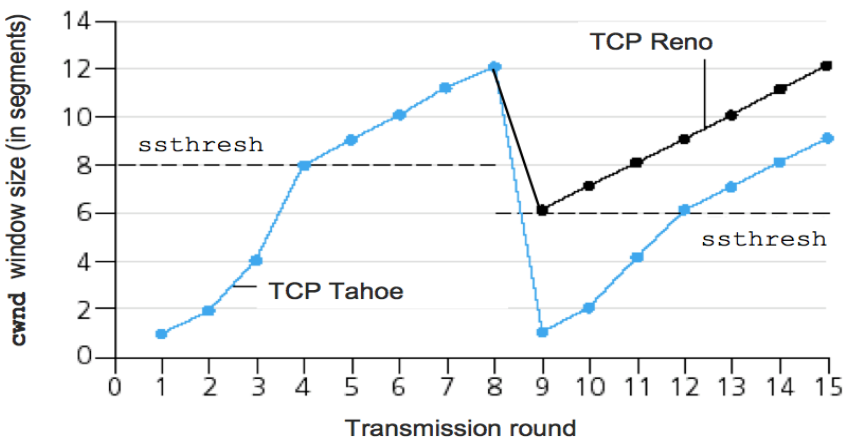

Figure 4: Comparison of TCP Tahoe and TCP Reno

Initially the threshold is equal to 8 MSS. The congestion window increases fast exponentially during slow start, and reaches to threshold at fourth round, the cwnd increases up to 3 duplicate ACKNOWLEDMENTs occurs. After transmission round 8, cwnd is 12 MSS when the loss event occurs. At this point cwnd decreases to one half to 6. In TCP Reno the cwnd=6 MSS, and in TCP Tahoe the cwnd is set to 1MSS and then grows linearly until threshold is reached.

Most TCP implementation today uses TCP-Reno algorithm and many variations of Reno algorithm have been proposed.

3.3 TCP Vegas

The TCP Vegas algorithm attempts to avoid congestion while maintaining good through put[2]. The basic idea of Vegas is to

1. Detect the congestion in routers between source and destination before packet loss occurs.
2. Lower the rate when packet loss is detected. Packet loss is predicted by observing the RTT. The longer RTT means the greater the congestion is.

IV. Congestion Control In Wireless Networks

The TCP was designed on the assumption that, loss of data packets is caused by congestion. But in wireless networks data packets losses due to high bit rate, nodes mobility, one tomany links and link failure etc. So the conventional TCP mechanism is not appropriate for wireless adhoc networks. To the large extent TCP will reduce the network performance. This unmodified TCP is unable to distinguish packet loss caused by network congestion and transmission errors.

4.1 TCP In wireless networks

In wireless links TCP reacts to packet loss as it is in wired environment, and it drops the transmission window before retransmitting the packets and initiates congestion control or avoidance mechanism. This will unnecessarily reduce the link's bandwidthutilization and causes significant performance degradation in the form of throughput[2].

It is important to improve the performance of TCP in wireless networks because many network applications may be builted up on TCP. Several reliable transport layer protocols for n/w with wireless links have been proposed to improve the performance of TCP in wireless machine.

4.1.1 The split connection approach

This approach is used in indirect TCP (I-TCP). It includes splitting a TCP connection between a fixed host, and mobile host in to two separate connections. One TCP Connection is established between fixed host and base station, and other between base station and mobile host[1].

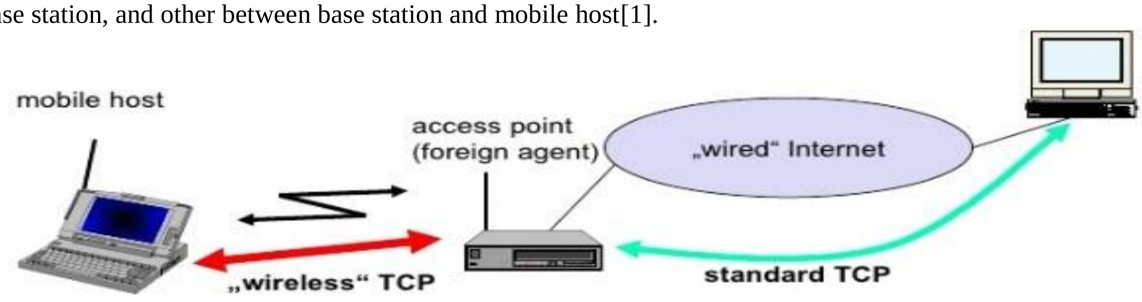

Figure 5: Split connection approach in I-TCP

On the second connection there is no need to use TCP, we may use a wireless specific protocol for better performance. The advantage here is the flow and congestion control of wireless link is separated from fixed host.

4.1.2. The Fast retransmit approach

This approach is used with TCP when communication resumes after a handoff. Unmodified TCP at the sender considers the delay caused by handoff process to be due to congestion, and when timeout occurs , reduces it's window size and retransmits unacknowledged packets[1].

The fast retransmit approach overcomes this problem by having the mobile host send a certain number of duplicateacknowledgments to the sender to immediately reduce it's window size and retransmit packets starting from missing one.

