Review Article **AJCE (2019) 2:5**

American Journal of Computer Engineering (AJCE)

Understand Congestion: It's Effects on Modern Networks

Omotosho Olawale, Oyebode Aduragbemi* and Hinmikaiye Johnson

Department of Computer Science Babcock University, Nigeria

Introduction

As Internet, can be considered as a Queue of packets, where transmitting nodes are constantly adding packets and some of them (receiving nodes) are removing packets from the queue. So, consider a situation where too many packets are present in this queue (or internet or a part of internet), such that constantly transmitting nodes are pouring packets at a higher rate than receiving nodes are removing them (Cardwell, Cheng, Gunn, Yeganeh, & Jacobson, 2016). This degrades the performance, and such a situation is termed as Congestion. Main reason for congestion in a network system is a greater number of packets into the network than it can handle. So, the objective of congestion control can be summarized as to maintain the number of packets in the network below the level at which performance falls off dramatically (Faisal Shahzad1, Ullah, Siddique, Khurram, & Saher, 2015). The nature of a Packet switching network can be summarized in following points:

• A network of queues

• At each node, there is a queue of packets for each outgoing channel

• If packet arrival rate exceeds the packet transmission rate, the queue size grows without bound

• When the line for which packets are, queuing becomes more than 80% utilized, the queue
length grows alarmingly

When the number of packets dumped into the network is within the carrying capacity, they all are delivered, expect a few that have to be rejected due to transmission errors). And then the number delivered is proportional to the number of packets sent (Evans & Filsfils, 2007). However, as traffic increases too far, the

***Correspondence to Author:**
Oyebode Aduragbemi
Department of Computer Science
Babcock University, Nigeria

How to cite this article:
Omotosho Olawale, Oyebode Aduragbemi and Hinmikaiye Johnson. Understand Congestion: It's Effects on Modern Networks. American Journal of Computer Engineering, 2019; 2:5.

eSciPub LLC, Houston, TX USA.
Website: http://escipub.com/

routers are no longer able to cope, and they begin to lose packets. This tends to make matter worse. At very high traffic, performance collapse completely, and almost no packet is delivered.

Congestion is an important issue that can arise in packet switched network. Congestion is a situation in Communication Networks in which too many packets are present in a part of the subnet, performance degrades (Floyd & Fall, 1999). Congestion in a network may occur when the load on the network (i.e. the number of packets sent to the network) is greater than the capacity of the network (i.e. the number of packets a network can handle). In other words, when too much traffic is offered, congestion sets in and performance degrades sharply (Jayakumari & Senthilkumar, 2015). Congestion collapse occurs when the network is increasingly busy, but little useful work is getting done.

Congested network refers to the moment in network's links when any new data entry to be sent to a destination will create instead a blocking effect into the transmission line. Thus, this may result primarily to throughput decrease with the data already admitted in process. Congestion occurs when the number of packets being transmitted through the network approaches the packet handling capacity of the network. It can be described is the reduced quality of service that occurs when a network node is carrying more data than it can handle.

Typical effects include queueing delay, packet loss or the blocking of new connections. A consequence of congestion is that an incremental increase in offered load leads either only to a small increase or even a decrease in network throughput.

Network protocols that use aggressive retransmissions to compensate for packet loss due to congestion can increase congestion, even after the initial load has been reduced to a level that would not normally have induced network congestion. Such networks exhibit two stable states under the same level of load. The stable state with low throughput is known as congestive collapse. Congestion collapse occurs when the network is increasingly busy, but little useful work is getting done.

2.Causes of Congestion

Congestion occurs when the request for packet service is greater than the available resource the network has at that present moment. Although congestion can be described as a resource allocation problem, making resources less expensive and more accessible is not a viable solution. Some of the solutions that has been widely provided include:

- Congestion is caused by lack of sufficient buffer memory so the solution to this is that memory has to become cheap enough thereby allowing for infinite storage.
- Slow links between connecting nodes is the cause of congestion. Having high speed connecting links will solve the problem.
- Having slow processing power will lead to congestion, having high speed processors will solve the problem.

All these solutions listed only provide just a temporary fix to the problem, if they are seen as fact by network administrators it could lead to improper network planning and implementation. On the contrary to all these assumptions, proper network protocol designing and implementation is needed to establish a network to combat congestion else, by implementing any of these assumptions above, it would reduce the overall throughput and also reduce performance thereby adding more congestion to the network system.

Infinite buffer storage although not a bad idea cannot offer a total solution to the problem, it gives just a temporal solution which will not last for long, although small memory buffers when faced with too much traffic suffer from buffer overflow and loss of packets but if we are to have infinite butter storage during heavy traffic would result in long queues of packets, delays and so on. Finally, when packets leave the queue, so much time would have passed and these packets would be times out.

Also, having high speed links between nodes will not make much difference to solve the problem of congestion, at the time when the internet began, the speed for nodes and workstation available as at the time was not more than 300 bits per seconds (300b/s). Gradually with increase in technology speeds of up to 1.5 megabits per seconds (1.5Mb/s) was attainable with the introduction of Local Area Networks (LANs) and also Ethernet speeds of about 10 megabits per seconds (10Mb/s) was possible (Jain, Congestion Control in Computer Networks: Issues and Trends, 1990).

Also, having high speed links between nodes will not make much difference to solve the problem of congestion, at the time when the internet began, the speed for nodes and workstation available as at the time was not more than 300 bits per seconds (300b/s). Gradually with increase in technology speeds of up to 1.5 megabits per seconds (1.5Mb/s) was attainable with the introduction of Local Area Networks (LANs) and also Ethernet speeds of about 10 megabits per seconds (10Mb/s) was possible (Jain, Congestion Control in Computer Networks: Issues and Trends, 1990) (Mohamed, Sahib, Suryana, & Hussin., 2016). It was at this point in history the issue of congestion started to gain recognition due to a mismatch such that fast-high speed LANs were connected via slow links and slow nodes that cannot keep up to the speed of these links. Today, when it come to the speed of data or internet connections we can have speeds up to Gigabits or even Terabits but the issue of congestion still persists, therefore having high speed links is not the solution. Also, the same can be taught of for slow processors, having high speed processors can lead to a mismatch.

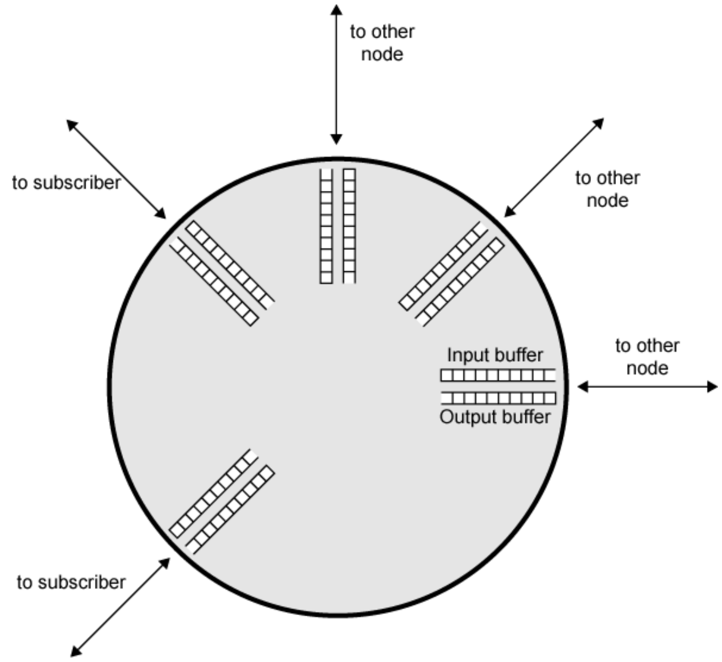

Figure 1: The logical description of a node. (Source Spring, 2003)

Figure 1 shows a model of node and how nodes generally behave. The input and output buffer store packets sent and also received from other nodes waiting to be transmitted. A network is comprised of various nodes and all this nodes account for packets queued up and waiting to be transmitted.