4.1.3. Link Level Retransmission

In this approach , the n/w is divided into two parts like fixed host to base station and from base station to mobile host. For fixed host to base station wired connection is used and for base station to mobile host wireless connection is used. The wireless link implements a retransmission protocol and forward error correction method, as well as TCP implements its own end-to-end retransmission protocol [3].

4.1.4 Snoop Protocol

The n/w applications that require reliable transmission uses TCP. So we may want o improve tcp performance without changing it's implementation. The only components that we may want to control are the base stations and mobile hosts. In snoop protocol modifications are done to the routing code of base stations only. These modifications are for catching unacknowledged TCP data and performing local retransmissions as soon as loss is detected [4].

This protocol will shield the sender regarding the packetloss, or problems occurred at the wireless link. This results in increased performance of the connection without sacrificing the end-to-end semantics of the TCP[7].

The operation of snoop protocol can be shown as follows

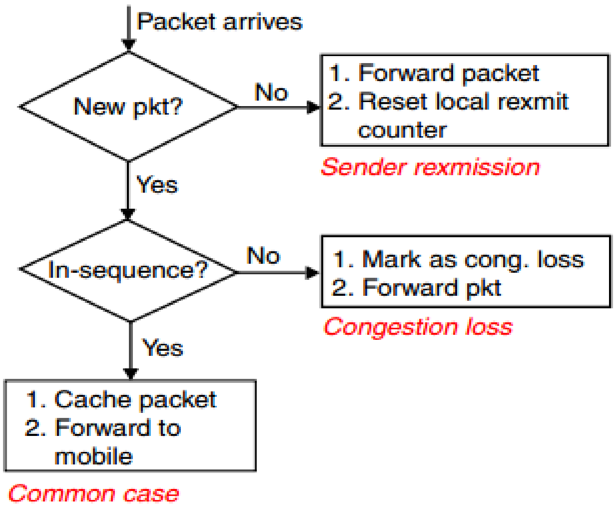

Figure 6: Operation of snoop protocol

4.1.5 Data Gram Congestion Control Protocol (DCCP)

DCCP is a message and connection oriented transport layer protocol, it differs from UDP because it includes congestion control and it differs from TCP, where it does not provide guaranteed reliability. DCCP is a general congestion control protocol for UDP-based applications. But DCCP provides semantics like in TCP, but does not provide reliable in order delivery, sequenced delivery within multiple streams as in SCTP. To make DCCP reliable and in-order delivery, DCCP allows the use of general TCP-friendly congestion control mechanisms like TCP-like and TCP-friendly [5].

DCCP is completely a new protocol; it is still under research that weather it can be used for real time applications practically [6].

V. Conclusion

The internet communication became a daily routine. And the major problem in the network is congestion which causes slower, disturbed communication. Congestions will occur in wireless and wired networks. Since wireless network has so many advantages over wired, so it is the mostly used technology. In this paper we have discussed various congestions and their controlling mechanisms in wired and wireless networks. But in future we may have more featured, fast networks which may also suffer from congestions, so there is a need to design a reliable congestion control mechanism for secure and fast communication.

REFERENCES

[1] Ning Jia, Lei An, Analysis of Congestion Control Strategy for Wireless Network, *Institute of Electrical and Electronics Engineers*, 978-1-4244-4994-1/09/$25.00 ©2009 IEEE

[2] Alhamali Masoud Alfrgani .Ali, Raghav Yadav**,** Hari Mohan Singh, Congestion Control Technique for Wireless Networks**,** *IOSR Journal of Computer Engineering (IOSR-JCE),* e-ISSN: 2278-0661, p- ISSN: 2278-8727

[3] Ali Ghaffari, Congestion control mechanisms in wireless sensor networks: A survey, A survey. *Journal of Network and Computer Applications* (2015), http://dx.doi.org/10.1016/j.jnca.2015.03.002i

[4] Alonso Silva, Patricio Reyes, Merouane Debbah, Congestion in Randomly Deployed Wireless Ad-Hoc and Sensor Networks, *Institute of Electrical and Electronics Engineers,* 9781-4244-3941-6/09/$25.00© 2009 IEEE

[5] Atilla Eryilmaz and R. Srikant**,** Fair Resource Allocation in Wireless Networks using Queue-length-based Scheduling and Congestion Control, *Institute of Electrical and Electronics Engineers,* 0-7803-8968-9/05/$20.0 ©2005 IEEE

[6] N. Thrimoorthy and Dr. T .Anuradha , A Review on Congestion control Mechanisms in Wireless Sensor Networks**,** *International Journal of Engineering Research and Applications, ISSN : 2248-9622*

[7] Raheleh Hashemzehi, Reza Nourmandipour,Farokh koroupi, Congestion in Wireless Sensor Networks and Mechanisms for Controling Congestion, *Indian Journal of Computer Science and Engineering*, ISSN : 0976-5166

[8] James E. Kurose, Computer Networking, (Pearson Education, Inc, Dorling, Kindersley(india)Pvt.Ltd)

[9] Behrouz A Forouzan, Data Communication and Networking (The MC Graw-Hill Education Pvt.Lid, New Delhi(India))

[10] Andrew S.Tanenbum, Computer Networks(Pearson Education ,Inc, Dorling, Kindersley(india)Pvt.Ltd)