3.Related Works

Over the years many algorithms, strategies coupled with procedures and also methods have been employed and adopted to control and manage congestion. This work carries out a comprehensive review on these methods.

In 2016, Xi Hu and Wei Guo who are Chinese scientists developed a “robust control model” to tackle congestion issues in wireless networks. Their work which was on congestion control in wireless software defined networks with propagation delay and external interference using a robust control approach, this was implement through the use of a centralized controller which was embedded into network simulations. This controller was tasked with the work of retrieving

information from all nodes and devices on the network that pertains to networks that are congested. This is very crucial due to the fact that the congestion control algorithm will use the information supplied as its parameters. A nonpreemptive scheme was used in the control system. The congestion problem is difficult has it is but considering a network environment where scheduling is also of importance, a nonpreemptive scheme is adopted which also offers stability to the network parameters gotten from various nodes and devices which will also maximize the overall throughput of the system. The solution provided by Xi Hu and Wei Guo who seem complex, requiring complex computations using mathematical theorems which might not be easily understood by numerous network administration.

Schemes and algorithms adopted to tackle the congestion problem must have what is known has a low overhead (Raj Jain 1990). This is also an important property of any algorithm to be adopted. It should not increase traffic on the network or reduce the overall throughput of the networks. The solution provided by Xi Hu and Wei Guo due to the computational complexity of the solution adopted would in theory greatly increase the workload on the system and thus the overall traffic therefore this solution may not be preferred in certain kinds of networks.

Saleem et al, in 2014 postulated that network efficiency can be improved by selecting and modifying congestion control constraints. They proposed and investigated an efficient mechanism for congestion control by the selection of appropriate congestion window size and proactive congestion avoidance, which improves system overall performance and efficiency. The work focused more on proactive congestion avoidance by postulating various algorithms and mathematical model to ensure the passage of data traffic through a congested network. It was not really offering solutions to the problem but offering a way to cope and improve the efficiency of a congested network.

Chegwen et all, in 2014 worked on a fast heuristic algorithm for minimizing congestion in MPLS networks. In the modern broadband communication networks, namely the multiple protocol label-switched networks (MPLS), data packets are transmitted through the label switched paths (LSPs). This algorithm is based on the strategy that reduces the whole feasible path set for each commodity, and a Round Robin strategy is also used. One main limitation of the new algorithm is that it needs to solve a mixed integer linear programming, when the number of commodities is sufficiently large, the running time of the algorithm may be large. However, the idea of the algorithm provides a good feasible direction for solving this problem.

Faisal et al, in 2015 worked on improving the queuing system throughput using distributed mean value analysis to control network congestion. The work used the distributed mean value analysis (DMVA) technique with the help of random observe property (ROP) and palm probabilities to improve the network queuing system throughput. In such networks, where finding the complete communication path from source to destination, especially when these nodes are not in the same region while sending data between two nodes. So, an algorithm is developed for single and multi-server centres which give more interesting and successful results. The network is designed by a closed queuing network model and we will use mean value analysis to determine the network throughput for its different values.

In 2017, Mohamed et al, worked on understanding the effect of congestion on network performance. The work analysed various Network Congestion types and also congestion control tools. Various causes of congestion were cited including node overload, packet over flow and so on. Although solutions were not specified, the work provided a broad overview to

the problem of congestion and how performance of networks is affected. The work reviewed papers written to uncover underlining reasons and also possible solutions to the problem of congestion.

Chengwen Jiao et al in 2014 did a study on the application of heuristics to the problem of congestion in networks. In the study, a fast heuristics algorithm and two mathematical theorems were employed to minimize congestion in networks. Data packets are transmitted via routes, the routes can be seen as paths from source to destination. If these routes can be known we can determine how busy the route is and if its busy we can transfer packets through other routes. The most feasible path for packets to travel is calculated and also other techniques such as round-robin in which token is given to nodes was also employed. Although these had a sufficiently high overhead, it would prove useful in real-time systems. The fact that the algorithm has a high overhead, it limits the whole application of the algorithms and cannot be employed by all networks.

Syed Afsar Shah et al in 2015, did a study on congestion control algorithms their main focus was on wireless sensor networks. Considering the delicate nature and also functionality of sensor networks, which operate and function in real-time, congestion control is of vital importance to sensor networks in which a down time of even seconds can be critical to network performance.

Taking into considerations all the various methods that have previously been used in sensor networks to combat congestion such as routing protocols coupled with features of congestion detection and also protocols dedicated to congestions control, the study reviews the various literature on how these methods have been implemented on various networks. Sensor networks consists of various individual devices all of which collect information and data from the external environment. A sink node is the node with which all information gathered by the various network elements are relayed to (Shar, Nazir, & Khan, 2016). This transfer is very crucial to the network, the information should be pass to the sink node with as little congestion as possible, which calls for an adequate congestion control mechanism. Although adequate reviews on existing techniques was done, most of these techniques will have a high overhead which will add sufficient congestion to the system.

4.Effect of Congestion

Congestion disrupts two important parameters of network performance; the two parameters are:

1. Network Throughput: This can be defined as the percentage utilization of the overall network capacity, it can be further viewed as the total amount of data moved or transferred successfully from a network element to another.

   The throughput of a network system increases as the traffic on the system increases, ideally the relationship between traffic and throughput is linear because the utilization is on the rise, but as traffic increases beyond certain limits, the throughput begins to drops and if the traffic continues to rise unchecked, a point will be reached where no the network will not be able to deliver any single packet. This situation can be referred to as network deadlock.

2. Network Delay: With increase in network traffic, packets are delayed and arrive late at the destination node. These causes network lag, network performance becomes poor, no matter the strategy or technique used for congestion control, the delay will continue to grow as the network traffic increases and also as it reaches network capacity. Also, it

is worth stating that there will be longer delays when congestion control policies are applied due to the overhead it will cause on the network.

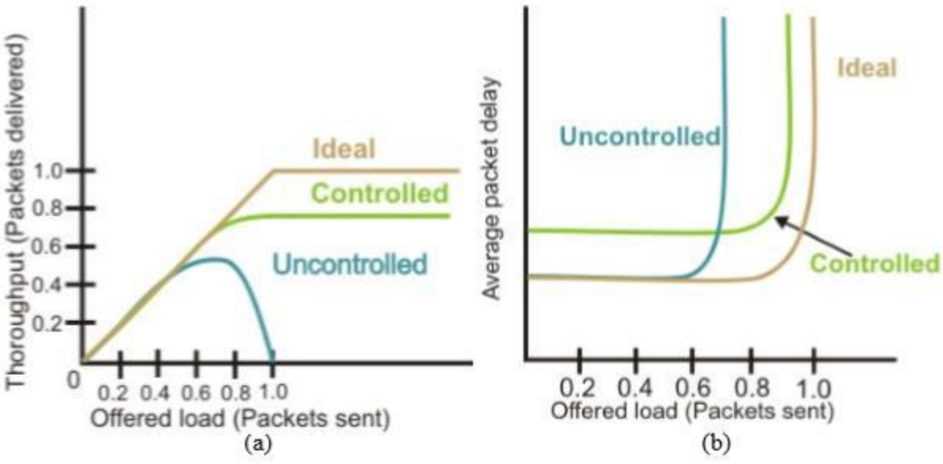

Figure 2.2 (a) Effect of Congestion on throughput (b) Effect of Congestion on delay. Source (Kharagpur, 2016)

6.Trends

Congestion control modulates traffic entry into a telecommunications network in order to avoid congestive collapse resulting from oversubscription. This is typically accomplished by reducing the rate of packets and it should not be confused with flow control, which prevents the sender from overwhelming the receiver (Jiao, Gao, Yang, Xia, & Zhu, 2014). Today, various researches in been carried out in this field, trying to come up with numerous algorithms for the control of congestion.

Congestion control algorithms care classified in the following:

• By type and amount of feedback received from the network: Loss; delay; single-bit or

multi-bit explicit signals

• By incremental deployability: Only sender needs modification; sender and receiver

need modification; only router needs modification; sender, receiver and routers need modification.

• By performance aspect: high bandwidth-delay product networks; lossy links; fairness; advantage to short flows; variable-rate links

• By fairness criterion: max-min, proportional, "minimum potential delay"

A few mechanisms have been invented to prevent network congestion or to deal with a network collapse:

• Network scheduler – active queue management (that is, the arbitrary reorder or drop of network packets under overload)

• Explicit Congestion Notification – an extension to IP and TCP communications protocols that adds a flow control mechanism

• TCP congestion control – various implementations of efforts to deal with network congestion

The correct endpoint behavior is usually to repeat dropped information, but progressively slow the repetition rate. Provided all endpoints do this, the congestion lifts and the network resume normal behavior. Other strategies such as slow-start ensure that new connections don't overwhelm the router before congestion detection initiates.

The most common router congestion avoidance mechanisms are fair queuing and other scheduling algorithms, and random early detection, or RED, where packets are randomly dropped, proactively triggering the endpoints to slow transmission before congestion collapse occurs. Fair queuing is most useful in routers at chokepoints with a small number of connections passing through them. Larger routers must rely on RED.

Some end-to-end protocols behave better under congested conditions. TCP is perhaps the best behaved. The first TCP implementations to handle congestion were developed in 1984,[citation needed] but Van Jacobson's inclusion of an open source solution in the Berkeley Standard

Distribution UNIX ("BSD") in 1988 first provided good behavior.

UDP does not control congestion. Protocols built atop UDP must handle congestion independently. Protocols that transmit at a fixed rate, independent of congestion, can be problematic. Real-time streaming protocols, including many Voice over IP protocols, have this property. Thus, special measures, such as quality-of-service routing, must be taken to keep packets from being dropped.

In general, congestion in pure datagram networks must be kept at the periphery of the network, where the above mechanisms can handle it. Congestion in the Internet backbone is problematic. Cheap fiber-optic lines have reduced costs in the Internet backbone allowing it to be provisioned with enough bandwidth to keep congestion at the periphery.

Connection-oriented protocols, such as the widely used TCP protocol, generally watch for packet errors, losses, or delays to adjust the transmit speed. Various network congestion avoidance processes, support different trade-offs (Kharagpur, 2016).

The TCP congestion avoidance algorithm is the primary basis for congestion control in the Internet. Problems occur when concurrent TCP flows experience port queue buffer tail-drops, defeating TCP's automatic congestion avoidance. All flows that experience port queue buffer tail-drop begin a TCP retrain at the same moment – this is called TCP global synchronization.

One solution is to use random early detection (RED) on the network equipment's port queue buffer. On network equipment ports with more than one queue buffer, weighted random early detection (WRED) could be used if available (Agarwal, 2000).

RED indirectly signals to sender and receiver by deleting some packets, e.g. when the average queue buffer lengths are more than a threshold (e.g. 50%) and deletes linearly or cubically more packets, up to e.g. 100%. The average queue buffer lengths are computed over 1 second intervals.

Some network equipment is equipped with ports that can follow and measure each flow (flow based-RED/WRED) and are thereby able to signal a too big bandwidth flow according to some quality of service policy. A policy could then divide the bandwidth among all flows by some criteria.

7.Conclusion

In summary, reviews from literature show that a vast amount of research has been carried out on the problem of congestion but the problem still persists. The solutions provided does not guarantee a system void of congestion. Transmission Control Protocol (TCP) Congestion Control provides a way to help minimize the effect of congestion, but it this approach favours reliability over performance. Transmission Control Protocol (TCP) Congestion Control cannot handle huge burst of data, it ensures a reliable system where all packets are accounted for, and if any is lost, the packet is retransmitted. Long-Term Evolution networks provide broadband services and performance is paramount, TCP would not be a viable option.

Congestion avoidance can be achieved efficiently by reducing traffic. When an application requests a large file, graphic or web page, it usually advertises a "window" of between 32K and 64K. This results in the server sending a full window of data (assuming the file is larger than the window). When many applications simultaneously request downloads, this data creates a congestion point at an upstream provider by flooding the queue. By using a device to reduce the window advertisement, the remote servers send less data, thus reducing the congestion. This technique can reduce network congestion by a factor of 40.

Congestion still remains a major problem in the world of computing. It's not expected to disappear anytime soon, due to the fact that resources are limited and, in any situation, where resources are limited congestion is bound to occur. Works are being done on improving the way in which congestion is managed and handled. Algorithms are always modified and changed to enhance efficiency.

References

1. Agarwal, A. (2000). Quality of Service (QoS) in the New Public Network Architecture. *IEEE, 1*(3), 62-69.
2. Cardwell, N., Cheng, Y., Gunn, C. S., Yeganeh, S. H., & Jacobson, V. (2016). Congestion-Based Congestion Control. *acmqueue, 6*(3), 96-99.
3. Evans, J., & Filsfils, C. (2007). *"Deploying IP and MPLS QoS for Multiservice Networks: Theory and Practice.* Morgan Kaufmann.
4. Faisal Shahzad1, M. F., Ullah, S., Siddique, M. A., Khurram, S., & Saher, N. (2015). Improving Queuing System Throughput Using Distributed Mean Value Analysis to Control Network Congestion. *International Journal Communication, Network and System Sciences*.
5. Floyd, S., & Fall, K. (1999). Promoting the Use of End-to-End Congestion Control in the Internet. *IEEE*.
6. Hussain, M. W., Jamwal, S., & Zaman, M. (2015). Congestion Control Techniques in a Computer Network. *International Journal of Computer Applications, 5*(4), 101-104.
7. Jain, R. (1990). Congestion Control in Computer Networks: Issues and Trends. *IEEE*.
8. Jain, R. (1990). Congestion Control in Computer Networks: Issues and Trends. *IEEE*.
9. Jayakumari, B., & Senthilkumar, J. (2015). Priority Based Congestion Control Dynamic Clustering Protocol in Mobile Wireless Sensor Networks. *Hindawi Publishing Corporation, 2*(1), 30-36.
10. Jiao, C., Gao, S., Yang, W., Xia, Y., & Zhu, M. (2014). A Fast Heuristic Algorithm for Minimizing Congestion in the MPLS Networks . *International Journal Communication, Network and System Sciences*.
11. Kharagpur. (2016). Congestion Control. *IIT, 3*(4), 96-100.
12. Mohamed, Sahib, S., Suryana, N., & Hussin., B. (2016). Understanding Network Congestion Effects On Performance. *Journal of Theoretical and Applied Information Technology*.
13. Moses, K. B. (2010). Optimization models for minimizing congestion in Global System for Mobile Communications (GSM) in Nigeria. *Journal Media and Communication Studies, 3*(6), 32-39.
14. Nortel. (2008). Long-Term Evolution (LTE): The vision beyond 3G. *Nortel Networks, 2*(5), 36-42.
15. Ramb, D. (2011, 03 01). *Congestion Control.* Retrieved from Word Press: https://dayaramb.files.wordpress.com/2011/03/computer-networks-notes-pu.pdf
16. Rysavy Research. (2014, March 09). *LTE Congestion Management.* Retrieved from Rysavy Research:http://www.rysavy.com.
17. Shar, S. A., Nazir, B., & Khan, I. A. (2016). Comgestiton Control Algorithms in Wireless Sensor Networks. *Journal of King Saud University*.
18. Ullah, S., Shahzad, F., Khurram, S., & Anwer, W. (2014). Improving Network Efficiency by Selecting and Modifying Congestion Control Constraints. *International Journal Communication, Network and System Sciences*.